

## A Prosodic Theory of Nonconcatenative Morphology

Most structuralist accounts of morphological structure overtly or implicitly make a distinction between two formal morphological types. *Concatenative morphology*, which in the more familiar languages appears almost exclusively, involves prefixation or suffixation only. Thus, morphemes are discrete elements linearly concatenated at the right or the left end of the base of the morphological operation. Morphology of this type is subject to analysis by a relatively simple discovery procedure. Given an adequate phonological representation, concatenative morphemes can be recovered by a left-to-right (or right-to-left) parse of words searching for invariant recurrent partial strings, possibly with constant meaning or function (Hockett (1947)).

The other type, *nonconcatenative morphology*, has remained rather more mysterious until now. Generally, in structuralist treatments we find only a list of the residue, those morphological operations that cannot be analyzed by the method of recurrent partials. These include reduplication, infixation, morphologically-governed ablaut, and suprafixation. All of these terms are in common use except the last, which refers to, for example, the variation in tonal pattern of the stem as a mark of verbal aspect inflection in Tiv (McCawley (1970), Goldsmith (1976)). Although nonconcatenative morphology as a whole has received less attention than concatenative, this is not for lack of exemplification. In a number of languages, processes like reduplication are the primary or sole morphological operations.

This residual status accorded nonconcatenative morphology in structuralist analyses extends to generative theories as well. All generative treatments known to me have relied entirely on the extremely rich transformational notation of Chomsky and Halle (1968). What is offered here instead is a new theory of nonconcatenative morphology, one which owes a great deal to Harris's (1941; 1951) notion of long components. It is a prosodic theory in the sense that it uses the devices of autosegmental phonology,

This article is a revised version of portions of chapter 4 of McCarthy (1979). Some of the material in section 3.3 was originally presented at the Fifth North American Conference on Afro-Asiatic Linguistics in 1976. I am grateful for the assistance of Lee Baker, Nick Clements, Morris Halle, Jay Keyser, Paul Kiparsky, Alan Prince, Ellen Woolford, and an anonymous reviewer for *Linguistic Inquiry*.

The system for transcribing Arabic used here has its familiar values, with the following exceptions.  $\text{ʔ}$  and  $\text{ħ}$  are the voiced and voiceless pharyngeal glides, respectively.  $\text{ġ}$  is the voiced velar spirant, and  $\text{j}$  is the voiced alveopalatal affricate. A subscripted dot in  $\text{t}$ ,  $\text{d}$ ,  $\text{s}$ , and  $\text{ʃ}$  indicates pharyngealization, also known as emphasis. Vowel length is represented bimoraically as  $\text{VV}$ .

which are most familiar through studies of tone and other prosody. This theory is justified extensively in this article by an analysis of the formal properties of the system of verbal derivation and aspect and voice inflection in Classical Arabic. A similar treatment of other verbal inflection and of nominal derivation and inflection can be found in McCarthy (1979).

To conclude this introduction, I will map out the overall geography of this article. Section 1 outlines the problem of the Arabic verb and its relevance to a theory of nonconcatenative morphology. Sections 2.1 and 2.2 present and partially justify much of the formal apparatus that is essential to the later analysis. Section 3 contains the analysis of the Arabic verb, with an occasional excursus into related issues in Tiberian Hebrew. Section 4 deals with the question of the form of morphological rules in this model. It also has some particular observations on reduplication and the extension of this treatment to non-Semitic languages. The appendix sketches two earlier analyses of Semitic morphological systems.

### 1. Statement of the Problem

One of the classic linguistic issues is that of providing an account of the nonconcatenative morphological system prevailing in most members of the Semitic language family. Unlike the more familiar basically concatenative morphology of the Indo-European languages, Semitic morphology is pervaded by a wide variety of purely morphological alternations internal to the stem. In Arabic, for instance, there is a clear sense in which the forms in (1) are morphologically related to one another, although they do not share isolable strings of segments in concatenated morphemes:<sup>1</sup>

- (1) a. kataba 'he wrote'
- b. kattaba 'he caused to write'
- c. kaataba 'he corresponded'
- d. takaatabuu 'they kept up a correspondence'
- e. ktataba 'he wrote, copied'
- f. kitaabun 'book (nom.)'
- g. kuttaabun 'Koran school (nom.)'
- h. kitaabatun 'act of writing (nom.)'
- i. maktabun 'office (nom.)'

Even the fairly elaborate paradigm in (1) is far from exhaustive; for instance, it does not

<sup>1</sup> Here and subsequently I abstract away from certain generally accepted phonological processes. Forms with initial clusters, if not preceded by a vowel in the same phonological phrase, receive epenthetic ?V. Intervocalic glottal stop and a following vowel are deleted in some forms. Some other rules apply only with roots of particular phonological types. Except in a few cases I will have nothing to say about these rules below, and I assume that they are formulated essentially as in Brame (1970), perhaps with some occasional notational adjustments for the analysis developed here. Some explicit suggestions about the expression of phonological processes on prosodic morphological representations can be found in Halle and Vergnaud (in preparation).

include inflectional alternations like *kutiba* 'it was written' and *makaatibu* 'offices (nom.)'.

Certain observations about this morphological system, crucial to an understanding of it, date from a very early period. It has long been known that at its basis there are roots of three or four consonants which cluster around a single semantic field, like *ktb* 'write'. Certain changes in these roots, like gemination of the middle radical in (1b), yield derivatives such as causative or agentive. Moreover, some vowel patterns seem to bear consistent meaning, like the difference in stem vocalism between active *kataba* and passive *kutiba*.

In the very earliest studies—the treatments by medieval Arabic and Hebrew grammarians, generally adopted in the work of Western Orientalists—an elaborated morphophonemic theory is complemented by only the most rudimentary analysis of paradigms like (1). This approach is usually a fairly superficial taxonomy, mediated by a notation that simply shows the citation root *fʕl* (Hebrew *pʕl*) 'do', with appropriate stem modifications. So the basic insight of these classical grammarians was to abstract away from the particular root, but with no richer understanding of the formal morphological system than this. So far as I know, there was no general treatment of relations between vowel patterns except as instantiated on a particular root.

The first modern insights into these problems appear in Harris's (1941) analysis of Biblical Hebrew and Chomsky's (1951) grammar of Modern Hebrew, both of which are discussed in some detail in the appendix to this article. The fundamental characteristic of Chomsky's proposal is a rule moving (or intercalating) long component vowel patterns into triconsonantal roots ((61) in the appendix), relying crucially on transformational rule notation and integral subscripts on segments in the structural description. In view of the fact that Chomsky (1951) contains all the notational apparatus later adopted by Chomsky and Halle (1968), it could reasonably be claimed that transformational morphological rules, essentially similar to Chomsky's, form the basis of the analysis of Semitic nonconcatenative morphology within the generative tradition.

A problem closely related to the formal character of morphological rules is the formal character of morphemes, the units that those rules manipulate. Again the standard theory makes a fairly explicit proposal: a morpheme is a string of segments delimited by the symbol "+" which contains no internal "+". A somewhat richer notion of the morpheme is proposed and justified in section 2.1.

Another necessary characteristic of a morphological analysis is a theory of the structure of the lexicon and of lexical entries. The basic view, adopted by Chomsky and Halle (1968), that the lexicon is a list of single morphemes only and that these units are subject to lexical insertion, has been convincingly dismissed by Halle (1973), Jackendoff (1975), and Aronoff (1976). There is no need to repeat these arguments here, so I shall simply take it for granted that the lexicon is composed of words rather than morphemes. Therefore, the processes described here can be seen as applying redundantly rather than generatively, except in the case of neologisms. Nothing of significance in what follows hinges on this assumption, however.

## 2. Formalism

### 2.1. *The Representation of Morphemes*

It is well known that a number of idiosyncratic morphological and phonological properties cluster around words like *permit*, *subsume*, and *submit*, with Latinate prefixes and stems. In the verb form, stress invariably falls on the final syllable in spite of the possibility of further retraction. Certain special assimilation and deletion rules apply at the boundary between the prefix and stem; compare *admit*, *assume*, *attempt*, *appear*, *accept*. Finally, as Aronoff (1976) notes, the types of nominalizations of these forms are determined entirely by the stem morphemes: *submission*, *permission* with *mit* versus *assumption*, *consumption* with *sume*.

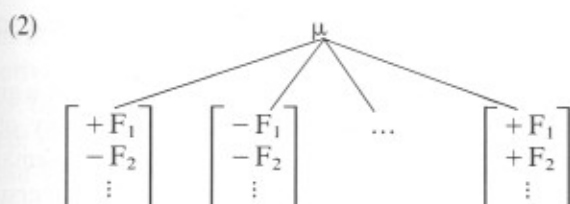
This clustering of properties means that the grammar must be able to recognize words of this type as a class composed of Latinate prefix and stem morphemes. But the exact delineation of morphemes in the representation of these words is an empirical question for which there are two alternative solutions.

One theory, essentially the one followed by Chomsky and Halle (1968), would analyze *permit* as a sequence of two morphemes separated by a boundary but without internal hierarchic or cyclic structure: *per* + *mit*. (It is irrelevant here whether this class has a special boundary like "=" or not.) The boundary allows us to recognize *permit* words as a class—they contain an internal boundary but have no other structure.

In some interesting proposals for the treatment of various junctural phenomena, Rotenberg (1978) and Selkirk (forthcoming) present convincing arguments against the use of boundary symbols in phonological representations. They claim instead that junctural rules actually refer not to boundaries but to hierarchic morphological structure itself, structure that results from deriving one word from another. Notice that here we have an obvious problem for this theory: there is no likely internal hierarchic structure in *permit* class words, but nevertheless several rules must have access to some sort of morphological analysis of them.

There is, however, a third formal possibility. This alternative is implicit in work by Zellig Harris (1951) and essentially involves an extension of his notion of the *long component*. While the boundary solution basically says that morphemes are delimited by symbols in the segmental string, the long component theory claims that the string of segments is uninterrupted, but the morphological analysis is given by another, simultaneous level of representation. Harris's long components were designed to handle discontinuous phenomena—in particular, the Semitic roots that figure prominently in this article. But it requires very little to extend a long component analysis to include segmentally continuous morphemes like *per* or *mit*.

The formal basis of this interpretation is essentially the notation of autosegmental phonology (Goldsmith (1976)). Formally, I will define a morpheme as an ordered string of  $1 \times n$  feature matrices associated autosegmentally with a root node  $\mu$ . This is schematized in (2):



The root node  $\mu$  identifies this string as a particular morpheme. Moreover,  $\mu$  bears all nonphonological information associated with the morpheme, such as rule diacritics, whether it is a root or an affix, and in fact its identity as a morpheme. Note that this is not intended as a substitute for hierarchic structure where that structure is motivated. It does, however, replace all delimitation of morphemes by boundary symbols like "+". A similar proposal, though not cast in autosegmental terms, was made by Pyle (1972).

Any basically concatenative morphological system, like ordinary English morphology, has a very simple translation into this notation. For any  $1 \times n$  feature matrix dominated by  $\mu$ ,  $n$  equals the cardinality of the set of all phonological features, and the daughters of any  $\mu$  form a continuous segmental string. So, for example, *permit* will be represented as in (3):



This sort of representation achieves the desired end. The grammar can refer to *per* and *mit* as separate morphemes with special phonological and morphological properties, without reference to boundary symbols. Because separate nodes  $\mu$  dominate *per* and *mit*, they are necessarily interpreted as distinct morphemes. Clearly, this proposal will trivially extend to the rest of English morphology as well.

A number of arguments can be developed in support of this position. The first type consists essentially of formal arguments, presented in some detail by Pyle (1972). The second type, given here, consists of actual cases where the  $\mu$ -notation is richer than the boundary notation in ways that are essential to the expression of linguistic generalizations.

What is perhaps the most compelling argument for this characterization of the morpheme is the basic organization of the Arabic (and Semitic) lexicon around the consonantal root. All verb forms of Arabic can be partitioned into fifteen derivational classes, which I will refer to by the Hebrew term *binyānīm* (singular *binyān*). I will deal with the formal properties of the binyanim in detail below. What we will be concerned with here is the derivational source of the various binyanim—what other forms in the language they appear to be most closely related to and derived from. This question is very difficult to answer for the first Arabic binyan. It is probably never derived from a verb of some other binyan, but it is usually impossible to say whether some nouns are

derived from this binyan or this binyan from the nouns. Consequently, I will not discuss the source of the first binyan further in this section.

However, there is often clear evidence of a particular derivational source for a given verb of some other binyan. This sort of evidence includes the absence of any other binyanim (including the first) formed on a particular root, as well as specific semantic relationships to related nouns or verbs. It is this sort of evidence that is uncontroversially reflected in the following generalizations.

The forms in most binyanim, except the first, are derived from other binyanim of the same root or from nouns of the same root. For instance, some representative derivational relationships are exemplified in (4):

(4)	<i>Derived Form</i>	<i>Derivational Source</i>
a.	Second Binyan	First Binyan
	ʕallam 'teach'	ʕalim 'know'
	kaððab 'consider a liar'	kaðab 'lie'
		Noun
	marraḍ 'nurse'	mariḍ 'sick'
	kabbar 'say battle-cry'	ʔalaahu ʔakbar 'Allah is great'
b.	Third Binyan	First Binyan
	kaatab 'correspond'	katab 'write'
		Fourth Binyan
	raasal 'correspond'	ʔarsal 'dispatch'
		Noun
	saafar 'travel'	safar 'a journey'
c.	Fourth Binyan	First Binyan
	ʔajlas 'seat'	jalas 'sit'
	ʔaʔkal 'feed'	ʔakal 'eat'
		Noun
	ʔašʔam 'go to Syria'	šaʔm 'Syria'
d.	Tenth Binyan	First Binyan
	stawjab 'consider necessary for oneself'	wajab 'be necessary'
		Fourth Binyan
	staslam 'surrender oneself'	ʔaslam 'surrender'
		Noun
	stawzar 'appoint as vizier'	waziir 'vizier'

Several interesting properties of the binyanim emerge from (4). First, it is clear that these four derived binyanim allow both nominal and verbal derivational sources for the



forms of different roots. In the examples given, the first and fourth binyanim both occur as derivational sources, as do a number of different noun patterns. The second property is that there is no relationship between the form of the source and the form of the output except for the root consonants. Therefore, a fourth binyan verb could come from a first binyan verb *CaCaC* or from a noun of the pattern, say, *CaCC*. Every property of the source except its root is ignored in the form of the derived binyan. This striking fact is perhaps the most interesting characteristic of the distinctive Semitic root and pattern morphology.

Formally, this means that whatever sort of rule relates a derived verb to its source, that rule will have to ignore the formal characteristics of the source except for the root. It will have to be able to isolate the root from the vowel quality and from the canonical distribution of consonants and vowels. Under the theory proposed here, the solution to this problem is almost trivial: the root can be isolated by any rule as the morpheme marked  $\mu$ . Without this notation in the theory, the derivational relationships like [root]

those in (4) which are richly attested throughout the language would be entirely incoherent.

Another argument which supports the notion that the root consonantism is a single unit at some level of representation comes from a language game of Bedouin Hijazi Arabic, a fairly conservative modern Arabic dialect described by al-Mozainy (in preparation). In this game, the consonants of the root may be freely permuted into any order, though nonroot consonants and the canonical pattern of the form remain unchanged. Vowel quality, which is subject to regular phonological effects under the influence of neighboring consonants, varies correspondingly. For example, the possible permutations of *difaʕna* 'we pushed' from the root *dfʕ* appear in (5):

- (5) a. *daʕafna*
- b. *fidaʕna*
- c. *ʕadafna*
- d. *faʕadna*
- e. *ʕafadna*

These permutations can apparently be performed and decoded with some fluency. They clearly demand that the grammar treat the discontinuous string of root consonants as a unit, as is ensured by the  $\mu$ -notation.

Still another consideration lies in the realm of morpheme structure constraints. The Semitic root is subject to a number of rules governing the cooccurrence of consonants within it, a fact originally noted by the classical grammarians. For instance, Greenberg (1978) observes that, with a single exception, no root of a verb contains both *f* and *h*, the voiced and voiceless pharyngeal glides, respectively. Similar distributions hold for other points of articulation, though no such constraints apply to consonants outside the root. The conclusion must be that morpheme structure in Arabic refers to the root

specifically, despite the fact that it is a discontinuous morpheme. Similarly, the vocalism—what I call the *vowel melody*—is not freely distributed among the vowels. For example, it is a fact that no Classical Arabic word (with the possible exception of some loans) has the vocalism *i-u*, nor does any verb have a melody that begins with *i*. Generalizations of this sort cannot be expressed without access to a notation like  $\mu$  in the formulation of the morpheme structure constraints of Arabic.

There is another class of data that is richly attested in Arabic and other Semitic languages. In the standard phonological theory, phonological rules that are restricted to some morpheme or morpheme class must refer to  $+$ -boundary and perhaps also to some set of morphological diacritic features. In a nonconcatenative system,  $+$ -boundary is clearly unavailable, so such rules could not be formulated. I present three cases of this sort below in support of the  $\mu$ -notation. These rules must, however, be taken as preliminary, since they would necessarily be rewritten in the light of the more elaborated analysis of Arabic (and Semitic) morphology in subsequent sections of this article. This consideration does not affect the argument.

The first case is an assimilation rule peculiar to the eighth binyan of the Arabic verb. One characteristic of this derivational class is a *t*-infix between the first and second consonants of the root: /frq/  $\rightarrow$  f $\bar{t}$ araq 'to part', /ʕrd/  $\rightarrow$  ʕ $\bar{t}$ araq 'to place something before one'. But in verbs whose first root consonant is *w* or *y*, the high glides, we find initial geminate *t* in the eighth binyan: /wʕd/  $\rightarrow$  t $\bar{t}$ aʕad 'to receive a promise', /ysr/  $\rightarrow$  t $\bar{t}$ asar 'to play with a dreydl'. This assimilation process is demonstrably unique to precisely this set of morphological circumstances. A root-initial high glide does not assimilate to a following *t* which is part of the same root rather than the eighth binyan infix: /wtd/  $\rightarrow$  ʔawtaad 'tent pegs', /ytm/  $\rightarrow$  yaytim 'to be an orphan'. Assimilation also fails to apply in roots whose third consonant is *w* or *y* when followed by an agreement desinence such as *ta*: /ǧzw/  $\rightarrow$  ǧazawta 'you (m. sg.) made a raid', /rmy/  $\rightarrow$  ramayta 'you (m. sg.) threw'.

The upshot of these facts is that, to apply the assimilation rule correctly, the grammar must be able to identify the *t*-infix of the eighth binyan exclusively. Under a boundary-based theory, though, there is no way to locate an infix as distinct from the unit that contains it. Infixes are not delimited by  $+$ -boundary—this is an incoherent and ad hoc suggestion that would lead to such absurdities as a morpheme apparently composed solely of the first root consonant, preceding the infix:  $+w+t+aʕad$ .

With the  $\mu$ -notation, this rule can be formulated as (6), where the *t*-infix is characterized as a reflexive morpheme:

$$(6) \begin{bmatrix} - \text{cons} \\ - \text{syll} \\ + \text{high} \end{bmatrix} \rightarrow t / \begin{array}{c} \text{---} t \\ | \\ \mu \\ [\text{reflexive}] \end{array}$$



There is, then, no logical or empirical problem with this particular case of morpheme discontinuity, even though this rule could not be expressed in a boundary-based theory.

Another interesting illustration of the necessity of the  $\mu$ -notation arises in the Akkadian reflex of this binyan, as well as in the Hebrew one. Akkadian also has a *t*-infix in the so-called Gt and Gtn (passive and iterative) verbal classes: /mḥs/ → *miḥas* 'to be struck (Gt)', *mitaḥḥas* 'to strike repeatedly (Gtn)'. But in forms where the first root consonant is a coronal spirant, we find that the spirant and the *t* exchange positions by a metathesis rule: /šbt/ → *šitbutum* → *tišbutum* 'to seize one another', /zqr/ → *zitqurum* → *tižqurum* 'to be elevated'. This metathesis proceeds only across an intervening vowel; thus, *ištabbat* 'he will seize' remains unchanged.

Again, it can be shown that this rule is restricted to a particular conjunction of morphological circumstances that require us to be able to identify the *t*-infix. In the notation proposed here, this rule is formulated as (7):

$$\begin{array}{ccccc}
 (7) & \begin{array}{c} \text{C} \\ \left[ \begin{array}{l} + \text{cor} \\ - \text{son} \\ + \text{cont} \end{array} \right] \\ 1 \end{array} & \text{V} & \begin{array}{c} t \\ | \\ \mu \\ \left[ \left\{ \begin{array}{l} \text{passive} \\ \text{iterative} \end{array} \right\} \right] \\ 3 \end{array} & \rightarrow & 3 \ 2 \ 1
 \end{array}$$

Another rule of Akkadian also provides support for recognizing the root as a discontinuous constituent. The nominal prefix *ma* is dissimilated to *na* in any form containing a labial root consonant: *naphar* 'totality', *neereb* 'entrance', *narkabt* 'chariot'. Only elements of the consonantal root suffice to trigger this dissimilation; it fails before a labial stem vowel (*mazuukt* 'mortar') or a labial desinential consonant (*meriit-um* 'pasture'). Therefore, this rule must refer directly to the nonconcatenative root morphemes of Akkadian:

$$\begin{array}{ccc}
 (8) & \begin{array}{c} \text{ma} \rightarrow \text{na} / \text{---} \\ \downarrow \\ \mu \end{array} & \begin{array}{c} \text{X[+labial]} \\ | \\ \mu \\ \text{[root]} \end{array}
 \end{array}$$

As in the Arabic derivational relationships, language game, and morpheme structure constraints, the grammar must have access here to the root as a string-discontinuous constituent.

In section 3, I will develop some further rules of this sort, and we will see reference to discontinuous morphemes as the basis of the analysis of Arabic word formation. The fact that it allows us to deal with these morphemes and their complex interrelations is the strongest confirmation offered for the  $\mu$ -notation.

## 2.2. Theoretical Framework

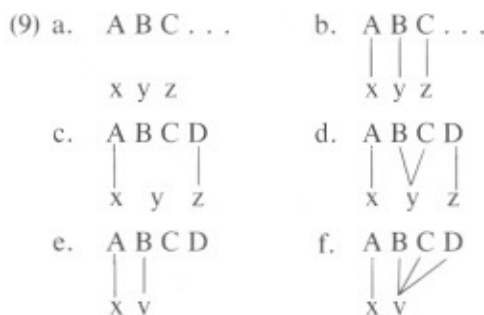
The foundation of the analysis presented here is the theory of autosegmental phonology as described by Clements and Ford (1979). I will assume some familiarity with this theory, and I will outline briefly only those points where it differs from the more familiar proposals of Goldsmith (1976) in ways relevant to this analysis.

The universal conventions for association are cast in terms of the mapping of melodic elements (units on an autosegmental tier) onto melody-bearing elements (units on the segmental tier). There are three such conventions, illustrated schematically by the association of lower case melodic elements with upper case melody-bearing elements in (9).

i. If there are several unassociated melodic elements and several unassociated melody-bearing elements, the former are associated one-to-one from left to right with the latter. This transforms a representation like (9a) into the one in (9b).

ii. If, after application of the first convention, there remain one unassociated melodic element and one or more unassociated melody-bearing elements, the former is associated with all of the latter. This transforms (9c) into (9d).

iii. If all melodic elements are associated and if there are one or more unassociated melody-bearing elements, all of the latter are assigned the melody associated with the melody-bearing element on their immediate left if possible. This principle, which has the effect of automatic spreading, will alter (9e) to (9f).



Contrary to earlier versions of this theory, however, no provision is made for automatic association of an unassociated melodic element with a melody-bearing element that already has an association. Therefore, the representation in (10) is well-formed in this new model:



Only by a language-particular rule can the floating melodic element *z* be anchored to a melody-bearing element. If *z* remains unassociated throughout the derivation, then it receives no phonetic realization or, equivalently, is deleted in the surface representation.

The ordinary case in nontonal autosegmental systems like the one to be developed for Arabic is that floating melodic elements like  $z$  in (10) are never anchored. I will refer to this characteristic informally as the *prohibition against many-to-one associations*. It is thereby ensured that segments with multiple specifications for point and manner of articulation features do not arise in the usual course of derivations.

In a few other respects, however, I will go beyond the theoretical apparatus in the cited literature. The chief difference lies in the somewhat richer notion of *autosegmental tier* presupposed here. It has been assumed that the autosegmentalization of some feature or bundle of features defines a single tier on which all and only those features are represented. I will claim instead that each language has the option of restricting every tier to autosegments which are members of a particular morpheme or morpheme class. Since a morpheme, as we have seen, is a set of feature matrices dominated by a single node  $\mu$ , we can say that a morphologically defined tier contains all and only the feature bundles that are daughters of a single  $\mu$ . In this way, as we will see, consonantal roots and vocalic melodies in Arabic, although they contain bundles of the same distinctive features, can nevertheless be represented on separate autosegmental tiers. This ensures that the association conventions for melodies can operate independently on these two tiers. Association of autosegments from different tiers to the same segments will be subject to the natural restriction that no segment receive multiple associations for the same nontonal feature. This is, in a sense, a generalization of the prohibition against many-to-one associations.

It should be noted that the original definition of an autosegmental tier is not supplanted in this model. Only one set of phonological features can appear in any column of a particular tier. Moreover, different tiers cannot contain the same features unless those tiers represent different morphemes, and then only if a particular grammar stipulates that the tiers are morphologically determined. Finally, as in the familiar version of autosegmental theory, each autosegmental tier will designate a natural class on the segmental tier as its set of tone-bearing elements, the units with which it is to be associated.

The other addition to autosegmental theory followed here is a revised version of Leben's (1973) *Obligatory Contour Principle*. Leben's principle says that no tonal melody can contain adjacent identical elements. Thus, a melody HHL is automatically simplified to HL, but HLH remains unchanged. The revisions of this principle involve two points. First, in the light of autosegmental representation of melodies, I will state it as a constraint on contiguous elements in any autosegmental tier rather than on the tonal melodies of Leben's theory. Second, in view of Goldsmith's (1976) demonstration that such a constraint alone is too strong for some aspects of Tiv conjugation, I will make the weaker claim that it operates as part of the evaluation metric rather than as an absolute universal principle. This seems to accord with the facts of Arabic, as we shall see, since the Obligatory Contour Principle is observed in all forms except for a few loan nouns.

Since we will have occasion to refer to this principle later, let us formulate it now:

(11) *Obligatory Contour Principle* (revised)

A grammar is less highly valued to the extent that it contains representations in which there are adjacent identical elements on any autosegmental tier.

This completes the summary of the theoretical apparatus needed in this analysis.

## 3. The Classical Arabic Verb System

3.1. *Outline*

The verb system of the trilateral root is based on fifteen derivational categories and that of the quadrilateral root on four—these are the binyanim mentioned above. Although the Arabists' nomenclature refers to them as conjugations, they are in no way similar to the more familiar conjugational types of Latin or Greek. In fact, each binyan is inflected in almost the same way as all the other binyanim. What they differ in is the arrangement of root consonantism with respect to characteristic affixes and vowel positions.

The first binyan is a possible category for nearly all roots that can appear as verbs. It is relatively unmarked morphologically, at least in the finite forms, and it has no special semantic properties. This is roughly true as well for the first quadrilateral binyan, QI. But the others, the derived binyanim, generally involve some special modification of the meaning of a related noun or verb or of the basic meaning of the root. So, for instance, the third trilateral binyan is usually reciprocal, while the sixth is usually the reflexive or effective of the reciprocal. It is, in general, an idiosyncratic property of any root whether it can appear in a particular binyan. Nevertheless, neologisms abound, loanwords are easily incorporated into the system, and speakers of Modern Standard Arabic report a reasonable facility in extending a root to other binyanim and interpreting the result.

Subject to these lexical idiosyncrasies, the binyanim cross-classify the roots morphologically and semantically, where the root supplies the basic meaning and the binyan (except for the first binyan) supplies some modification of this meaning or of the verbal diathesis. The meaning of any verb is not a composition of the meaning of root and binyan, but there is a reasonable amount of predictability. For instance, the root *ktb* expresses a notion like 'write'. This root occurs in eight binyanim, reflected by the following uninflected forms of the perfective active:

(12) *Binyan*

I	katab	'write'
II	kattab	'cause to write'
III	kaatab	'correspond'
IV	?aktab	'cause to write'
VI	takaatab	'write to each other'
VII	nkatab	'subscribe'
VIII	ktatab	'write, be registered'
X	staktab	'write, make write'

Table 1

<i>Perfective</i>		<i>Imperfective</i>		<i>Participle</i>	
Active	Passive	Active	Passive	Active	Passive
<b>Triliteral Roots</b>					
I katab	kutib	aktub	uktab	kaatib	maktuub
II kattab	kuttib	ukattib	ukattab	mukattib	mukattab
III kaatab	kuutib	ukaatib	ukaatab	mukaatib	mukaatab
IV ʔaktab	ʔuktib	uʔaktib	uʔaktab	muʔaktib	muʔaktab
V takattab	tukuttib	atakattab	utakattab	mutakattib	mutakattab
VI takaatab	tukuutib	atakaatab	utakaatab	mutakaatib	mutakaatab
VII nkatab	nkutib	ankatib	unkatab	munkatib	munkatab
VIII ktatab	ktutib	aktatib	uktatab	muktatib	muktatab
IX ktabab		aktabib		muktabib	
X staktab	stuktib	astaktib	ustaktab	mustaktib	mustaktab
XI ktaabab		aktaabib		muktaabib	
XII ktawtab		aktawtib		muktawtib	
XIII ktawwab		aktawwib		muktawwib	
XIV ktabab		aktanbib		muktanbib	
XV ktabay		aktanbiy		muktanbiy	
<b>Quadriliteral Roots</b>					
QI dahraj	duhrij	udahrij	udahraj	mudahrij	mudahraj
QII tadahraj	tuduhrij	atadahraj	utadahraj	mutadahrij	mutadahraj
QIII dhanraj	dhunrij	adhanrij	udhanraj	mudhanrij	mudhanraj
QIV dharjaj	dhurjij	adharjij	udharjaj	mudharjij	mudharjaj

The characteristic morphology of these forms—permutations of vowels and consonants and so on—will emerge shortly.

Besides the binyanim, this analysis will attempt to account for several other properties of the Arabic verb system. There is a basic division into two aspects, perfective and imperfective. Voice is active or passive, with slightly different morphology for voice in the two aspects. For reasons of brevity, no account will be given here of verbal agreement, nor of mood or verbal clitics. (A full discussion of agreement can be found in McCarthy (1979).) In all other respects, though, this analysis strives for a complete account of the formal characteristics of Arabic verbal morphology.<sup>2</sup>

Table 1, which will serve as the basis for much of the analysis, displays the citation

<sup>2</sup> Since the forms in table 1 involve a considerable degree of abstraction, a little caution is in order. First, the roots *ktb* and *dhrj* may happen not to occur in particular binyanim, although formally equivalent roots do. Thus, V *takattab* is not a real verb, although V *takassab* 'to earn' is one. In the first binyan, different roots belonging to different ablaut classes, treated in section 3.4, yield different vocalism from that of *ktb* in the perfective and imperfective active. Finally, the forms in table 1 are all stems, so they do not show mood, agreement, or case, gender, or number marking, which are not dealt with in this article.

triliteral root *ktb* in all fifteen triliteral binyanim and the root *dhrj* 'roll' in the four quadriliteral binyanim, organized as in any traditional grammar. Here and later, each triliteral binyan is referred to by the appropriate Roman numeral of the Orientalists' system, while the quadrilaterals have a prefixed Q. The major aspect and voice inflections of the finite and nonfinite verb forms head the columns. Gaps in the passive inflections indicate binyanim that are regularly intransitive and stative, and therefore not susceptible of passivization for nonmorphological reasons.

### 3.2. Consonantism

Let us consider the differences among the various binyanim in just the perfective active, where the vowel characteristics are most muted. As a kind of minimal, barely adequate account of these differences, we would have to answer the following questions:

- (A) How are the consonants arranged with respect to the vowels—what is the canonical syllable pattern of the form?
- (B) How are prefixes and infixes like *t* or *n* arranged among the root consonants?
- (C) How are the root consonants arranged with respect to each other? That is, where do clusters or geminates occur?
- (D) How is one binyan related to or derived from another?

This last question, which would take us rather far from the purely formal issues here into the function of the various binyanim, is dealt with in McCarthy (1979).

On the other hand, a preliminary answer to the entirely formal question (A) is much easier to get. The inventory of canonical patterns in the perfective of the triliteral binyanim is listed in (13), where C denotes any [–syll] segment, including consonants and glides:

- |               |            |
|---------------|------------|
| (13) a. CVCVC |            |
| b. CVCCVC     | f. CCVCVC  |
| c. CVVCVC     | g. CCVCCVC |
| d. CVCVCCVC   | h. CCVVCVC |
| e. CVCVVCVC   |            |

Certain obvious regularities appear in (13) which the grammar ought to take account of. First, the stems of all binyanim invariably end in closed syllables (CVC). Second, there is no binyan with a sequence of two light syllables like CVCVCVC. Third, no binyan contains a light syllable after a heavy syllable like CVCCVCVC. Fourth, no binyan which begins with a consonant cluster is three or more syllables long overall.

To minimally express these regularities, the grammar should contain some sort of rules regulating the canonical distribution of consonants and vowels in the binyanim. The template (14a) generates all and only the observed canonical patterns of the binyanim in (13), provided that we exclude sequences of two light syllables by the rule (14b):<sup>3</sup>

<sup>3</sup> The template schema in (14a) could conceivably be analyzed further. There is some evidence for a [CV] template prefix in V *takatab* and VI *takaatab*, which are fairly regularly derived from II *katab* and III *kaatab*. A fuller treatment of this observation would necessarily take us rather far afield into the interrelationships of the various binyanim, however.

I am indebted to Morris Halle and Alan Prince for their suggestions about the proper formulation of (14).



- (14) a.  $[(\begin{Bmatrix} C \\ CV \end{Bmatrix}) CV([+seg])CVC]$   
 b.  $V \rightarrow \phi / [CVC \text{ \_\_\_\_ } CVC]$

The notation  $[+seg]$  indicates an element that may be either a consonant or a vowel, depending on the binyan. The first expansion of the curly brackets in the template allows all and only the patterns in the first column of (13) and the second expansion allows all the patterns in the second column of (13), plus the illicit  $[CVCVCVC]$ . The rule (14b), which eliminates this last possibility, can be thought of as applying redundantly to the set of templates generated by (14a). We will, however, see evidence of alternations supporting (14b) in section 3.4.

Since it specifies the overall prosody, or syllable pattern, of a form, I will refer to the schema in (14a) as a *prosodic template*, although the term *CV-skeleton* adopted by Halle and Vergnaud (1980) may be more evocative. Prosodic templates are composed solely of the features [segmental] and [syllabic], the appropriate values of these features being abbreviated by C and V. Each binyan characteristically stipulates one expansion of this schema, choosing optional elements and consonantal or vocalic values for those units marked only as  $[+seg]$ . Therefore, we can say that one aspect of the specification of any given binyan in the grammar is an indication of the prosodic template of that binyan chosen from the set abbreviated by (14a). The stem patterns of Arabic verbs must be selected from this restricted group of possibilities and no others.

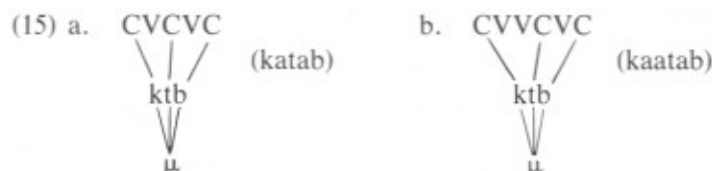
It is proposed here that the prosodic template corresponds to the segmental level in more familiar autosegmental analyses. Thus, the segmental level will contain only the features [segmental] and [syllabic], and all other features will be autosegmental. This leads to a straightforward analysis of the problem in (B) and (C) of arranging root and affixal consonantism with respect to the C-slots of the prosodic template.

Let us assume that the Arabic trilateral root is represented formally as a melody on a single, morphologically defined autosegmental tier which takes as its melody-bearing elements the  $[-syllabic]$  positions of the prosodic template. This melody contains three melodic elements composed of all features except [segmental] and [syllabic]. In this way, the root tier will provide all the information needed to distinguish consonants from one another by point and manner of articulation. Rather than list all these features, I will informally abbreviate them as *ktb* and so on, although strictly speaking *k*, *t*, and *b* in this sense are not ordinary segments but rather archisegments unspecified for [segmental] and [syllabic]. Similarly, affixes like *n* or *t* will appear on separate autosegmental tiers. These affixal tiers involve the same distinctive features as the root tier, but they are distinct because the tiers are morphologically defined, as described in section 2.2. The significance of this move will emerge shortly.

The problem now is to account for the mode of association between the melody-bearing  $[-syllabic]$  slots of the prosodic template and the autosegments of the various consonantal tiers. We will begin by considering some cases in detail.

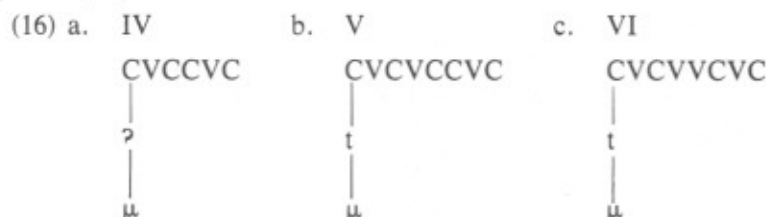
For the templates (13a) and (13c), the problem of association is trivial. A triconsonantal root will, by the first universal convention in section 2.2, associate from left

to right, resulting in a simple one-to-one association with the three C-slots of the template. This result appears in (15):



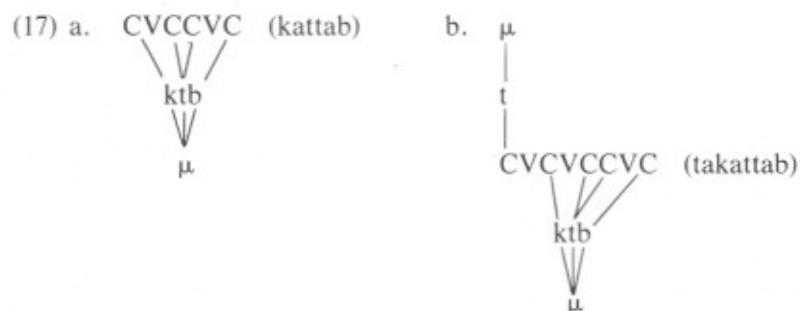
Consequently, these two cases involve no complications in root-to-prosodic template association.

Now let us examine the forms that have an affix—a consonant which is demonstrably not part of the root—mapped onto one of the slots in (13). Each of the binyanim IV, V, and VI has additional morphological material, either *ʔ* or *t*. For these binyanim, it suffices to associate this affixal material with the initial consonant in the template, yielding the outputs in (16):



At this stage, the remaining C-slots in (16a) and (16c) can be unambiguously associated with the elements on the root tier from left to right.

But a problem remains in treating forms like the second and fifth binyanim. Even after affixation as in (16b), the templates of these two categories have four slots to accommodate just three root consonants. What actually occurs is gemination of the middle root consonant, in effect expanding the triliteral root to fit four consonantal slots. I interpret this gemination formally as a one-to-many mapping of the single middle root consonant onto two slots in the prosodic template:



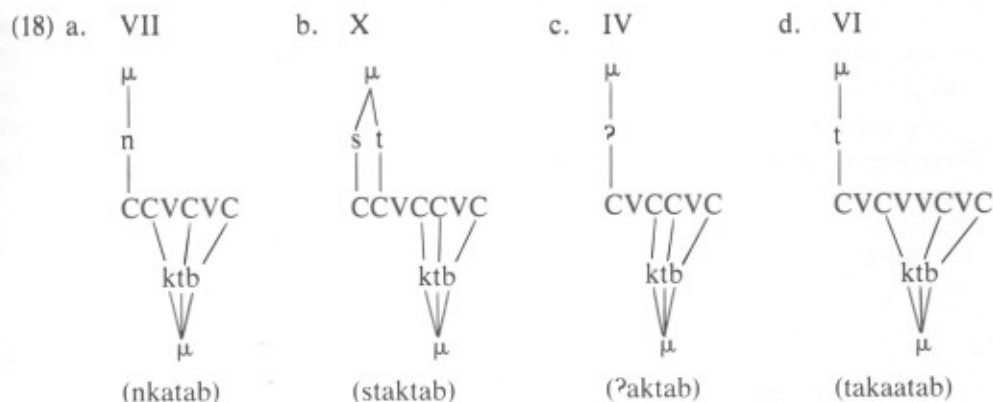
Note here that the affix *t* is represented on a separate tier from the root *ktb* as a consequence of saying that the tiers are morphologically defined.

The structures in (17) represent the presumed output of the processes forming the second and fifth binyanim. The question we have to answer is how the grammar produces these particular associations of root consonants with slots, and not ones where, say, the final root consonant is in a one-to-many relationship. We must consider the other binyanim before we can respond to this.

Template (13f) appears in the seventh binyan with an *n*-prefix, in the eighth with a *t*-infix after the first radical, and in the ninth with gemination of the final root consonant. (13g) appears in the tenth binyan with prefixed *st*, while (13h) appears in the eleventh binyan also with a geminated final radical.

First the affixal material must be dealt with. It suffices to say that *n*, like the 2-affix, is associated with the first consonant of the template. This property—association of the affix with the first consonantal slot of the prosodic template—is observed consistently by the fifth and sixth binyanim for the affix *t*, by the fourth binyan for the affix ʔ, and by the seventh binyan for the affix *n*. This holds as well for the complex affix *st* of the tenth binyan, since it lodges on the first two consonantal slots of the prosodic template. If we suppose that material on an affixal tier is applied to the prosodic template before material on any root tier, then, as an automatic consequence of this ordering and of the conventional left-to-right association, affixes will without further stipulation appear on the leftmost consonantal slots of the prosodic template.

The output of left-to-right association on both tiers is shown in (18):

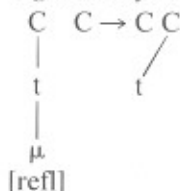


First the material on the affixal tier is mapped onto the template, selecting the leftmost slots. The remaining slots receive a left-to-right mapping of the root tier melody, subject, of course, to the condition that there be no many-to-one associations with the segmental level.

There is one systematic exception to this pattern of affixation. The reflexive morpheme *t*, which is prefixed in the fifth (*takattab*) and sixth (*takaatab*) binyanim, is infixed in the eighth binyan. That is, it is associated with the second consonant slot of the prosodic template and not the first. Here we can say that left-to-right association applies

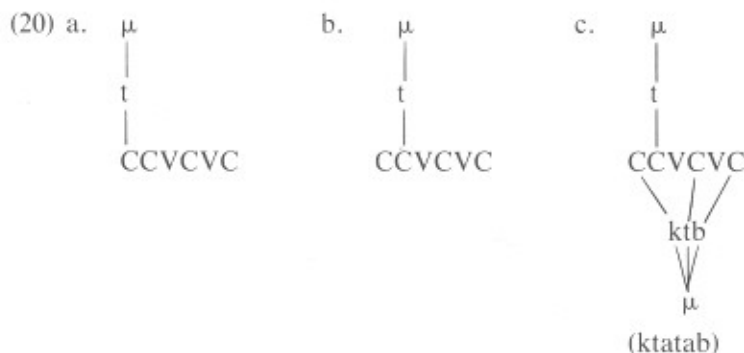
in its usual fashion, but that a subsequent rule, restricted to this affix and a particular prosodic template, flops the association of the affix from the first to the second slot of the template. Rules of this type are fairly common in tonal systems (Goldsmith (1976)). Formally, the Arabic rule reads as follows:

(19) *Eighth Binyan Flop*



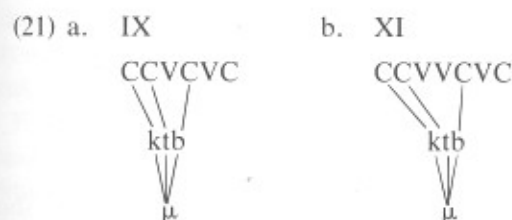
By moving the association of *t* to the right, this flop rule correctly makes *t* an infix in the eighth binyan only. The morphological feature [reflexive] identifies this particular morpheme with the phonological shape *t*, distinguishing it from the *t* of, say, the agreement system. The requirement that the two consonants of the prosodic template be adjacent ensures that the same reflexive *t* will not flop in the fifth and sixth binyanim, where the consonants are separated by an intervening vowel.

At this point we can see the effect of the notion of morphologically defined autosegmental tiers. The affix *t* is on a separate tier from the root *ktb*, since they are different morphemes. The affix is first associated with the initial *C* of the template [CCVCVC] (20a), and then rule (19) shifts its association to the second slot (20b). At that point, mapping of autosegments from the root tier is effected, in accordance with the left-to-right association convention. The slot with which affixal *t* is associated is already filled, and the prohibition against many-to-one associations will not allow it to be doubly filled. Therefore, the root must associate with the other available slots, yielding the representation in (20c):

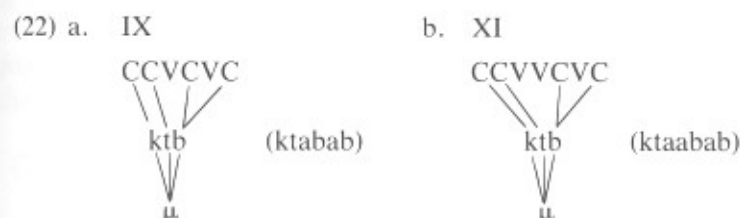


The morpheme *ktb* in (20c) does not contain the affix *t* in the strict sense; rather, they are distinct representations on separate tiers which have contact with each other only by way of association with the same prosodic template.

Further consequences of left-to-right association of roots with prosodic templates arise in the ninth and eleventh binyanim. These are formed on the templates (13f) and (13h). Simple association yields (21a,b):

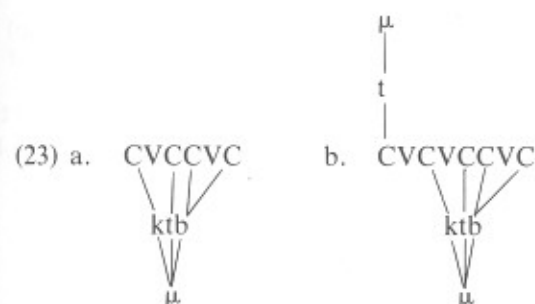


The unassociated final C-slot is now associated with the melodic element bound to the C-slot on its left, in this case *b*. This is a consequence of the third universal convention described in section 2.2. This convention yields the representations in (22a,b):



Consequently, this sort of automatic spreading is sufficient to generate the gemination displayed by these two binyanim without any additional stipulations.

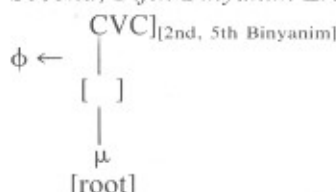
In a similar way, we can derive the gemination of the medial radical in the second and fifth binyanim, *kattab* and *takattab*. Association of the affix *t* and left-to-right association of the root consonantism yield structures like those in (23a,b):



Then a new, morphologically restricted rule erases the association of the final root consonant with the medial C. This now empty C is conventionally subject to reassociation with the autosegment associated with the consonant slot on its left: in this case, the medial radical *t*. This is the same universal mechanism of automatic spreading responsible for the ninth and eleventh binyanim, though in this case it presupposes prior application

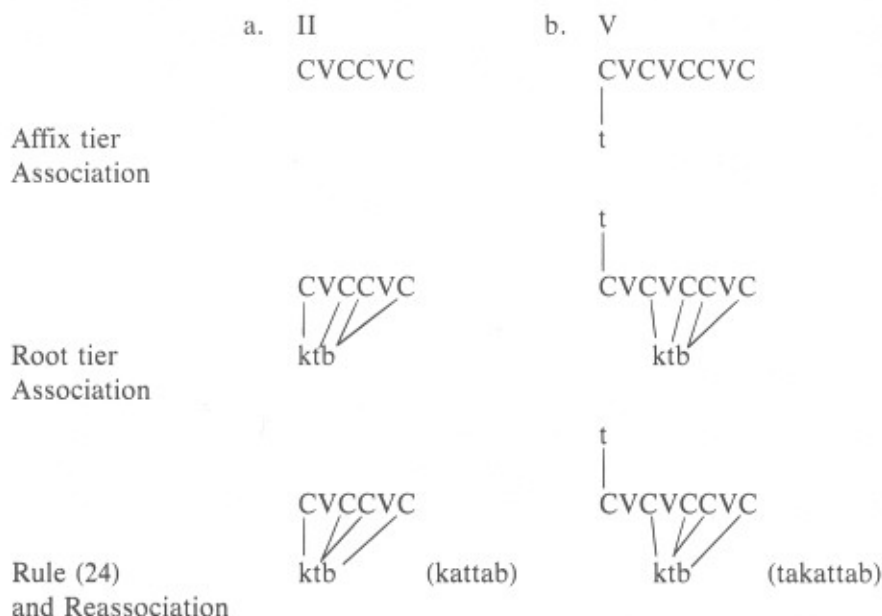
of rule (24):<sup>4</sup>

(24) *Second, Fifth Binyanim Erasure*



So a partial derivation of the perfective forms of these binyanim will proceed as shown in (25):

(25)



In sum, the basic formal apparatus that is specific to Arabic grammar (rather than being part of the universal theory of autosegmental phonology) that generates the binyanim is the list in (26):

- (26) a. The prosodic template (14a) and rule (14b)  
 b. The affixes  $\text{ʔ}$ ,  $t$ ,  $n$ , and  $st$   
 c. The Flop and Erasure rules (19) and (24)

In addition, the grammar must contain a specification for each binyan of its choice from the vocabulary of prosodic templates and of affixes. For example, the sixth binyan will

<sup>4</sup> Further analysis of the relation between the second and fifth binyanim, as described in note 3, would probably eliminate the need to refer to both of them in the formulation of (24).

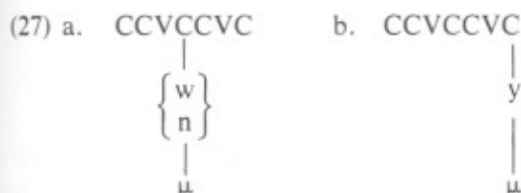


select the template [CVCVVCVC] generated by (14a) and the affix *t*. The only other formal device needed is, obviously, a list of triconsonantal roots.

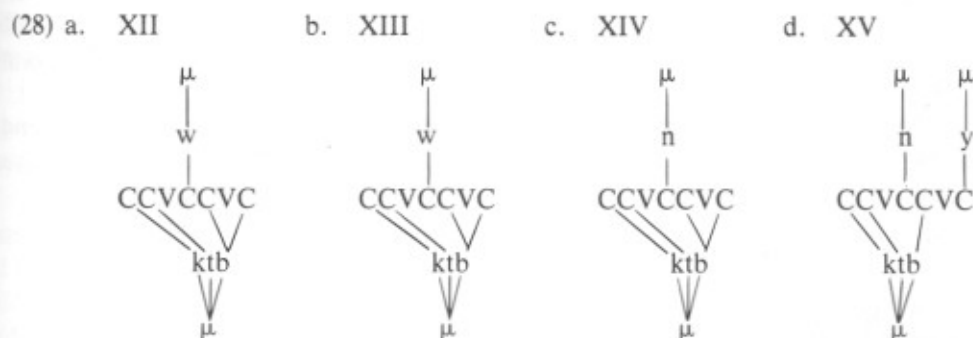
Considering the complexity of the phenomena, it is remarkable that so few stipulated mechanisms are needed to capture a great number of generalizations. Interestingly, this analysis has quite a number of specific empirical consequences other than those already discussed.

First, consider the triliteral binyanim XII–XV. These are indisputably rare; nevertheless, they do occur, they were recognized as binyanim in the classical grammatical tradition, and they usually are fairly transparently related to a verb of the first binyan or perhaps a noun. They are almost always intransitive.

They form a natural class in the prosodic template notation, since all of them are formed on the template [CCVCCVC]. They are also peculiar in having affixal material— infixes *w*, *n*, suffix *y*—that is never prefixed, unlike the reflexive morpheme *t* of the fifth, sixth, and eighth binyanim. Therefore, there seems to be no reason to suppose that a flop rule is operating here, so the additional complication of these very rare conjugations is that rules of association must indicate where the affixes are to be fixed on the prosodic template:

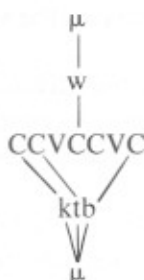


After these two special association rules on the affixal tiers, the usual association conventions apply to the root tier, yielding the intermediate representations in (28):



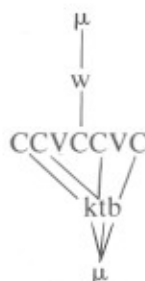
The forms *ktanbab* in (28c) and *ktanbay* in (28d) are correct results and so require no further comment. (28a) and (28b), on the other hand, are subject to a generalized version of the Erasure rule (24), which has the same effect on both forms, yielding (29):

(29)

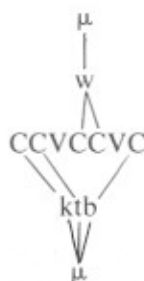


After Erasure, we expect reassociation from the nearest consonant slot on the left—in this case, *w*. But since the root and the infix are representations on separate autosegmental tiers, it is possible to reassociate either from the infixed *w* or from the second root consonant *t* and still produce a well-formed representation. In fact, the twelfth and thirteenth binyanim differ on exactly that point—on whether the infix or the second root consonant is geminated: XII *ktawtab*, XIII *ktawwab*. The final result is the representations in (30a,b):

(30) a. XII



b. XIII



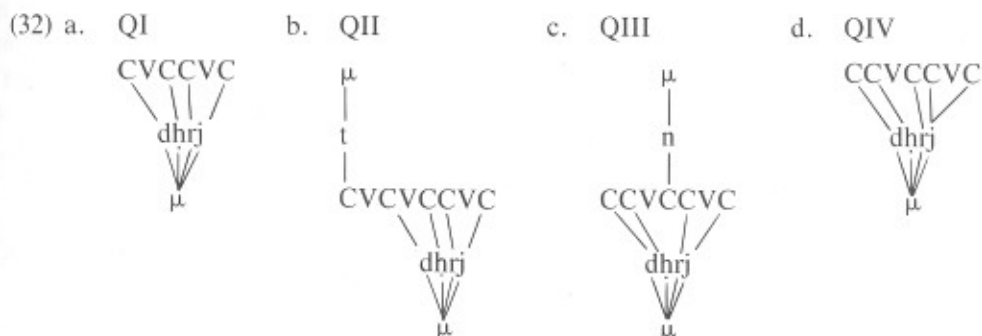
My general conclusion is that these rare binyanim require no more theoretical or grammatical apparatus than the more common binyanim, other than the peculiar affixes in (27). They can be subsumed under basically the same rubrics. The same is more significantly true for the quadriliteral verb forms.

Arabic recognizes four quadriliteral binyanim, the first two fairly common and the last two rather rare. In several respects, we can identify all the quadriliteral binyanim with corresponding triliteral ones. Consider the parallel characteristics in (31):

- |         |          |          |
|---------|----------|----------|
| (31) a. | II       | QI       |
|         | kattab   | daḥraj   |
| b.      | V        | QII      |
|         | takattab | tadaḥraj |
| c.      | XIV      | QIII     |
|         | ktaḥbab  | dḥanraj  |
| d.      | XI       | QIV      |
|         | ktaabab  | dḥarjaj  |

The formal similarities between corresponding triliteral and quadrilateral binyanim are quite clear in terms of the analysis proposed here. In every case, the corresponding forms in both columns are built on the same prosodic template and have the same affixes *t* and *n*. A partial exception to the overall similarity in (31) is (31d), where both forms result from the same prosodic template but with different realizations of the template slot that is designated only as [+seg].<sup>5</sup>

For these reasons, we need not stipulate four other binyanim that are restricted to quadrilateral roots. Rather, it is enough to notate four of the triliteral binyanim as also allowing the application of quadrilateral roots to their templates: binyanim II, V, XIV, and XI (where [+seg] is C). The direct result of left-to-right association of affixes and of the four-consonant root *dhrj* is shown in (32):



The gemination in (32d) is a familiar result of spreading from the left. One question raised by these forms is, If QI and QII are formally instances of the second and fifth triliteral binyanim, why does the Erasure rule (24) not apply in (32a) and (32b)? Since these forms are in the second and fifth binyanim, we would expect erasure of the association between the root consonant *r* and its slot on the template. Actually, we can allow Erasure to apply in quadrilaterals. Its output will be subject to the second universal association convention in section 2.2, immediately undoing the effects of Erasure and yielding (32a) and (32b).

In sum, the whole quadrilateral scheme requires no elaboration of the apparatus and bears clear and demonstrably correct formal relationships to corresponding triliteral binyanim.

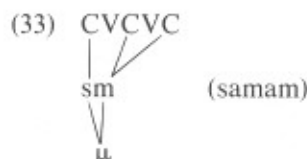
Another empirical consequence of this theory lies in the treatment of so-called geminate roots in Arabic. There is quite a number of roots (perhaps 200) whose second and third radicals are apparently identical: *smm*, *hll*, *mdd*, etc. Greenberg's (1978) statistical study of native vocabulary also found about 20 verb roots with identical first and third radicals: *qlq*, *ndn*. There is also a large number of roots restricted to nouns with identical first and third radicals: *θalaaθ* 'three'. But certainly in Arabic, and reasonably

<sup>5</sup> Further similarities hold at other levels. Although QI is not generally causative like the second triliteral binyan, the other quadrilateral binyanim share some semantic correspondences with trilaterals. The second quadrilateral binyan is, like the fifth triliteral, generally reflexive or resultative. Like their triliteral counterparts, QIII and QIV are usually intransitive and stative.

confidently in the other major Semitic languages, there are no verbal or nominal roots with identical first and second radicals, except for the unique Arabic noun *dadān*, a nursery word for 'plaything'. The grammars also note a unique Arabic root *yyy*, which means 'to write the letter y'.

This asymmetry in distributional restrictions between first and second root position versus other positions has not yet received a satisfactory explanation. Consider two representative roots with identical radicals in the permitted positions, like *qlq* and *smm*. The first, *qlq*, is unremarkable in the autosegmental treatment, and is formally indistinguishable from entirely regular roots like *ktb*. But the second, *smm*, as well as all other geminate roots, must be represented formally as a biliteral root *sm* according to the revised Obligatory Contour Principle presented in section 2.2, in the most highly valued grammar. This holds for each morpheme separately or, strictly speaking, for each morphologically defined autosegmental tier. Consequently, it does not apply to heteromorphic sequences of adjacent identical units. If there were a (traditional) root of the nonoccurring type designated as *ssm*, this root would be formally identical to *smm* because of the operation of the Obligatory Contour Principle. Given this apparatus, the convention of left-to-right association can explain the absence of verbs or nouns like *sasam* versus the existence of *samam*.

Now consider the mapping of the biliteral root onto the prosodic template of the first binyan perfective:



Because mapping is from left to right, only the second radical is geminated by automatic spreading. This gemination has nothing to do with the morphology of any binyan—it depends only on filling up the available slots. Given left-to-right association, though, there is no way, short of additional unmotivated rules, to induce gemination of the first radical, so we will never end up with first binyan verbs like *\*sasam*. This is, in fact, exactly the right result, and it clearly accounts for this tremendous skewing of the Arabic (and Semitic) lexicon.<sup>6</sup>

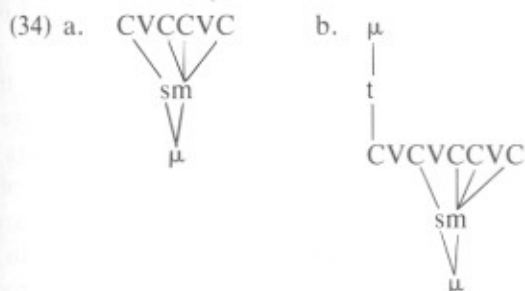
In brief, Arabic allows roots of two, three, and four consonants, all of them subject to the Obligatory Contour Principle. Biconsonantal roots are realized on the surface with gemination of the second consonant as a direct consequence of the universal left-

<sup>6</sup> This analysis of biliteral roots is further confirmed by data from the Bedouin Hijazi Arabic language game described in section 2.1. Under this game, a form with a biliteral root like *ḥall* 'he solved' can be transformed only to *lahḥ* and not to *\*lahḥ* or *\*lahl*. This is exactly what we would expect if the game permutes a biliteral root *hl*, and then this root is mapped onto a [CVCC] template by the association conventions.

I should note that this synchronic analysis is neutral with respect to the diachronic question of whether proto-Semitic had biliteral forms. This historical problem refers to actual surface representations, not to representations on an abstract autosegmental level.

to-right association convention. Note also that the Obligatory Contour Principle excludes quadriliteral roots with adjacent identical autosegments, like hypothetical \**ddrj* or \**drj*. In fact, this is the right result; there are no QI verbs of the type \**dadraj*.

This theory also predicts the occurrence of doubly reduplicated root consonants. The only limitation on such reduplication is the difference between the number of root consonants and the number of empty consonantal slots in the template. Arabic routinely shows double reduplication in the second and fifth binyanim with roots like *sm*: *sammam*, *tasammam*. These are represented formally as follows:



The representations in (34) are subject to the Erasure rule, but its effect is automatically reversed by the application of the third universal association convention in section 2.2. So (34) does, in fact, give the output form of the consonantism.

There is a further result of this analysis of biconsonantal verbal roots. Because of the autosegmental treatment, there is a particular formal characteristic shared by biliteral roots and those trilateral and quadrilateral roots that appear in binyanim with characteristic gemination. In every case, gemination is represented formally as a one-to-many association from the root tier to the prosodic template. This representation does not hold, however, of adjacent identical consonants that come from different morphemes and consequently from different autosegmental tiers, such as root and affix. This makes a difference in the conditioning of a phonological rule of some generality.

The alternations in inflected forms of a biliteral root in (35a) are paralleled by alternations of a trilateral root in the ninth and eleventh binyanim in (35b) and of a quadrilateral root in the QIV binyan in (35c):

- (35) a. *samamtu* 'I poisoned'      *yasmumna* 'they (f.) will poison'  
           *samma* 'he poisoned'      *yasummu* 'he will poison'  
       b. *šfarartu* 'I was yellow'  
           *šfarra* 'he was yellow'  
       c. *šmaʕaltu* 'I hastened'  
           *šmaʕalla* 'he hastened'

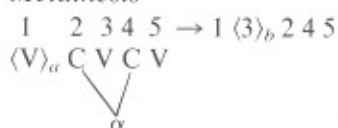
The alternations in (35) reflect a process applied to underlying stems like /*samam*/. If two identical consonants are separated by a short vowel, and if the second of them is also followed by a vowel, then the two consonants are joined into a geminate cluster.

Thus, underlying /samama/ becomes *samma* and underlying /yasmumu/ becomes *yas-ummu*, but /samamtu/ and /yasmumna/ remain unchanged. Similar effects appear in the derived binyanim from nongeminate roots in (35b) and (35c).

What is significant for the theory here is that this process does not apply to identical consonants that do not belong to the same root. Thus, the eighth binyan *ktatab* does not become *\*kattab*, since the first *t* is affixal and the second is radical. The same situation holds for V *yatatabbasu* 'he will pursue' and VI *yatataabas* 'he will succeed', where the second *t* is the first consonant of the root /tbs/. The process also fails with *maqataa* 'they (f. du.) detested', where the first *t* is part of the root *mq*t and the second is an inflectional affix of the feminine.

Although these facts seem to demand some baroque morphological conditions, there is in fact quite a simple solution under the analysis presented here. All cases where the cluster-forming process does apply are those in which the identical consonants are represented by the association of a single consonantal autosegment with two slots of the prosodic template. The process fails to apply when the identical consonants are in different morphemes, and consequently appear on different autosegmental tiers. In this case there is no one-to-many association. Therefore, it suffices to say that the process applies only to template positions that are associated with the same element on the autosegmental tier. If we suppose, following Brame (1970), that the cluster-forming process collapses metathesis and deletion rules, then it can be formulated as (36):

(36) *Metathesis*



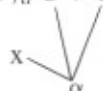
Condition:  $a \supset \sim b$

The angled brackets and the condition distinguish the two cases on the left and on the right in (35a). These aspects of the rule are not under consideration here, and could be reformulated. What is relevant, though, is the fact that both affected consonants must be associated with the same melodic element  $\alpha$ ; it does not suffice that they are merely identical. Metathesis will therefore apply to the geminated root consonants in (36), but it will be unable to apply to the forms cited in the text where the identical consonants are represented on separate autosegmental tiers, since they are in different morphemes.

A few additional facts suggest a slight complication of the Metathesis rule (36). Second and fifth binyan forms from geminate roots do not undergo Metathesis. Thus, /sammama/ and /tasammama/ do not yield *\*samamma* and *\*tasamamma*. Referring to the representation of these forms with double reduplication in (34), we can see a simple way to block the application of Metathesis. We can require that  $\alpha$  have no association lines to the left of the one associated with the C-slot in position 2 of the structural description. Following Kahn (1976), I will notate this as shown in (37):

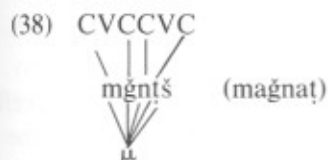


(37) *Metathesis* (revised)
$$1 \quad 2 \ 3 \ 4 \ 5$$

$$\langle V \rangle_a \ C \ V \ C \ V \rightarrow 1 \ \langle 3 \rangle_b \ 2 \ 4 \ 5$$
Condition:  $a \supset \sim b$ 

This addition provides that the first identical consonant will not itself be a geminate.

We have seen in some detail the behavior of biliteral, trilateral, and quadrilateral roots. Quinqueliteral roots, although they are not native to the Semitic languages, do nevertheless occur. Arabic has some quinqueliteral roots that appear in nouns. These are invariably loanwords or, in a few cases, acronyms. There are some examples of denominal verbs derived from these nouns quite transparently. When this happens, the final consonant of the root in the derived verb simply disappears, and the result is a typical quadrilateral verb: *mağnaṭiis* 'magnet', *mağnaṭ* 'to magnetize (QI)'; *qalansuw(at)* 'cap', *taqalnaṣ* 'to wear a cap (QII)'. According to left-to-right association, a root like *mğntš* will associate with the CVCCVC prosodic template as illustrated in (38):



What happens is that the universal association convention leaves *š* stranded at the right without a consonantal slot. It cannot attach to any of the already filled slots because of the usual prohibition against many-to-one association in nontonal systems. Consequently, final *š* remains unattached and receives no phonetic realization. The left-to-right mapping correctly predicts that the unassociated consonant will be at the right side of the root. Such behavior is attested as well in the nominal system, described in McCarthy (1979).<sup>7</sup>

### 3.3. *Vocalism*

As I have already noted, certain verbal categories such as aspect and voice are marked on the various binyanim not by the disarrangement of consonantism but rather by altering the quality of the vowels of the stem in a systematic way. This is untrue of the first

<sup>7</sup> Although quadrilateral roots are usually confined to their four binyanim, there is some evidence of the extension of quadrilateral roots to binyanim that provide only three C-slots in the prosodic template, with the expected loss of the supernumerary consonant. It has long been noted that many quadrilateral roots have near-synonymous trilateral doublets: *šamʕal* 'to be scattered', *šamaʕ* 'id.'; *xalbas* 'to deceive with soft words', *xalab* 'id.'. Usually the rightmost consonant is lost in the trilateral form, though not invariably so. These facts clearly could be analyzed as suggested in (38), although more study is needed.

triliteral binyan, so my subsequent remarks in this section are restricted to the other binyanim, and I will return to the problem of the first binyan later in section 3.4.

Let us examine the nature of this systematic variation in vowel quality. In the first column of table 1 above, the stem contains from two to four vocalic morae, all of which are *a*. In the second column, the last vowel is *i* but the other one to three vowels are *u*. We will skip the third column for the moment, proceeding in the same way with the remaining columns. The net result is the following set of vowel patterns associated with verbal categories:

(39) Perfective Active	$a_2^4$
Perfective Passive	$u_1^3 i$
Imperfective Passive	$u a_2^4$
Active Participle	$u a_1^3 i$
Passive Participle	$u a_2^4$

Each of these verbal vowel patterns serves for all binyanim but I. Each pattern has one vowel that spreads to fill up all the spaces in the stem except those that are occupied by other vowels fixed at either end of the stem.

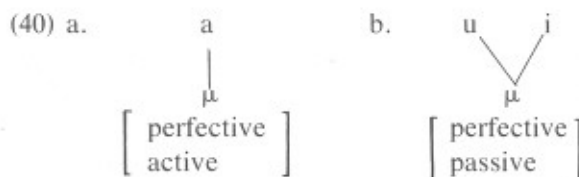
We now have two generalizations to account for:

(A) The categories in (39) do not alter the canonical shape of the stem.

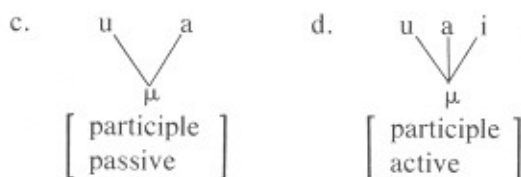
(B) The categories in (39) do alter vowel quality.

The one exception to the first of these generalizations is that the imperfective prefixes *V* and the participles prefix *mV* to the stems of the binyanim. Actually, both the imperfective and the participles prefix the prosodic template affix [CV]. The melody associated with *V* depends on the categories in (39), while the one associated with *C* is invariably *m* in the participle and varies with subject agreement in the imperfective. This phenomenon, which is discussed at greater length in McCarthy (1979), shows that we can have affixes composed solely of prosodic template material, like the prefix [CV].

Apart from this, it is apparent that the difference in the categories of (39) lies solely in the quality of the vowels. Consequently, we can isolate melodies from each of the vowel patterns in (39). These melodies are the morphemes that mark the indicated categories, and they all appear on a morphologically defined tier which takes [+syllabic] positions of the prosodic template as its set of tone-bearing elements:<sup>8</sup>



<sup>8</sup> Here, as in the preceding section, I represent melodic elements as the conventional segments *a*, *i*, and *u*. These are intended only as convenient abbreviations for the feature bundles of archisegments unspecified for [syllabic] and [segmental].

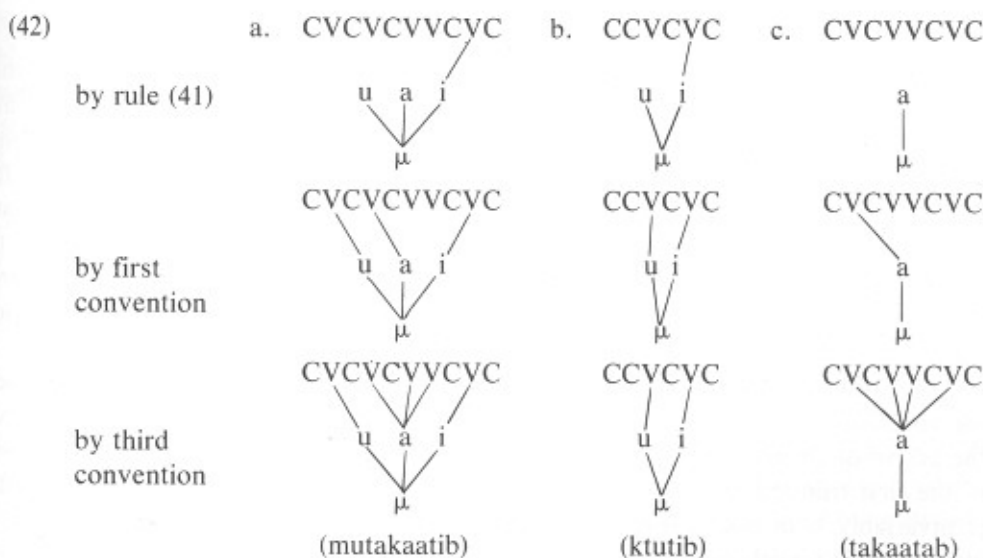


The universal conventions alone are not sufficient to ensure the correct association of these melodies with the V-slots of the prosodic template. We must first apply rule (41), which takes precedence over all the universal conventions:

- (41) *Vowel Association*  
 VC]  
 :  
 i

This rule says that the melodic element *i* of the perfective passive and active participle must be associated with the final vowel of the stem. The remainder of the association is accomplished by the first and third universal conventions, left-to-right association and spreading from the left.

A few sample derivations of the vocalism run as follows:<sup>9</sup>



The melodies of the imperfective active are somewhat more complicated, since they vary under phonological or morphological conditions. Three different melodies occur

<sup>9</sup> Strictly speaking, the full representations in (42) and later examples should contain the material on root and affix tiers analyzed in section 3.2. These additional tiers, however, cannot be conveniently depicted on a printed page. The reader may visualize them as appearing above the prosodic template and on another plane perpendicular to the one containing the rest of the representation.

on the surface, all of which are correctly associated with their templates by the procedure illustrated in (42):

- |      |   |  |
|------|---|--|
| (43) | <i>Binyanim</i>   | <i>Melodies</i>  |
| a.   | II, III, IV, QI   | $\begin{array}{c} u \quad a \quad i \\ \quad \diagdown \quad \diagup \\ \quad \mu \end{array}$ |
| b.   | VII, VIII, IX, X, XI,<br>XII, XIII, XIV, XV,<br>QIII, QIV | $\begin{array}{c} a \quad i \\ \quad \diagdown \quad \diagup \\ \quad \mu \end{array}$         |
| c.   | V, VI, QII  | $\begin{array}{c} a \\   \\ \mu \end{array}$   |

In McCarthy (1979) it is argued that all three melodies in (43) are derived from the underlying melody *u-a-i* (43a) by partly morphologically conditioned rules deleting *u* and *i* melodic elements. I will not repeat the details of this analysis here, but will simply observe that the single melody in (43a) is the basic form of the imperfective melody under this analysis.

### 3.4. *The First Binyan*

We will now turn to the issues presented by the somewhat more varied finite forms of the first triliteral binyan. (Discussion of the participles, which involve further complications, can be found in McCarthy (1979).) The first binyan is unique in that the canonical pattern of the perfective stem [CVCVC] differs other than in prefixation of [CV] from the canonical pattern of the imperfective [CVCCVC]. We can account for this alternation by rule (14b), which transforms an underlying [CVCVCVC] prosodic template to a derived [CVCCVC] one. Thus, the first binyan regularly receives the usual [CV] prefix in the imperfective and is then subject to elision of the middle vowel.<sup>10</sup> A conventional segmental rule with similar effect is formulated by Brame (1970).

A further peculiarity of the first binyan, and a much more complicated one, lies in the vocalism. We have isolated a single perfective and a single imperfective melody for the active of all other binyanim, but this result does not carry over to the active voice of the first triliteral binyan. First of all, in this binyan the vowel of the initial syllable is invariably *a* in both aspects. We will record this observation with a special rule inserting this vowel, associated with the first vowel of the stem:

- (44) 
$$\left[ \begin{array}{l} \text{First Binyan} \\ \text{Active} \end{array} \right] \begin{array}{c} [C \ V \\ \vdots \\ a \end{array}$$

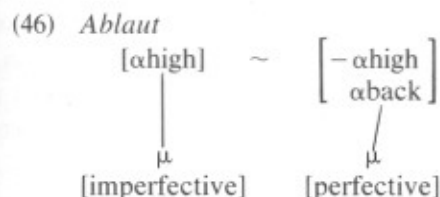
<sup>10</sup> The information about hierarchic structure encoded into (14b) by means of square brackets ensures that, although (14b) demonstrably applies in the template of first binyan imperfectives, it fails to apply in fifth and sixth binyan imperfectives *yatakattab* and *yatakaatab*.

Separate generalizations hold for the second syllable. It is subject to alternations in a complex set of ablaut classes, which are exemplified in (45):

(45)	Perfective	Imperfective	Examples
a.	a	i	ḍarab, yaḍrib 'beat'
b.	a	u	katab, yaktub 'write'
c.	i	a	ʔalim, yaʔlam 'know'
d.	u	u	ḥasun, yaḥsun 'be beautiful'

Some of these ablaut patterns are associated with verbs of a particular semantic class, though not strictly. Ordinarily, the first binyan form of a particular root is restricted to just one of these ablaut classes, but some slippage appears. There are also rare cases of anomalous ablaut, exhausting almost all the possibilities.

It is obvious that we can give only a lexical account of assignment of any given root to an ablaut class. It is further clear that there is no unambiguous ablaut function from perfective to imperfective or vice versa. That is, given any vowel in one aspect, we cannot uniquely determine its quality in the other aspect. Nevertheless, it is possible to relate imperfective to perfective if we exclude class (45d), which also has the regular semantic property of stativity. The ablaut redundancy rule (46), which reflects essentially the same observation as its counterpart in Chomsky and Halle (1968), invokes a polarity shift between aspects on the first binyan melody:

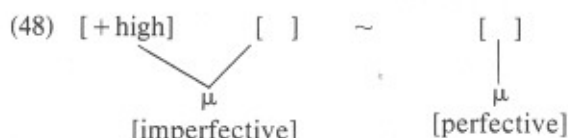


Unlike the formulation given by Chomsky and Halle, rule (46) is a generalization over the perfective and imperfective melodies, rather than the actual vowel segments of the stem. This has a few extremely interesting consequences for some facts we have already discussed.

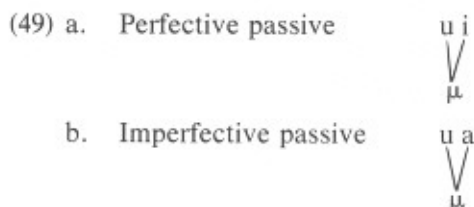
First, consider the underlying melodies of the perfective and imperfective active in the derived binyanim. They are repeated below for convenience:



Now if the Ablaut rule (46) is applied to the imperfective melody, it will relate the final *i* of the melody to *a*. Then, by the revised Obligatory Contour Principle discussed earlier in connection with the treatment of biliteral roots, this *a* collapses with the preceding identical melodic element into the single unit [ - high]. Therefore, it remains only to erase the initial *u* portion of the imperfective melody to yield the perfective of the derived binyanim. I will formulate this as the redundancy rule (48):



An even stronger argument for the prosodic analysis can be made from the imperfective and perfective passive melodies, repeated in (49):



The Ablaut rule (46) also expresses the relation between these two melodies, but with a further consequence when the melodies are mapped onto segments. The second element of the melody spreads in the imperfective passive, so it is impossible to state the polarity generalization just on vowels, since up to four morae might be associated with that melodic element. If Ablaut (46) were just a segmental rule (as is its counterpart in Chomsky and Halle (1968)), then it would systematically relate an imperfective passive form like *utakaatab* to the nonexistent perfective passive form *\*tukaatib* rather than the actual *tukuutib*. It is only at the level of the autosegmental melody that the Ablaut rule can express the aspectual relationships of the passive. This particular phenomenon, then, lends strong support to the prosodic analysis.

#### 4. Conclusions

What has emerged in the above discussion is a partial grammar of Arabic verbal morphology that captures a number of significant but otherwise inexpressible generalizations with a simple and elegant set of language-particular rules and representations and with the mostly independently motivated universal apparatus of autosegmental phonology. This analysis and its concomitant theoretical principles constitute, without elaboration, a contribution to the problem of nonconcatenative morphology as instantiated in a Semitic language.

We can, however, delineate more sharply some of the results for linguistic theory that follow from these considerations. Two main points are discussed below: the ap-



propriate formal power of morphological rules, and ways of extending this prosodic theory of morphology to the treatment of nonconcatenative phenomena, particularly reduplication, in languages other than Arabic. Some further results, concerned with reduplication ordering paradoxes and with the internal structure of the lexicon in a largely nonconcatenative morphological system, can be found in McCarthy (1979).

#### 4.1. *Formal Properties of Morphological Rules*

We have seen that, just at the level of surface phenomena, Arabic exhibits a wide variety of nonconcatenative morphology: ablaut processes, apparent movements of segments to restructure canonical patterns, reduplication, and infixation. One result of the prosodic theory is that all of this manipulation can be accomplished without recourse to transformational formalism. In generative studies of nonconcatenative morphological systems, the only means of describing phenomena like reduplication and infixation has been the use of transformational notation—ordinarily reserved for phonological rules of metathesis and coalescence—to copy or move segments. In the analysis presented here, however, it is sufficient to capture all the relevant generalizations if the theory provides morphemes on autosegmental tiers, morphological rules of the form  $A \rightarrow B / X$ , and the universal and partly language-particular apparatus of autosegmental phonology. No need was demonstrated for the richer transformational formalism, in spite of the complexity of the phenomena and the depth of the analysis.

In the light of these observations, I propose the following universal principle:

##### (50) *Morphological Rule Constraint (MRC)*

All morphological rules are of the form  $A \rightarrow B / X$ , where  $A$  is a single element or zero and  $B$  and  $X$  are (possibly null) strings of elements.

That is, morphological rules must be context-sensitive rewrite rules affecting no more than one segment at a time, and no richer type of rule is permitted in the morphology. It is to be assumed that the MRC applies to rules which have already been put in their most highly valued form according to the familiar procedure for minimization of features in Chomsky and Halle (1968). This is to eliminate the possibility of subverting the MRC by translating some morphological transformations into complex conjunctions of non-transformational morphological rules.

It is obvious that a theory that incorporates the MRC strongly generates a smaller class of grammars than a theory without this constraint. Morphological transformations potentially allow any arbitrary operation on a segmental string. For example, transformational morphological rules can freely move particular segments an unbounded distance within the word, copy all and only the vowels in a word, or reverse strings of finite length. If the segmental representation is further enriched by permitting integral indexing of segments, as in Chomsky's (1951) analysis of Modern Hebrew intercalation described in the appendix, then morphological transformations can perform their arbitrary oper-

ations on only the prime or factor-of-twelve numbered segments in the word with no further enrichment of the formalism.

These examples, although bizarre, are not facetious. It is a fact that a morphological theory without the MRC allows all of these types and in some cases values them more highly than morphological rules that actually occur in some language. A theory with the MRC is therefore significantly more explanatory than one without it.

Of course, one could object that although the MRC delimits a theory with lessened strong generative capacity, it has no corresponding effect on weak generative capacity. It is fine to eliminate morphological transformations, so the argument goes, but isn't it possible to encode the same effects into the phonological rules, which do allow transformational formalism?

The defect in this argument is that it takes no cognizance of the theory of phonological rule naturalness which, although only imperfectly understood at this point, nevertheless must be a part of linguistic theory as a whole. To see how this works, let us examine the archetypical phonological rules that must be formulated transformationally: rules of metathesis. It has been observed both traditionally and in more recent studies (Ulan (1971)) that only a very limited set of possible metathesis rule types exists, depending on phonetic properties of the affected segments. One type is vowel-liquid metathesis, represented, for example, by the Maltese rule of Brame (1972). This apparently reflects a more general type of metathesis between neighboring continuants of unequal sonority, as the Latvian vowel-glide metathesis of Halle and Zeps (1966) shows. Another sort is stop-spirant metathesis, like the Akkadian rule of section 2.1. An apparently distinct type, involving identical consonants separated by a vowel, is attested in the Classical Arabic rule of section 3.2.

It is fairly clear from these brief observations, as well as others by Ulan (1971), that there exists a quite limited set of possible metathesis rules, which we could characterize as a preliminary theory of natural metathesis. Although linguistic theory allows full transformational formalism in phonological rules, it is nevertheless subject to this sort of substantive constraint. Therefore, only a small subset of the formally possible metathesis rules will actually occur, since many possibilities will be excluded on phonetic grounds. Notice, however, that it is impossible to place any such constraints on the phonetic naturalness of morphological rules. It follows directly from *l'arbitraire du signe* that phonetically determined considerations of naturalness have no place in morphological rules. Therefore, any constraint on the morphology must be an essentially formal one, like the MRC.

I conclude, then, that a linguistic theory that incorporates the MRC is more constrained than and consequently superior to a theory that does not, all other things being equal. The most striking confirmation for the empirical validity of this restriction on linguistic theory is the grammar of the Classical Arabic verb developed above. Despite morphological phenomena that appear to invite analysis by morphological transformations, a revealing analysis was constructed that relies entirely on the universal apparatus

of a version of autosegmental phonology and language-particular context-sensitive rewrite rules.

#### 4.2. *Beyond Arabic*

It is clear that ordinary concatenative morphological processes can be formulated in a way entirely consistent with the MRC. The same is true of relatively simple ablaut rules, like those found in the English strong verb system. On the other hand, there are several types of phenomena that have usually been analyzed in other languages as the results of morphological transformations. This is particularly true of reduplication, which has received the most attention. Although the material in the literature is far beyond any individual's capacity for reanalysis, it is nevertheless possible to show that the prosodic theory proposed here accounts for a variety of observations that have not been adequately dealt with previously. I will conclude with a brief discussion of some recent refinements of this model of reduplication.

Let us consider the basic characteristics of reduplication in the prosodic model. The basis of Arabic morphology is a set of prosodic templates that vowel and consonant melodies are mapped onto by certain rules of great generality. Reduplication can be characterized formally as a one-to-many association of a single melodic element with the slots of the prosodic template. That is, reduplication is just an instance of the more general autosegmental phenomenon of spreading. This is the case, for example, with reduplication of the *u* portion of the perfect passive melody in sixth binyan *tukuutib* or of the final root consonant in ninth binyan *ktabab*. In every instance, the surface reduplication is not a consequence of a transformational rule but rather of the spreading of a particular melodic element to fill up the available slots of the template.

Although the bulk of Arabic reduplication results from spreading of melodies onto a template made up of V and C positions, this is not always true. Another kind of reduplication shows how far the notions of association and morphologically defined tier can take us in dealing with problematic morphological types. In Arabic a number of quadriliteral verbs are of the pattern  $C_iVC_jC_iVC_j$ : *ğarğar* 'to gargle', *waswas* 'to whisper', *zalzal* 'to shake'. As is apparent from the glosses, these forms have some sort of elusive onomatopoeic effect. These words are not generally related to any trilateral verbs, so there is little evidence here for a word-formation process. Therefore, I will concentrate on Biblical Hebrew, where this evidence does exist. My remarks about the formal character of this sort of reduplication hold equally well for Arabic, so no theoretical point hinges on switching languages here.

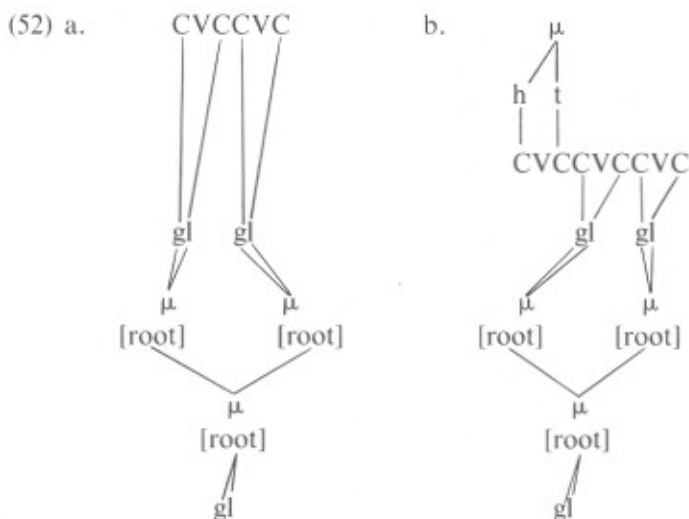
In Hebrew, traditional grammar recognizes a binyan known as the *piḥpel*, and a related reflexive *hitpalpel*. In attested cases, these are formed from biconsonantal root types:<sup>11</sup>

<sup>11</sup> The vowel *ā* in the Hebrew forms in (51b) is a very short, schwa-like vowel. It is introduced by a relatively late phonological rule described in detail in McCarthy (1979). It may be freely ignored in the morphological analysis. Further complications in the transcription of Biblical Hebrew vowels are irrelevant to the issues at hand.

- (51) a. root: gl  
 first binyan: gālāl 'to roll (intrans.)'  
 pilpel: gilgēl 'to roll (trans.)'  
 hitpalpel: hitgalgēl 'to roll oneself along'
- b. root: šf  
 first binyan: šāfaš 'to be smeared'  
 pilpel: šifšāšaš 'to stroke'  
 hitpalpel: hištašāšaš 'to indulge oneself'

Semantically, the pilpel generally has the usual transitivity or causative force of the piṣel (= Arabic second binyan), while the hitpalpel is a reflexive like the hitpaṣel (= Arabic fifth binyan). In formal terms, the pilpel and the hitpalpel are just instances of the Hebrew reflexes of the Arabic second and fifth binyanim, with which they share similar semantics and identical prosodic templates.

The autosegmental interpretation of these facts is that a biconsonantal root is expanded to fit a prosodic template—the [CVCCVC] template of the causative and [CVCCVCCVC] of the reflexive—with four empty slots. However, in this case the expansion is not effected by reduplicating a single root consonant, but rather by reduplicating the entire root. Now, with a slight enrichment of the notion of a morphological tier, it is possible to speak of a mapping between morpheme positions rather than directly between elements of a morpheme and the corresponding template. That is, the root is reduplicated by a one-to-many morpheme-to-morpheme association, and then the elements of these morphemes are mapped onto the prosodic template. I will represent this formally in the following way:

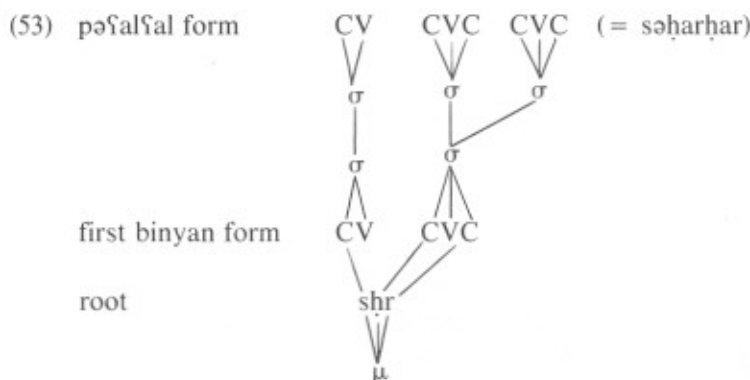


That is, reduplication is accomplished here by mapping one root morpheme onto two root morpheme positions in a separate tier. The units contained in these derivative

morphemes are then mapped onto the prosodic template. All of this mapping follows directly from the usual conventions. The sole thing that is stipulated is that verbs of this type in Hebrew (or in Arabic) have associated with them two positions labeled  $\mu$ ,<sub>[root]</sub> so the root can be reduplicated. This extra stipulation is justified because the usual result of mapping a biconsonantal root onto a four-slot template is double reduplication, like *sibbēb* 'he surrounded' (cf. also (34)). Reduplication of the entire root is limited to a lexically-governed class of verbs in Hebrew, but formally similar morpheme reduplication is encountered frequently in other languages.

Clearly this mechanism will work in Arabic; moreover, Arabic has some additional evidence that verbs like *zalzala* constitute a definable class within the lexicon. One bit of evidence is the semantic consistency of this class alluded to earlier: these forms seem to refer to repeated, iterative operations. A much stronger argument lies in the formation of gerunds or infinitives from verbs of this class. Verbs like *zalzala* often form gerunds of the pattern *zalzaal*, *ḡalḡaal*, and so on. However, no other trilateral or quadrilateral verb can form a gerund of this pattern. Therefore, the rule responsible for just this type of gerund must be able to refer directly to verbs with reduplicated biconsonantal roots. The theory offered here allows exactly this, since verbs of this type all have two  $\mu$ ,<sub>[root]</sub> slots associated with them.

A further extension of this theory also handles the forms in a very rare binyan of Hebrew. This is the *pəʕalʕal*, which seems to be connected with intensification of some sort. For instance, corresponding to the first binyan form *sāḥar* 'to go about' is the *pəʕalʕal* form *səḥarḥar* 'to palpitate'. Clearly, it is not the whole root that is reduplicated here, but rather the final syllable of the stem. Now the prosodic template of the *pəʕalʕal* is somewhat anomalous in Hebrew, since it involves an otherwise nonoccurring [CVCVCCVC] prosodic template. I suggest that it is derived from the [CVCVC] template of the first binyan by suffixation of the syllable [CVC], and that then the syllables of the first binyan are mapped—as always, from left to right—onto the syllables of this new template. I will characterize this process using the notation for syllable structure developed by Kahn (1976):



In view of the paucity of relevant examples in Hebrew, this analysis must, of course, be considered tentative. Nevertheless, it suggests a range of possibilities where prosodic units other than C/V and  $\sigma$  are represented in one-to-many associations. This theory predicts that, in principle, any labeled prosodic category could be subject to reduplication. Thus, we might conjecture that reduplication of the prosodic category foot (Selkirk (forthcoming)) is responsible for sporadic English formations like *higgledy-piggledy*. This irregular English process is clearly not compelling evidence, but the prediction made by the theory is confirmed by Nash's (1980) extensive analysis of verbal reduplication in Warlpiri. Nash demonstrates that Warlpiri verbs productively form a sort of intensive or distributive derivative by reduplicating exactly the first metrical foot. Besides the foot, if the syllable contains labeled internal subconstituents such as onset, rhyme, or coda, we would plausibly expect these prosodic categories to reduplicate as well. This indicates that there may be a much richer variety of units subject to reduplication than is represented by C/V prosodic templates.

In general, then, the formal basis of reduplication is the specification of a template composed of positions such as V and C or  $\mu$  and the regular autosegmental mapping onto that template. No special rules of reduplication are needed—the phenomenon simply arises whenever the universal or language-particular rules of association yield a one-to-many association between the melody and the template. Morphological categories with characteristic reduplication, like the Arabic verbs of the *zalzal* type and the related Hebrew *pilpel*, simply stipulate a template in which this sort of association necessarily arises.

Not surprisingly, there are several interesting empirical consequences of this very reduced apparatus for describing reduplication phenomena.

First, the directionality of reduplication is, in general, invariant. Since the direction of reduplication—the position of the reduplicated string with respect to the rest of the form—is a direct consequence of the direction of association, a left-to-right rule of association yields reduplication at the right end of the stem. Clearly, other rules of association, right-to-left in particular, could yield other directions of reduplication. However, the prediction, generally borne out by the Semitic verb data as well as by casual observations of other languages, is that the apparent direction of different reduplication phenomena should be invariant. Languages can deviate from this only at greater cost. Thus, it requires the stipulation of an additional rule, the Second, Fifth Binyanim Erasure rule of section 3.2, to yield medial reduplication in the forms *kattab* and *takattab*. As in these forms, medial reduplication, which is apparently quite rare, will always require an additional stipulation. Thus, the unmarked case under the prosodic theory is for a language to reduplicate exclusively at the left or right stem boundary.

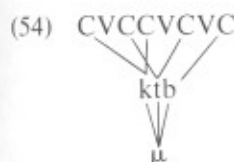
Second, there is only a very limited possibility in the prosodic theory of restricting reduplication to particular phonologically defined classes of forms. To see the significance of this, consider two putative reduplication rules formulated transformationally. One rule reduplicates any final string CVC, while the other reduplicates that string only



if the final consonant is a lateral. These two rules are equally valued in the transformational theory; the first applies to CV[+cons], the second to CV[+lat]. This is, however, almost certainly the wrong prediction, and clearly the first rule should be much more highly valued if the second is possible at all. In fact, one result of Moravcsik's (1978) survey of a number of reduplication phenomena is that no phonetic specification of the reduplicated string is ever necessary except its composition in terms of V and C. This observation is obviously supported in detail by the analysis of Arabic presented here.<sup>12</sup>

Under the prosodic model, a morphological category which characteristically reduplicates simply stipulates an output template composed of V/C or  $\mu$ . The template cannot refer to the whole rich set of phonological features. It is therefore impossible to restrict reduplication to forms sharing some other properties, short of additional arbitrary restrictions on the mapping rules.

A kind of corollary to this property of the theory is the result that reduplication is limited to strings that form constituents at some level of representation. The notions of mapping and spreading are meaningful only insofar as they involve the association of constituents at one level (like morphemes, syllables, or individual elements of the autosegmental melody) with units at another level (like V, C,  $\sigma$ , or  $\mu$  positions in the prosodic template). Association of a nonconstituent string on one level with a constituent string on another level is excluded formally because it necessarily leads to an ill-formed representation with lines crossing. By this logic, then, there can be no Arabic binyan characteristically formed like \**katakatab* from the root *ktb*. The only possible representation for this hypothetical binyan would be (54):



Since *k* and *t* do not exhaust a constituent on any tier, there is no way to derive \**katakatab* without crossing association lines.

In sum, we see that the prosodic model of morphology not only provides a revealing account of the complexities of the Arabic verb, but also yields a rich variety of results concerning the universal properties of nonconcatenative morphological phenomena, particularly reduplication. An earlier version of this work has generated several responses which, while occasionally suggesting some refinements of the theory to handle some circumstances not attested in Arabic, have largely confirmed its basic insights.

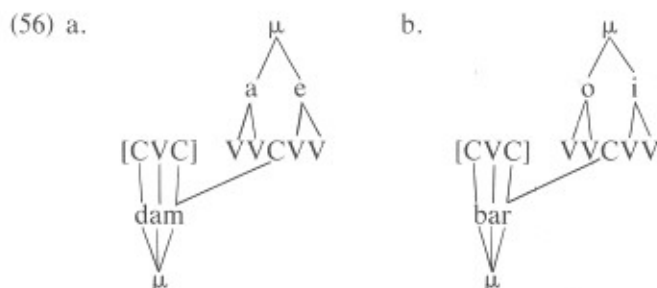
<sup>12</sup> A particularly compelling result of notating reduplication by means of a prosodic template arises in Cupeño, as described by Hill (1970). Hill argues that the habilitative construction is formed from consonant-final stems by repeated reduplication until an output target (that of having two syllables follow the stress) is reached. The template, then, can encode the output target, and automatic spreading effectively reduplicates until the template is filled. A formal analysis of these observations can be found in McCarthy (1979).



Harris (1980) and Halle and Vergnaud (1980) apply the prosodic model to some relatively intractable problems in Spanish and Hausa plural formation with great success. For example, Hausa has a fairly large class of nouns which reduplicate the stem-final consonant in the plural:

- (55) *Singular Plural*  
 dāmóo dāmàamée 'land monitor'  
 báráa báróorí 'servant'

These forms can be analyzed as [CVC] stem templates with [VV] template suffixes in the singular and [VVCVV] template suffixes in the plural. The full representation of the two plural forms appears in (56):



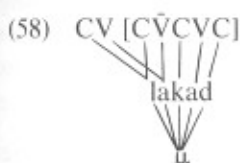
The stem-final consonant is automatically associated with the unspecified C-slot of the template suffix, capturing the fact that these plurals have characteristic reduplication. Notice the role played by morphologically defined tiers in this representation: since the plural suffix melodies *a-e* and *o-i* are represented on a separate tier from the stem melodies *dam* and *bar*, the association of the final consonant can spread without inhibition.

These forms and the other Hausa data cited by Halle and Vergnaud suggest the first refinement of the prosodic theory. Recall how the grammar of Arabic ensures that vowel melodies are associated with V-slots and consonant melodies with C-slots. Since vowel melodies and consonant melodies invariably appear on different morphologically defined tiers in Arabic, it suffices to indicate for each tier what its melody-bearing elements will be, either syllabic or nonsyllabic positions in the prosodic template. This proposal clearly will not suffice for Hausa and, in general, any language without the Semitic consonantal and vocalic morphemes. A simple enrichment of the model provides a ready account of this observation. Suppose we retain the prosodic template unaltered, but require, at least in some cases, that each melodic element bear an indication of whether it is to be associated with a C or a V in the template. In effect, the melodic elements will be specified as [+syllabic] or [-syllabic], and the association conventions must match up values of this feature between the melodic and template tiers. An explicit procedure for this matching can be found in Halle and Vergnaud (in preparation).

Another possible refinement of the prosodic theory comes from some reduplication phenomena of Tagalog. Lieber (1980) and Marantz (1980) have argued that data like that in (57), based on work by Carrier (1979), present apparent counterexamples to simple reduplication by one-to-many association.

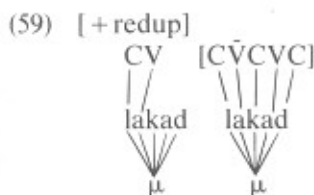
- (57) a. um-lākad 'walk'      pag-lalākad 'walking'  
       b. kandīlah 'candle'      pag-kakandīlah 'candle vendor'

It is clear that the reduplication process in (57) copies a CV sequence, concomitantly shortening the vowel. Even with the enrichment of the prosodic theory suggested above in connection with Hausa, we cannot represent reduplication of a CV sequence where both consonantal and vocalic melodic elements appear on the same morphologically defined tier:



Like (54), this ill-formed representation with association lines crossing is a consequence of attempting to reduplicate a nonconstituent on a melodic tier.

Marantz (1980) proposes a straightforward modification of the prosodic theory that permits an account of data such as those from Tagalog. I will deviate somewhat from his suggestion in the following formulation. Let us suppose that some prosodic template affixes bear the feature [+reduplication], which induces special behavior. This feature should not be taken as one of the familiar triggers of a reduplication transformation in standard generative analyses. Rather, it has the effect of causing automatic copying of all the melodic elements in some morpheme—formally, all the daughters of some  $\mu$  in a particular tier. This copied material is then associated in the familiar way with the C/V positions of the prosodic template affix. As in the analysis of Arabic, material remaining unassociated is considered to be deleted or without phonetic effect. A sample representation of the form (pag-)lalākad will appear as shown in (59):



The melodic elements *kad* which remain unassociated have no effect on the phonetic output.

This rather small elaboration apparently solves the problem presented by the Tagalog data and similar facts in a way which is well within the spirit of the prosodic theory

of nonconcatenative morphology advanced here. It has most of the desirable characteristics of that theory, particularly the lack of transformational morphological rules. The copying induced by the presence of the feature [+reduplication] is part of universal grammar, not part of some language-particular reduplication transformation, and consequently it is irrelevant to the whole problem of restrictiveness. This formulation also allows an elegant characterization of the apparent differences between reduplication in Classical Arabic and Tagalog: in the former, the feature [+reduplication] plays no role, so all reduplication is a consequence of one-to-many associations derived by the usual conventions. It remains to be seen whether other alternatives along these general lines for data of the Tagalog type can be found.

In sum, this success in applying and extending a theory that was originally justified on the basis of Arabic morphology to such typologically diverse languages strongly suggests that this model will yield rich insights into a wide variety of morphological phenomena.

## Appendix

In view of the apparent similarities between many of the notions of autosegmental phonology and Harris's long components, we could reasonably expect the theory developed here to have been prefigured somewhat by earlier work. In fact, there exist fairly detailed accounts of Biblical and Modern Hebrew in terms of the theory of long components. I summarize these analyses in this appendix and I also include some criticisms and other observations about them.

The first modern insights into Semitic morphology appear in Harris's (1941) analysis of Biblical Hebrew. Harris proposes a list of morphemes divided into three types on formal and semantic grounds. The consonantal roots like *ktb* have the sort of general meaning alluded to earlier. Morphemes of the second class, patterns, are composed of vowels plus symbols from the set " \_ ", " : ", and affixal consonants. The dash marks "the presence of some phoneme, usually a consonant, in close juncture" (Harris (1941, 152)). The colon is the familiar notation for consonant length. The meaning of a pattern is essentially a modification of the meaning of the root. So, for instance, the pattern of *kattab* would be notated  \_ a \_ : a \_ with the meaning 'intensive, causative'. The third class of morphemes is relatively uninteresting, consisting of those function words and loans not obviously susceptible to root and pattern analysis.

The relationship between morphemes of the root class and those of the pattern class is expressed by a single statement of morpheme order: members of the root class are intercalated in patterns. This statement suffices, since any pattern will contain three dashes, one for each of the consonants of the root, so the mapping of consonants to slots is unambiguous. Thus, Harris has a very simple expression of the fundamental morphological process of Hebrew. The cost of this simplicity is a significant loss of generality in the characterization of patterns. It is, therefore, an accident under this theory that nearly all verb patterns contain a portion of the form  \_ V \_ V \_, or that the

vowels in all patterns with two vowels are placed in the same way with respect to the dashes for the root consonants. The actually attested possibilities of intercalating roots and patterns are much more limited than this apparatus allows.

Chomsky's (1951) analysis of Modern Hebrew eliminates this defect, though at the cost of greater complexity in the intercalation process. Chomsky (1951, 17) offers general schemata for roots and patterns of the form (60):

- (60) a.  $R \rightarrow C_1^R C_2^R C_3^R$  (:, sometimes, if  $C_2 = Y_2$ )  
 b. Vowel Pattern:  $\alpha_1 - \beta_2$  where  $\alpha_1, \beta_2 = V$  or  $\phi$

The notation  $C^R$  in the definition of a root refers to a set of morphophonemes that can occur in roots. The parenthesized material refers to a special case where the medial root consonant is a high glide (hollow root). The definition of a vowel pattern is quite general; the dash serves only to separate the two vowels, and not to indicate the position taken by a consonant. In practice, although not in this formal definition, Chomsky also allows patterns with the symbol "::" immediately preceding  $\beta_2$ , indicating gemination of a consonant.

Since Chomsky's analysis is one of the earliest and most extensive demonstrations of rule ordering within a modified structuralist framework, we can coherently speak of a morphophonemic derivation. At the earliest stage of this derivation, there is a linear concatenation of morphemes from the different classes. So, for instance, the stem of *kattab* will have the remote representation  $ktb + a - :a$ . Several morphophonemic rules apply to representations of this sort. These rules must, by Chomsky's ordering argument, crucially precede a morphophonemic rule of intercalation, formulated as (61) (Chomsky (1951, 23)):

$$(61) C_1 C_2 C_3 \left\{ \begin{array}{c} \vdots \\ - \end{array} \right\} + Q_1 - (:) Q_2 \left[ \begin{array}{c} \vdots \\ - \end{array} \right] \rightarrow \\ C_1 Q_1 C_2 (:) Q_2 C_3 \left\{ \begin{array}{c} \vdots \\ - \end{array} \right\} \left[ \begin{array}{c} \vdots \\ - \end{array} \right], \text{ and } :: \rightarrow : \\ \text{where } Q_i = V_i \text{ or } \phi [i = 1, 2]$$

Since the mode of application of this rule may not be entirely perspicuous, I will attempt to paraphrase it.

The consonants of a root and the vowels of a pattern are indexed by subscript integers from left to right. In concatenation, the first vowel ( $Q_1$ ) is placed after the first consonant ( $C_1$ ). If the second vowel is preceded by a colon, then the colon is placed after  $C_2$ , indicating gemination of the second root consonant. The colon is itself followed by the second vowel ( $Q_2$ ) and then by the third root consonant ( $C_3$ ). Curly brackets and square brackets both are identical in effect to the curly brackets of Chomsky and Halle (1968), except that the former are expanded before the latter. The result of these notations in (61), along with the reduction of "::", is that length of either  $C_3$  or  $Q_2$  or both in the input is realized by length of  $C_3$  in the output.

In essence, then, the operation of intercalation in Chomsky's analysis is a transformational rule that refers to indices on vowels and consonants according to their positions in the stems and roots. While Harris stipulates for each pattern where consonants will fall within it by the dash notation, Chomsky abstracts away to a generalized vowel pattern and writes a rule to indicate the relative ordering of members of roots and vowel patterns.

Although a model of insightful and compact statement, Chomsky's analysis is descriptively inadequate on a few relevant points. One of these is the treatment of quadriliteral roots. Although he disavows an explicit treatment of them, Chomsky does tentatively suggest that these roots are accommodated by replacing ":" with a root consonant in vowel patterns of the form  $V_1—:V_2$ . That is, a root consonant is substituted formally for medial gemination. For example, replacement of ":" by *g* in the pattern  $i—:e$  would yield the quadriliteral verb *tirgem* 'he translated'. Apart from the obvious fact that this requires a new, ad hoc rule to deal with quadriliteral roots, it also apparently makes the incorrect claim that these roots are derived from triconsonantal roots by augmentation. It is not possible to substitute any consonant for ":"; only *g* will do if the rest of the root is *trḡm*. I conclude, then, that the mode of intercalation in (61) is inadequate for roots of four consonants.

Chomsky's analysis also fails, as does Harris's, to provide a means of stating generalizations about vowel patterns independently of generalizations about geminate root consonants. Both theories stipulate gemination by including a colon in a particular vowel pattern. Although this is adequate for Hebrew, it misses an important property of Arabic verb forms like *katab* versus *kattab*. In these words, all vocalism is *a* as one manifestation of the perfective active. The vocalism can be changed independently of the gemination of the medial consonant; compare the corresponding passives *kutib* and *kuttib*. This cannot be accounted for in a model like Harris's or Chomsky's that marks consonant gemination on vowel patterns.

Finally, Chomsky's analysis is subject to the same criticism of insufficient restrictiveness as the classical generative morphological theory. The rule of intercalation in (61) is, clearly, a transformation, implying an apparatus with corresponding descriptive power. A rule formulated along these lines seems unavoidable in an explicit long component analysis of any morphological system of the Semitic type.

## References

- al-Mozainy, H. (in preparation) *The Phonology of a Bedouin Hijazi Dialect*, Doctoral dissertation, University of Texas at Austin, Austin, Texas.
- Aronoff, M. (1976) *Word Formation in Generative Grammar*, Linguistic Inquiry Monograph 1, MIT Press, Cambridge, Massachusetts.
- Brame, M. (1970) *Arabic Phonology: Implications for Phonological Theory and Historical Semitic*, Doctoral dissertation, MIT, Cambridge, Massachusetts.

- Brame, M. (1972) "On the Abstractness of Phonology: Maltese  $\mathfrak{f}$ ," in M. Brame, ed., *Contributions to Generative Phonology*, University of Texas Press, Austin, Texas.
- Carrier, J. (1979) *The Interaction of Morphological and Phonological Rules in Tagalog: A Study in the Relationship between Rule Components in Grammar*, Doctoral dissertation, MIT, Cambridge, Massachusetts.
- Chomsky, N. (1951) *Morphophonemics of Modern Hebrew*, Master's thesis, University of Pennsylvania, Philadelphia, Pennsylvania.
- Chomsky, N. and M. Halle (1968) *The Sound Pattern of English*, Harper and Row, New York.
- Clements, G. N. and K. Ford (1979) "Kikuyu Tone Shift and its Synchronic Consequences," *Linguistic Inquiry* 10, 179-210.
- Goldsmith, J. (1976) *Autosegmental Phonology*, Doctoral dissertation, MIT, Cambridge, Massachusetts.
- Greenberg, J. (1978) "The Patterning of Root Morphemes in Semitic," in S. Al-Ani, ed., *Readings in Arabic Linguistics*, Indiana University Linguistics Club, Bloomington, Indiana.
- Halle, M. (1973) "Prolegomena to a Theory of Word Formation," *Linguistic Inquiry* 4, 3-16.
- Halle, M. and J.-R. Vergnaud (1980) "Three Dimensional Phonology," *Journal of Linguistic Research* 1, 83-105.
- Halle, M. and J.-R. Vergnaud (in preparation) *Three Dimensional Phonology*.
- Halle, M. and V. Zeps (1966) "A Survey of Latvian Morphophonemics," *Quarterly Progress Report of the Research Laboratory of Electronics* 83, 105-113.
- Harris, J. (1980) "Nonconcatenative Morphology and Spanish Plurals," *Journal of Linguistic Research* 1, 15-31.
- Harris, Z. (1941) "Linguistic Structure of Hebrew," *Journal of the American Oriental Society* 62, 143-167.
- Harris, Z. (1951) *Methods in Structural Linguistics*, University of Chicago Press, Chicago, Illinois.
- Hill, J. (1970) "A Peeking Rule in Cupeño," *Linguistic Inquiry* 1, 534-539.
- Hockett, C. (1947) "Problems of Morphemic Analysis," *Language* 23, 321-343.
- Jackendoff, R. (1975) "Morphological and Semantic Regularities in the Lexicon," *Language* 51, 639-671.
- Kahn, D. (1976) *Syllable-based Generalizations in English Phonology*, Doctoral dissertation, MIT, Cambridge, Massachusetts.
- Leben, W. (1973) *Suprasegmental Phonology*, Doctoral dissertation, MIT, Cambridge, Massachusetts.
- Lieber, R. (1980) *On the Organization of the Lexicon*, Doctoral dissertation, MIT, Cambridge, Massachusetts.
- Marantz, A. (1980) "Reduplication and C-V Skeleta," unpublished paper, MIT, Cambridge, Massachusetts.
- McCarthy, J. (1979) *Formal Problems in Semitic Phonology and Morphology*, Doctoral dissertation, MIT, Cambridge, Massachusetts.
- McCawley, J. (1970) "A Note on Tone in Tiv Conjugation," *African Linguistics* 1, 123-129.
- Moravcsik, E. (1978) "Reduplicative Constructions," in J. Greenberg, ed., *Universals of Human Language* III, Stanford University Press, Stanford, California.
- Nash, D. (1980) *Topics in Warlpiri Grammar*, Doctoral dissertation, MIT, Cambridge, Massachusetts.
- Pyle, C. (1972) "On Eliminating BM's," in P. Peranteau, J. Levi, and G. Phares, eds., *Papers from the Eighth Regional Meeting of the Chicago Linguistic Society*, Chicago Linguistic Society, Chicago, Illinois.

- Rotenberg, J. (1978) *The Syntax of Phonology*, Doctoral dissertation, MIT, Cambridge, Massachusetts.
- Selkirk, E. (forthcoming) *Phonology and Syntax: The Relation between Sound and Structure*.
- Ulan, R. (1971) "A Typological View of Metathesis," *Working Papers on Language Universals* 7, 1-44.

*Department of Linguistics*  
*University of Texas at Austin*  
*Austin, Texas 78712*