Prosodic Morphology and Templatic Morphology

John McCarthy
University of Massachusetts, Amherst

Alan Prince
Brandeis University

PROSODIC MORPHOLOGY AND TEMPLATIC MORPHOLOGY

JOHN McCARTHY
UNIVERSITY OF MASSACHUSETTS, AMHERST

ALAN PRINCE
BRANDEIS UNIVERSITY

1. Introduction

Much of the time, morphology is just word-syntax. That is, the morphological grammar of a language reduces to statements like "<u>ness</u> is a Level 2 suffix". But this is not always true, and the cases where it is not true reveal a great deal about morphological structure and its relation to phonology.

In many languages, morphological categories are expressed not by conventional affixes but by morphemes whose only constant is a fixed canonical pattern -- what might be called shape-invariant morphology. The most common kind of shape-invariant morphology is reduplication, but it is also central to the somewhat rarer templatic morphology of Arabic. In Arabic, various morphological distinctions are expressed by specifying a fixed canonical form of the stem that does not vary despite independent morphological or lexical changes in the consonants or vowels that fill this canonical form. For example, (1) demonstrates the property of shape-invariance for the Arabic causative, known as the <u>faʕʕala</u> or Form 2. (Here and throughout this article, unless otherwise indicated, Arabic words are given in their stem form, which abstracts away from the effects of phonological rules and the addition of inflectional affixes from the agreement, mood, and case-marking systems):

University of Massachusetts Occasional Papers in Linguistics, Vol. 13 [1990], Art. 8

(1) Shape Invariance in Arabic Form 2 (faʕʕala)

	/ktb/	/drs/	/ʕlm/	/sm/
	'write'	'study'	'know'	'poison'
	kattab	darras	ʕallam	sammam
perfect active				
	kuttib	durris	ʕullim	summim
perfect passive				
Cu+	kattib	darris	ʕallim	sammim
imperfect active				
Cu+	kattab	darras	ʕallam	sammam
imperfect passive				

Taken in the context of the fuller analysis of the verbal system in McCarthy (1981), this small array of facts is sufficient to demonstrate the property of shape-invariance in Arabic templatic morphology. Moving across the columns of (1) changes the consonantal root, the fundamental lexical unit of the language. Despite this change in the consonants, the canonical pattern remains the same. Similarly, moving down the rows of (1) changes the vocalism: voice goes from active to passive or aspect from perfective to imperfective. Again the canonical pattern remains the same. Similar regularities are met with throughout the Arabic system of verbal conjugations.

Clearly, as more information about shape-invariant morphology in general and templatic morphology in particular becomes available, it becomes increasingly important that a satisfactory theory of these phenomena underlies the analysis. In recent research (McCarthy and Prince 1986, 1988, 1990, forthcoming), we have developed an approach to shape-invariant morphology that is fundamentally founded in the phonology of prosody. It is called "Prosodic Morphology". The properties attributed to reduplicative and templatic morphology in Prosodic Morphology are independently motivated by their role in the characterization of phonological processes, stress, and versification. Our first task in this article is to lay out briefly the fundamental tenets of this theory.

The article continues with an extended analysis of the templatic morphology of Standard Arabic. We begin by sketching in a very brief and superficial way the nature of the prosodic analysis of Arabic templatic morphology. This is followed by detailed treatment of the most significant issues, demonstrating that the prosodic theory is not only a viable alternative to its predecessors but is in fact superior to them, revealing and capturing regularities that have played no role in previous treatments.

The investigation is guided by one fundamental idea, called the Prosodic Morphology Hypothesis. It asserts that the templates of reduplicative or templatic morphology are defined in terms of the authentic units of prosody: the mora, the syllable, the foot, and the phonological word. In other words, the Prosodic Morphology Hypothesis demands that the vocabulary of templates is the same as the vocabulary of prosody in general, including stress, syllabification, epenthesis, compensatory lengthening, rhyme, "counting rules", and poetic metre.

The prosodic constituents are arranged in a hierarchy of exhaustive domination (cf. Selkirk 1980):

(2) Prosodic Hierarchy

Phonological Word	W
	$|$
Foot	F
	$|$
Syllable	σ
	$|$
Mora	μ

The hierarchy is read from top to bottom, so the units at a higher level only contain units from lower levels. The phonological word corresponds roughly but not exactly to the grammatical or syntactic word; it is typically the domain of main-stress assignment. The foot is a constituent composed of at least one stressed syllable and usually an unstressed syllable as well. For example, the single phonological word indefensibility contains three feet dominating seven syllables, as (3) shows:

(3)

indefensibility
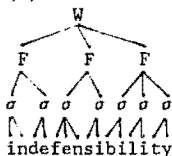
The mora is the unit by which syllable weight is measured -- its role in a theory like (2) has been explored by Prince (1983), Hyman (1985), McCarthy and Prince (1986 et seq.), Hayes (1989), Archangeli (1988), and Itô (1989). Investigation of systems of stress assignment, versification, and other phenomena reveals a fundamental distinction between two types of syllables, heavy and light. Usually heavy syllables are those that contain a long vowel (Cvv) or are closed by a final consonant (CvC), while light syllables are open with a short vowel (Cv), but occasionally

McCarthy & Prince

syllables contain two moras, while light syllables contain only
one.

This establishes the set of descriptive primitives and the
hierarchy in which they are arranged. The theory also includes
constraints on the combination of prosodic units, but we will not
give a comprehensive treatment of that topic at this time (see
McCarthy and Prince (1986, 1988, 1990, forthcoming) for additional
development), introducing only those notions that become necessary
as particular examples are treated.

3. Prosodic Morphology in Arabic

In McCarthy (1981), it is shown that a form like kuttib
simultaneously expresses three different morphemes: the
consonantal root /ktb/ 'write', the vowel melody /u_i/ 'perfective
passive', and the templatic morpheme CVCCVC 'Form 2' or
'causative/factive'. It is the templatic morpheme CVCCVC --
called a CV skeleton -- that accounts for morphological shape
invariance. The coordination of these three levels is by the
principles of autosegmental phonology (Goldsmith 1976, Clements
and Ford 1979), so the form kuttib is represented as follows:

(4)
Vowel Melody	u i	'perfective passive'
	| |	
CV Skeleton	CVCCVC	'causative (Form 2)'
	| |/ |	
Root	k t b	'write'

The Form 2 template is, in CV skeleton theory, a string of
segment-sized units C and V. The Prosodic Morphology Hypothesis
requires that a very different vocabulary be used to characterize
templates like this one: it is a sequence of two heavy syllables.
In prosodic morphological terms, then, kuttib is represented as:

(5)

(This will later be refined somewhat.) Since moras are the
prosodic unit of syllable weight, a syllable dominating two moras
is heavy, like the syllables kat and tab of the Form 2 verb
kattab.

What are the differences between these two seemingly equivalent characterizations of the Form 2 template? First, the prosodic characterization is motivated in prosody. This is not a mere tautology, since independent motivation for the segment-sized units of CV skeleton theory is difficult to come by and often, if not always, subject to plausible reanalysis. Unambiguous evidence for segment-sized skeletal units is nonexistent. Second, as we will see below, prosodic templates reveal connections with other aspects of Arabic prosody, particularly foot structure and minimality, that could not be obtained from a CV template. Third, the Prosodic Morphology Hypothesis often forces the correct analysis in cases where CV skeletal theory is confronted with an array of incompatible and inadequate options, as we show in our study of the Arabic broken plural (McCarthy and Prince 1990). Finally, prosodic morphological theory is more restrictive than CV skeletal theory (since the units of prosody are needed independently in either theory), and is therefore to be preferred to it on general grounds of parsimony and learnability.

4. Moras and Extrametricality

Moraic theory provides us with certain basic tools for characterizing the syllable types of a language. A syllable normally may contain one mora or two; a monomoraic syllable is called light and a bimoraic one heavy. Peripheral elements -- those at the left or right edge of a stem, word, or other domain -- may be extrametrical, not participating in the overall prosody of a word.
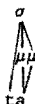
In medial position, where extrametricality is not a factor, Standard Arabic has just three types of syllables: Cv (ta), Cvv (taa), and CvC (tab). On the basis of cross-linguistic comparison, our normal expectation is that Cv syllables are light or monomoraic while Cvv and CvC syllables are heavy or bimoraic. We therefore represent these syllable types as follows:
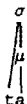
(6)
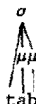a. Light Cv   b. Heavy Cvv   c. Heavy CvC

The representations in (6) segregate syllables into two classes, lumping Cvv and CvC syllables together as bimoraic. Evidence for this classification is abundant.

The first set of arguments comes from an aspect of prosody, the classical system of versification. In Al-Xalil's analysis of this system, a fundamental distinction is made between two kinds

consists of any syllable, heavy or light. But a peg is composed
specifically of a sequence of a light syllable followed by a heavy
one. (The peg, then, is an iambic foot.) In other words, the
characterization of a peg must count moras: it is a syllable with
one mora followed by a syllable with two moras (with the usual
moraic equivalence of the heavy CvC and Cvv syllables). An even
better case for moras comes from the phenomenon of resolution,
which appears in the meters called kaamil and waafir. In these
meters, in certain positions in the verse the poet may use either
two light syllables or a single heavy syllable (mutafaaSilun and
mufaaSalatun are the mnemonic examples). This too is an instance
of mora counting -- since a light syllable occupies one mora and a
heavy syllable occupies two, the equivalence between two light
syllables and one heavy syllable is precisely what is expected.
Finally, the traditional theory of the rhyme in poetry and rhymed
prose (sajS) relies crucially on the notion "heavy syllable":

(7) Traditional Typology of Rhymes

Rhyme Type   Rhyming Words

mutawaatir   Zulmi, siHru, šaybaanaa, Zunuunii
mutadaarik   ?al-mubaasilu, yazuurahaa, haykali
mutaraakib   walaa faraqaa, qad Husiduu

Each of the four types of rhyme is distinguished from the others
by the position of the rightmost nonfinal heavy syllable in the
verse: in mutawaatir it is the penultimate syllable, in mutadaarik
it is the antepenultimate, and in the rare mutaraakib it is the
preantepenultimate.[1]

        Another aspect of prosody, stress, leads to exactly the same
conclusion as the rhyme facts do. There is considerable
discrepancy in the stressing of standard Arabic words between
different areas of the Arab world, and no direct testimony on this
subject exists from the Classical period. Nevertheless, a
plausible norm with wide geographic and ethnic distribution is
represented by the data in (8):

(8) Stress Placement

Final	Penult	Antepenult
yaquul	yaquulu	kataba
qaanuun	yaqulna	katabat
sirHaan	qaalat	katabuu
tarjamt	ramaa	

---

[1] Another, rare rhyme type, mutaraadif, is distinguished by final
CvvC or CvCC syllables.

(9)

a. Stress the final syllable if it is CvvC or CvCC.
b. Otherwise stress the penultimate syllable if it is Cvv or CvC
(or the word is disyllabic).
c. Otherwise stress the antepenultimate.

Leaving aside the cases of final stress in (9a) (which occur only pre-pausally in Classical Arabic, under loss of case and agreement desinences), the basic observation is that the penult is stressed if it is bimoraic, otherwise stress falls on the antepenult.

   In the typology of metrical stress feet introduced by Hayes (1987) and McCarthy and Prince (1986), based on Hayes's (1985) survey, this type of stress pattern is derived by a foot called the moraic or quantitative trochee. This foot type contains exactly two moras and is stressed on the left:

(10) Quantitative Trochee

$$\begin{array}{c} F \\ \bigwedge \\ \mu \ \mu \end{array}$$

For purposes of stress assignment, final (light) syllables are not included in the application of this foot. They are therefore extrametrical with respect to foot assigment.[2]

   The segmental phonology of Arabic also provides direct evidence of the light/heavy distinction. Consider first a process originally described in generative phonological terms by Brame (1970). The alternation in vowel length in hollow verbs[3] in (11a) is a typical example of a well-established phonological phenomenon: vowel shortening in closed syllables. The derivation proceeds as in (11b):

---

   [2] An interesting case is presented by forms like tárjama 'he translated', with a heavy antepenult followed by a light syllable. With final extrametricality, the metrical portion of the word is tarja. A final bimoraic foot cannot be placed on this word to give ta[rja]$_F$ because this would violate the prosodic hierarchy. A final monomoraic foot (tar[ja]$_F$) is impossible, because the quantitative trochee is exactly two moras. Therefore the right-to-left operation of foot-assignment must move on to yield [tar]$_F$ja, correctly resulting in antepenultimate stress.

   [3] Hollow verbs are those whose medial root consonant is a high glide -- /qwl/ in 11. The w appears overtly when geminate (qawwaal 'garrulous') or syllable-final (qawl 'word'); it is otherwise usually subject to complex morphophonemic processes which will not be discussed here.

(11)
a. qaala   qultu   'he/she said'
     yaquulu   yaqulna   'he/they (f) say'
b.
     /yaquul+u/   /yaquul+na/ Underlying Form
     [yaquulu]   [yaqul+na]   Shortening Rule

The rule of vowel shortening in closed syllables has a
straightforward interpretation in moraic terms. The heaviest
syllable in Arabic is one with two moras, and a long vowel
occupies both of them. But a CvC syllable also has two moras.
Taking these two facts together, it is apparent that a long vowel
should normally be incompatible with a syllable-final consonant.
In Arabic and usually elsewhere, the length of the vowel gives way
under the pressure of the consonant that would otherwise be
unsyllabifiable.[4] Example (12) shows in a somewhat informal way
what happens:

(12)

Input Form   ya qu l na (= /yaquul+na/)

Derived Form   ya qul na (= [yaqul+na])

Moraic theory provides a straightforward account of this common
phonological rule.[5]

_____

[4] Makkan Arabic (Abu-Mansour 1987:163) takes another option. Instead
of shortening the long vowel before an unsyllabifiable consonant, it
epenthesizes a vowel: /muftaaH+kum/ -> muftaaHakum 'your (pl.) key'. The
third logical possibility, loss of the unsyllabifiable consonant after a
long vowel, is attested in no language known to us.

[5] There is an interesting aspect to the treatment of CVVC in word-
final position. The jussive of yaquulu is yaqul, with vowel shortening in
a closed syllable. But the pausal form of yaquulu is yaquul, which retains
the long vowel, as predicted by the conditions on extrasyllabicity
developed below. There is evidence (from the jussives of III-w, y roots)
that the jussive is formed by a morphological truncation of the final
vowel, and this may be responsible for the lack of final extrasyllabicity
in jussives.

Another sort of phonological process that provides similar evidence for the mora in Arabic is the phenomenon of compensatory lengthening, in which deletion of a consonant is compensated for by lengthening of the preceding vowel. (Compensatory lengthening is treated briefly in McCarthy and Prince (1986); a recent comprehensive review of the topic in moraic terms is provided by Hayes (1989).) In Arabic, compensatory lengthening is quite common; it shows up most obviously in the derivation of Form 4 (ʔafʕala) from roots whose first consonant is ʔ. Example (13) contains the evidence and an informal statement of the rule:

(13)
a.
Underlying   ʔaʔθar (Form 4 (ʔaCCaC) of root /ʔθr/)
Derived   ʔaaθar

b. ʔ Deletion
ʔ -> φ / ʔV__

In moraic terms, compensatory lengthening is simply an exchange of one type of heavy syllable for another -- the moras remain the same but the segments associated with them change. This is shown in (14):

(14)

Underlying Form

Derived Form

Deletion of the ʔ leaves a mora stranded; this mora is then filled by spreading of the vowel a. The equivalence of the two types of heavy syllables is apparent in this example.

        In summary, these arguments all point toward the central importance of the notion mora in Arabic. The evidence indicates a fundamental classification into light (monomoraic) Cv syllables and heavy (bimoraic) Cvv and CvC syllables. This typology, though, holds only of syllables in medial position. Initially or finally, extrametricality provides a richer array of options. Let us consider, then, what is special about initial or final position.

initial sequences of two consonants occur. These appear in verb
forms and their derivatives that have what is traditionally called
<u>hamzatu l-waSli</u>, the "elideable" glottal stop. Examples include
Form 7 <u>ʔinfaʕal</u>, Form 8 <u>ʔiftaʕal</u>, and Form 10 <u>ʔistafʕal</u>. The
distribution of this property forces any generative phonological
analysis to say that the initial glottal stop and the vowel
following it are not in fact elided, but rather inserted in the
course of syllabification. For example, the underlying
representation of the Form 8 stem is <u>ftaʕal</u>, although on the
surface this word in isolation is pronounced as <u>ʔiftaʕal</u>.

        The following examples show what happens to this form in
different phonological contexts within the utterance or major
phonological phrase:

(15) The Phenomenon of <u>hamzatu l-waSli</u>

a. Postpausally (that is, utterance initially)

    σ σ σ
    Λ Λ Λ
    ʔiftaʕal

b. Postconsonantally

    σ σ σ σ
    Λ Λ Λ Λ
    qad iftaʕal

c. Postvocalically

    σ σ σ σ   σ σ σ
    Λ Λ Λ Λ   Λ Λ Λ
    huwwa ftaʕal   waftaʕal

The examples indicate the syllabic affiliation of every segment in
the three possible phonological contexts. In postvocalic
contexts, underlying <u>ftaʕal</u> emerges unchanged. In postconsonantal
contexts, a triconsonantal cluster is broken up by an epenthetic
vowel -- <u>i</u> before <u>a</u> or <u>i</u>, <u>u</u> before <u>u</u>. Postpausally, the initial
biconsonantal cluster of <u>ftaʕal</u> requires an epenthetic vowel, and
the epenthetic vowel itself requires a preceding consonant, <u>ʔ</u>,
since all Arabic syllables must begin with a consonant. The
appearance of <u>ʔ</u> and the epenthetic vowel are fully predictable
from the underlying representation <u>ftaʕal</u>. For that reason we
cannot speak of elision, but rather of insertion.

        Having established the existence of underlying
representations with initial consonant clusters, we must now
integrate them into the moraic model. Since this phenomenon is
limited to stem-initial position, extrametricality, inherently
restricted to the periphery (Hayes 1982, Harris 1983), is the
mechanism that presents itself.

As a first approximation, the initial consonant of ftaʃal can be analyzed as an extrametrical mora, one that is not linked to any syllabic or extrasyllabic constituent, symbolized by parentheses in representations like the following:

(16)


```
        σ   σ
        |   |
(μ)  /μ  /μμ
     | /|  /||
     f ta  ʃal
```
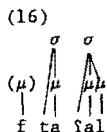
When f becomes intrametrical, either by prefixation or in the postlexical phonology, it remains moraic but must be fully integrated into a complete syllabic structure. A preceding vowel supplies that in (15c), while epenthesis is necessary in (15a, b). In all cases in (15), the f is indeed in a moraic position, closing a heavy syllable.

But this by itself is not proof that f is linked to an extrametrical mora in underlying representation -- the phonology of Standard Arabic (specifically, the epenthesis rule) could simply stipulate the position of the epenthetic vowel, since there is little evidence for epenthesis elsewhere in the language. This is emphatically not the case in Egyptian Colloquial, however. As Broselow (1976) and Selkirk (1981) have shown in considerable detail, the treatment of otherwise unsyllabifiable consonants in Egyptian Arabic follows a very regular pattern in which a vowel is always inserted after the consonant, as (17a, b) show. Indeed, as Itô (1986) shows, a single parameter in the grammar of Egyptian Arabic, left-to-right syllabification, accounts for this consistent placement of the epenthetic vowel relative to the otherwise unsyllabifiable, or "stray", consonant.

(17) Egyptian Colloquial Treatment of Stray Consonants

a. Vowel Insertion in CC___C Context
    /katabtlu/ -> katabtilu

b. Vowel Insertion in #C___C Context in Loans
    plastic -> bilastik

c. Vowel Insertion in #___CC Context in Templatic Verbs
   /gtamaʕ/ -> igtamaʕ (-> ʔigtamaʕ)

It is puzzling, then, that just the opposite treatment is accorded the initial extra consonant in templatic Form 8 verbs like (17c). Yet this is exactly what is expected if the g of ʔigtamaʕ is an underlying extrametrical mora -- it must remain a mora despite the normal vowel insertion process of the language.

This evidence shows, then, that the extrametrical initial
mora has some independent motivation. Prosodic theory forces us
to posit this analysis in both Standard Arabic and the Cairene
colloquial. Other facts of Standard Arabic are consistent with
this approach, but they do not prove its correctness. But in
Cairene, the difference between epenthesis in templatic verbs and
elsewhere in the language requires the extrametrical initial mora.

At the right edge of stems, we also find both a more limited
and a richer structure than the Cv, CvC, and Cvv medial syllables
would allow. All stems of Arabic must end in a consonant; thus,
Cvv and Cv stem-final syllables are prohibited. Furthermore, noun
stems can end in CvCC (baHr 'sea') or CvvC (qaamuus 'ocean'), with
a heavy syllable followed by an extra consonant. Stem-finally,
then, the only permitted sequences are CvC, CvCC, and CvvC. (In
word-final position, because of affixes, Cv, Cvv, and prepausally
CvvC are also permitted.)

The licit stem-final sequences can be analyzed as a sequence
consisting of any possible medial syllable followed by an
obligatory consonant: Cv+C, CvC+C, and Cvv+C. The obligatory
stem-final consonant is plausibly analyzed as extrametrical but
not as moraic, since it becomes an onset before vowel-initial
suffixes or words: cf. (15b) and katab+a 'he wrote', qaamuus+un
'ocean (nom. indef.)'. We might, then, regard the position to
which this consonant is linked as an extrametrical final syllable,
as in the following representations:

(18)

ka ta b   baH r   qa mu   s
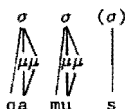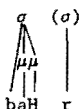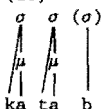
(The examination of biliteral roots in the next section will show
why final consonants must be linked to a skeletal position; they
cannot simply float.) For stems, the final extrametrical syllable
is required; thus, all stems must end in a consonant. The
following rule records this:

(19) Final Incompleteness

$\phi \rightarrow (\sigma) \, / \, \_\_\_]_{Stem}$

Thus far, we have a significant asymmetry between initial
and final position. Initially, we posit an extrametrical mora to
bear the parenthesized extrasyllabic consonant of forms like
(f)taʕal. Finally, there is an extrametrical syllable to bear the
extrasyllabic consonant of kata(b), baH(r), and qaamuu(s). The
unifying observation is that the inital extrasyllabic consonant

has a characteristic property of syllable-final position -- it is moraic -- while the final extrasyllabic consonant is not moraic, as if it were syllable-initial.

This paradoxical behavior of extrasyllabic consonants has been noted before, originally by Itô (1982: 13-14) for Russian and later by Borowsky (1986: 197-199) for English. In these languages which, unlike Arabic, allow complex onsets and codas, the permitted sequences of final extrasyllabic consonants are just exactly the permitted onsets and the permitted sequences of initial extrasyllabic consonants are the permitted codas. Itô's general schema for a Russian word, then, is (Coda)σ^*(Onset), where σ^* denotes a string of zero or more syllables. The same can be said of an Arabic stem, except that the initial Coda appears only in certain morphological classes of the verb and the final Onset is obligatory in all stems.

We can now incorporate these results into prosodic theory. The fundamental insight is that syllables at the periphery of a stem, word, or other domain may be incomplete, consisting solely of a moraic consonant (a Coda) or a nonmoraic consonant (an Onset). Clearly these incomplete peripheral syllables are what we have been calling up until now extrametrical ones, and we can continue to denote them by (σ). We stipulate that (σ) is vowel-less (in Arabic), but derive its positional characteristics (initial (σ) is a moraic coda, final (σ) is a nonmoraic onset) from the following principle:

(20) Contiguity Constraint

Syllabic well-formedness is enforced over contiguous strings of subsyllabic elements.

The Contiguity Constraint entails that vowel-less syllables (σ) can be found only at the periphery of words, since a representation like $...[CvC]_\sigma[C]_\sigma[CvC]_\sigma...$ violates well-formedness with respect to the string CCC. Similarly, an initial or final $[CC]_\sigma$ syllable violates the Contiguity Constraint. Furthermore, a representation like $[C]_\sigma[Cv]_\sigma...$ respects the contiguity constraint if and only if the initial C is analyzed as a moraic position in the vowel-less syllable, because syllabic well-formedness demands that, in any heterosyllabic CC sequence, the first C be a mora. Likewise, $...[CV]_\sigma[C]_\sigma$ is well-formed if and only if the final C is nonmoraic (an onset), since it is syllable-initial and follows another syllable.

The Contiguity Constraint we take to be universal, although there is some variation in how it is enforced. Morphological templates like those of Arabic enforce it absolutely, but subsequent morphological (Archangeli 1988) or phonological (Itô 1986, 1989) developments may respect it only by requiring the

We tentatively suggest that the limitation of the
incompleteness property to vowel-less syllables is part of the
particular grammar of Arabic, rather than universal. There is
some evidence from other languages that initial onset-less
(therefore incomplete but voweled) syllables may also show
extrametrical behavior. In the Timugon Murut diminutive and
instrumental morphology (Prentice 1971, McCarthy and Prince 1986),
initial CV reduplication is the norm for consonant-initial words:
bulud 'hill', bu-bulud; dondo? 'one', do-dondo?. But vowel-
initial words disregard the entire first syllable, so the
reduplicative morphology is infixed: ulampoy (not glossed), u-la-
lampoy; indimo 'five times', in-di-dimo 'dim./inst.'; ompod, om-
po-pod 'dim./inst.'. And in Western Aranda (Strehlow 1942, Davis
1988, Archangeli 1986, Halle and Vergnaud 1987), main stress falls
on the first syllable if the word begins with a consonant (túkura
'ulcer') and on the second if the word begins with a vowel (ergúma
'to seize').[6] Both of these phenomena have been treated as
effects of syllable extrametricality; in our conception, this
correlates with the incompleteness of these peripheral syllables.

To sum up this discussion, we have argued that the treatment
of extrasyllabic consonants at the edge of Arabic stems is the
reflection of a far more general property, the option for
languages to have incomplete syllables at the periphery. We
obtain the particular properties of these elements in Arabic --
moraicity initially, non-moraicity finally -- from the universal
Contiguity Constraint in (20) and the stipulation that incomplete
syllables are vowel-less. These extrametrical syllables therefore
fit-in nicely with the general view of Arabic syllable prosody
offered here.

5. Minimality in Arabic

Since the word dominates the foot in the prosodic hierarchy,
the smallest word will be a single foot. We call a word, stem, or
other top-level category that exactly meets this criterion
minimal. Since Arabic requires quantitative trochaic stress feet,
the minimal stem (and therefore word) will be a single foot of
this type, or two moras. These two moras can be contained in a
single heavy syllable or distributed between two light syllables.
Final incomplete syllables, although required in Arabic stems, do
not contribute to the fulfillment of the minimal stem requirement,
and so they are in addition to the two moras required by

---

[6] Western Aranda disyllabic words nevertheless always have initial
stress.

minimality. Some typical examples of minimal stems - that is, forms with exactly two moras - appear in (21):

(21)

baH r   ba z   ka ta b
'sea'   'vulture'   'wrote'

Modulo final extrametricality, each of these forms minimally satisifies the two-mora requirement. They (and of course many others like them) are all uncontroversially words of the language, abstracting away from the addition of case and agreement affixes.

There are, however, a few apparent counterexamples to the minimal stem requirement. Some candidate words that are too small appear in (22):

(22) Apparent Monomoraic Words

Non-words	Biliterals	Imperatives
wa 'and'	?ab	li (imperative /wly/)
qad 'past'	bn	daʕ (imperative /wdʕ/)
bi 'in, with'		ktub (imperative /ktb/)

All of these forms have at most one mora by the criteria established; in fact, in the case of bn there are evidently no metrical moras at all in underlying representation (since this form contains no intrametrical syllables). These apparent counterexamples must obviously be dealt with.

Those in the category "non-words" in (22) are exactly that. All of these forms are in the so-called nonlexical vocabulary -- they are not members of the major lexical categories noun, verb, and adjective.[7] Cross-linguistic investigation reveals that nonlexical vocabulary rarely has the phonological or grammatical properties of ordinary nouns and verbs. (For example, the only "words" of English beginning with voiced th are non-lexical: the, this, that, thou, then, etc.) The minimal word constraint is a prosodic constraint on the lexicon: it therefore does not apply to nonlexical vocabulary. Furthermore, there is no reason to believe that these nonlexical words are independent prosodic words in any case.

---

[7] The traditional conception of "word" implicit in the orthography evidently counts letters; wa and bi require a single letter and are written as prefixes, while laa and qad require two letters and are written as separate words.

Cvv words like laa 'no' do respect the two-mora minimality requirement, they run afoul of another requirement on Arabic stems -- they must end in a consonant by (19), a constraint that is required independently of minimality.

The examples in the second column of (22) are lexical vocabulary items -- they are nouns -- but they too are not compelling evidence against the minimal word requirement. The reason is that they come from a very small, closed class of items that never reflected a productive pattern of the language. In the 1-mora class with ʔab are Ham, ʔax, dam, fam, and yad. And in the 0-mora class with bn is sm.[8] These lists are exhaustive, so the numbers are obviously quite small. Moreover, these words are quite irregular whenever they participate in any of the truly productive morphology of the language, and the irregularities they display always make the stem larger, so that it satisfies the minimality requirement. ʔab, for instance, receives an added aw in the dual and nisba (a productive denominal adjective obtained by suffixing iiy): ʔabaw+aan, ʔabaw+iiy. It also has this w in the plural /ʔaʔbaaw/ (which becomes ʔaabaaʔ by regular phonological rules). And, interestingly, it lengthens the case suffix (making the word bimoraic) in the definite singular: ʔal-ʔabuu, ʔal-ʔabii. These observations indicate that these monomoraic words are in fact exceptional in nearly all respects; it is no surprise, then, that they are exceptions (rather than counterexamples) to the minimality requirement.

The examples in the third column of (22) are all imperatives. Traditionally, imperatives are special in two respects, both of which involve morphological truncation or deletion processes. First, the imperative, like the jussive, deletes the final vowel of the indicative imperfective. Second, the imperative is derived from the jussive by deleting the agreement prefix. What has happened in these forms is that application of these morphological truncation processes creates the apparent violation of the minimal word constraint. The constraint, then, must be enforced at a relatively early stage of the derivation before these truncations.

Apart from these basic observations, there are at least four other arguments in support of the bimoraic minimal stem in Arabic. First, it is clear that CvC stems like ʔab are abnormal even when

---

[8] Interestingly, all of these except for sm 'name' are words for near kin (ʔab 'father', Ham '(woman's) father-in-law', ʔax 'brother', bn 'son') or body parts (dam 'blood', fam 'mouth', yad 'hand'). In many languages, such words, including 'name', are inalienably possessed, requiring possessive pronouns. If this situation obtained in an earlier, unattested stage of Arabic, it would account for the violation of minimality here; with possessive suffixes, all these words would be at least two moras long.

the root is biconsonantal. Versus the tiny number of words like
?ab, whose lexical material contains only one monosyllable like
barr, buzz, or tall, in which biliteral roots like /br/ or /bz/
must satisfy the minimum of two metrical moras via gemination of
the final radical. For this reason, too, the bimoraic minimality
requirement is not reducible to counting root consonants, as
traditional accounts would have it. If all Arabic roots had three
consonants, as the tradition assumes, then a CvCC/CvCvC minimum
would follow simply from the need to find positions for all of
them. But biliteral roots are a prominent feature of the Arabic
lexicon (McCarthy 1979, 1981, 1986), and so the prosodic
requirement of bimoraicity is essential.

Second, many roots whose initial consonant is w lose this w
in the masdar (a kind of nominalization) by a partly phonological
rule, as (23) shows:

(23)

Perfective	Masdar	
waθiq	θiq+at	'rely'
wadaʕ	daʕ+at	'put'
wada(y)	diy+at	'pay wergild'
wariθ	riθ+at	'inherit'
wazan	zin+at	'weigh'
wasiʕ	saʕ+at	'be wide'

The problem is why just these masdar forms require the feminine
suffix +at.

The obligatory feminine suffix in these masdars is explained
by the bimoraic minimum. With the loss of the root-initial w, a
form like θiq is simply too small, since it contains only a single
mora. Addition of the feminine suffix augments it to make it
bimoraic, as (24) shows:

(24)
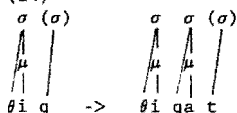
The traditional idea (Wright 1971: 118) that the feminine suffix
compensates for the loss of the first radical is expressed
formally by the bimoraic minimal stem requirement. An additional
bit of evidence in support of this analysis comes from the nisba
derived from these masdar forms. Since the feminine suffix can
never precede the nisba suffix, the feminine suffix must be lost.
The result is that the base is then too small. This problem is
resolved by introducing a final aw, just as in the case of ?ab:
perfective waʕad 'make a promise', imperfective yaʕid, masdar
ʕid+at 'a promise', nisba of masdar ʕidaw+iy. 'promissory'

of this line of evidence in support of the bimoraic minimal
word comes from the treatment of borrowed words that would
otherwise be too small. A few examples, gathered at random,
appear in (25):

(25)
Source   Arabicized form
bar   baar
jazz   jaaz
gas   gaaz
Shem   saam
Gaul   gaal
shawl   šaal

Words that would be monomoraic when borrowed into Modern Standard
Arabic are made bimoraic, satisfying minimality, by lengthening
the vowel. Along the same lines, Broselow (p.c.; cf. Broselow
1982: 124) observes that the English word bus, which would be
monomoraic in Arabic, is borrowed into the Palestinian colloquial
variously as baas, bass, and basi, all bimoraic. In fact, Smeaton
(1973: 87), in his comprehensive treatment of loan words in a
Saudi Bedouin dialect, proposes a rule of Arabicization by which
all CvC monosyllables are borrowed with gemination of the final
consonant: baSS 'bus', natt 'nut', rigg 'rig'.

     Similar regularities are even more profoundly integrated
into the phonology of the modern Arabic dialects. Broselow 1982
notes that in Iraqi Arabic initial epenthesis is obligatory for
sub-minimal CCvC imperatives but optional in longer ones. And
Kenstowicz 1981 has argued that vowel-length alternations observed
in Lebanese Arabic imperatives like ktoob 'write! (m. sg.)', ktibu
'write! (pl.)' also demonstrate a two-mora minimality requirement.

     A final phenomenon demonstrating the role of the bimoraic
minimal word is found in the remarkable behavior of the truncated
vocative. Cross-linguistically, truncated hypocoristics or
vocatives (nicknames) often are based on the minimal word or,
equivalently, the foot (McCarthy and Prince 1986, forthcoming).
Arabic has truncated vocatives occasionally in classical verse
(though not in the contemporary literary language). These are
discussed by Wright (1971:2.88) and Howell (1986:1.1.191-4).
Representative data, all proper nouns, appear in (26):

(26)
a. CvvCvC nouns

maazin	maazi
maalik	maali
ʕaamir	ʕaami
Haariθ	Haari

b. CvCvvC nouns

suʕaad	suʕaa
majiid	majii
θamuud	θamuu

c. CvCCvC nouns

jaʕfar	jaʕfa

d. CvCCvvC nouns

ʕuθmaan	ʕuθma
marwaan	marwa
manSuur	manSu
miskiin	miski

Smaller nouns -- those with stems CvCC or CvCvC -- do not form
distinctive truncated vocatives. This is to be expected, if the
truncated vocatives are based on the minimal word; CvCC and CvCvC
stems are already minimal.

The most interesting contrast in (26) is between CvCvvC and
CvCCvvC stems; the former retain the length of the final vowel in
the truncated vocative, as in majiid/majii, while the CvCCvvC
forms do not, as in marwaan/marwa. The source of this difference
is clearly the weight of the initial syllable -- light in CvCvvC
and heavy in CvCCvvC.

If the minimal stem is bimoraic, then the truncated vocative
is a minimal stem followed by a vowel: [maji]i, [marw]a. Since
the vocative "stem" is always followed by a vowel, it is not
subject to Final Completeness (19). The vocative vowel is not some
arbitrary appurtenance to the bimoraic template. Rather, it is a
kind of simulation of the normal case-marking final short vowel
(usually the nominative +u) that untruncated vocatives have: yaa
Haariθ+u, yaa jaʕfar+u (yaa is the vocative particle). In fact,
the final vowel of the truncated vocative may assume the melody of
the nominative case-marking: yaa Haar+u, yaa jaʕf+u. Thus, the
truncated vocatives are minimal words to which the appearance, and
sometimes the reality, of normal vocative nominative case-marking
is added.

Let us now summarize the discussion up to this point. We
have a characterization of moras and extrasyllabicity in Arabic,
and we have seen how these notions play a role in the minimal word
constraint. Now we will turn to the templatic morphology, first
of the noun, then of the verb.

The problem now to be examined is the characterization of
shape-invariance in Arabic morphology. The analysis begins with a
look at the basic stem structures -- not including prefixes or
suffixes -- in the noun. We will have little to say about the
broken plural, which we have dealt with extensively elsewhere
(McCarthy and Prince 1988, 1990).

(27) contains a list of representative underived nouns of
all possible basic patterns. Since our concern is with overall
shape or canonical pattern here, differences in vowel quality have
been disregarded. The percentages in parentheses below each word
give a rough idea of how common each of these canonical patterns
is; they were obtained by counting all the nouns that form broken
plurals in the first half of Wehr's (1971) dictionary (about 2400
words):

(27) Basic Nominal Patterns

			Mora Count		
	Two		Three		Four
Biliteral root	barr (6%)	sabab (<1%)	jadiid (3%)	baarir (<1%)	jaaruur (<1%)
Triliteral root	baHr (27%)	badal (7%)	?ataan (18%)	kaatib (12%)	jaamuus (2%)
Quadriliteral root				xanjar (14%)	rasmaal (11%)

CvvC nouns like baab 'door' are all arguably derived from
underlying /CvGvC/; compare the plural ?abwaab.

The noun patterns have been sorted into columns according to
their canonical pattern and into rows according to the number of
root consonants. In the left two columns are words that are
exactly minimal -- with the final consonant extrasyllabic they
have only two moras subject to metrical scansion. In the middle
columns are the patterns that contain three moras. These three
moras can be divided among two syllables either as light-heavy
(?ataan) or heavy-light (kaatib). At the extreme right are the
noun patterns with both syllables heavy. These observations about
the moraic and syllabic composition of the forms in (27) are
summarized in (28):

McCarthy and Prince: Prosodic Morphology and Templatic Morphology

	Two Moras		Three Moras		Four Moras
barr	sabab	jadiid	baarir	jaaruur
baHr	badal	?ataan	kaatib	jaamuus
			xanjar	rasmaal

$\sigma(\sigma)$ $\quad\sigma\;\sigma(\sigma)$ $\qquad\sigma\;\;\sigma\;(\sigma)$ $\qquad\sigma\;\;\sigma\;(\sigma)$ $\qquad\sigma\;\;\sigma\;(\sigma)$

CvCC   CvCvC   Cv CvvC   CvCCv C   CvCCvvC

What the analysis must now explain is why these are the
possible basic noun patterns and no others are. Part of the
explanation comes from the minimal stem constraint -- it sets a
lower limit on stem size of two metrical moras that all noun
patterns must respect. The other half of the explanation comes
from the rule stated in (24):

(29) Maximal Stem Constraint

Templates are maximally disyllabic.

In other words, since all canonical noun stems are formed on
templates, no stem can exceed two intrametrical syllables. This
seemingly arbitrary disyllabic upper bound on templates is in fact
not arbitrary at all: as we show in McCarthy and Prince (1986,
forthcoming), general considerations of locality in linguistic
theory require that no rules count to greater than two. The rules
specifying the Arabic templates are subject to locality, and so
the Maximal Stem Constraint can be obtained from a principle of
much wider application.

The minimality constraint, on the one hand, sets a lower
bound in terms of moras; the maximal stem constraint, on the
other hand, sets an upper bound in terms of syllables
(incidentally showing that both levels of representation are
required). Between these two extremes, everything that is
possible is actually being quite heavily used by the language.
With the additional requirement in (19) that all stems end in an
incomplete syllable (equivalently, a consonant), (28) contains
everything expected given these two constraints. In a sense, this
is the ideal situation, where the analysis accounts for just
exactly the phenomena that the language displays.

There is some independent evidence for the Maximal Stem
Constraint, just as there is for the minimality constraint.
First, there is the problem of distinguishing diptotic from

(30) Diptotic Broken Plural Patterns

Pattern	Example
CuCaCaa?	xulafaa?
?aCCiCaa?	?aqribaa?
CaCaaCiC	jadaawil
CaCaaCiiC	Sanaadiiq
CawaaCiC	bawaaSiθ
CawaaCiiC	jawaamiis
CaCaa?iC	jazaa?ir

What the diptotic broken plurals all have in common is that they
are stems with three syllables -- one greater than the maximum.
(How they get that way while still respecting the requirements of
locality and the Prosodic Morphology Hypothesis is a topic treated
in McCarthy and Prince (1988, 1990).) In other words, only those
stems that do not exceed the maximality constraint are triptotic.
There are other, nonphonological criteria by which a noun can be
diptotic, but the maximal stem constraint provides an explanation
for why just this set of broken plurals should be consistently
diptotic.

    Second, the maximal stem constraint predicts that singular
noun stems of three syllables should be rare and extremely
irregular in their behavior, like the sub-minimal nouns like ?ab
or bn. A sample of some of these super-maximal noun stems appears
in (31):

(31) Trisyllabic Noun Stems

namuuδaj
Sankabuut
safarjal
barnaamaj

Such words are quite rare. More important, however, is the fact
that they are very irregular as well. The most important aspect
of their irregularity is that they are not templatic: they display
no regularities of formation other than respect for the

_____

    [9] The regular or triptotic ('three'+'case') declension of Arabic
distinguishes three cases in the singular, marked by suffixes +u
nominative, +i genitive, and +a accusative. The diptotic declension has
only +u nominative and +a genitive-accusative. The diptotic declension
also lacks the suffix +n that marks indefinite nouns. Apart from the
broken plurals discussed in the text, diptotic declension is restricted to
certain adjectival patterns and some proper nouns. The monoptotic or
indeclinable declension is phonologically explicable.

McCarthy and Prince: Prosodic Morphology and Templatic Morphology

phonotactics of the language. Another sign of the irregularity
appears with older words like those in (31). Although these nouns
form broken plurals and diminutives, they do so only with very odd
conditions on the treatment of vowel length and supernumerary
consonants. These are discussed in detail in McCarthy and Prince
(1988, 1990). The other sign of irregularity comes from more
recent loans that have three syllables, like tilifuun. In the
half of Wehr's (1971) dictionary that we have examined, no recent
trisyllabic loan ever forms a broken plural; instead, they have
the sound plural tilifuun-aat. This is a powerful indication of
how irregular these super-maximal words are, since the broken
plural system is otherwise so productive that it very quickly
assimilates borrowed words, like bank/bunuuk, malyuun/malaayiin,
and so on. The super-maximal noun stems -- those with three
syllables -- are quite clearly outside the Arabic morphological
system, as the Maximal Stem Constraint predicts.

The moraic and syllabic skeleta in (28) do not exhaust the
insights obtainable from an examination of the basic nominal
patterns. Other, more suprising results of prosodic analysis also
emerge.

The Prosodic Morphology Hypothesis asserts that templates
are composed of the units of prosody. The skeleta in (28),
composed as they are of the prosodic units mora and syllable,
satisfy this condition only weakly. But interpreted strongly, the
Prosodic Morphology Hypothesis requires templates that are exactly
specified by a single prosodic constituent. We have already seen
how the nouns in the two-mora class satisfy this requirement.
These nouns are minimal words, whose template is identical to a
prosodic constituent, the quantitative trochaic foot. In other
words, under the Hypothesis, the desirable equation is Template=P,
where P is any prosodic constituent.

There is an important and unexplained asymmetry between the
two types of trimoraic stems, CvvCvC baarir/kaatib and CvCvvC
jadiid/?ataan. Our statistical investigations reveal that CvCvvC
nouns are considerably more common and diverse than CvvCvC nouns.
The data are summarized in the following table:

(32)

CvvCvC			CvCvvC	
CaaCiC	263		CaCiiC	265
CaaCaC	7		CiCaaC	106
CaaCuC	1		CaCaaC	37
			CaCuuC	29
			CuCaaC	25
			CiCiiC	1
Total 271			Total	463

CvCvvC stems are much more common and occur in many more vocalic patterns in a more even distribution than CvCvCvC stems. On deeper analysis, the skew turns out to be even worse than this -- all CaaCiC nouns, constituting 97% of the CvvCvC class, owe their existence to a single morphological process, the formation of the Form I active participle (kaatib) from the corresponding finite verb (katab). Apart from this single source, there are practically no CvvCvC stems, while the iambic stems are abundant and diverse.

The explanation for this dramatic skew comes from the Prosodic Morphology Hypothesis. Pursuing the implications of Hayes's (1985) typological study, McCarthy and Prince (1986) and Hayes (1987) propose that there is a fundamental structural distinction between iambic and trochaic feet: the iambic foot is asymmetrically light-heavy, but the trochaic one consists of two equal parts -- two moras in the case of Arabic. On this view, the mirror-image symmetry of CvCvvC and CvvCvC is linguistically meaningless; the two have incommensurable prosodic structures. The form [$_F$CvCvvC] is an entire iambic foot (with the final consonant extrasyllabic), but CvvCvC is a bimoraic (trochaic) foot plus something more: [$_F$Cvv]CvC. The anti-iambic form [$_F$Cvv]CvC cannot be analyzed as a single prosodic constituent. It is therefore excluded from the list of nonderived stem types by the Prosodic Morphology Hypothesis.

The morphology shows that CvvCvC is indeed a derived stem type. Since it occurs in the noun system almost exclusively as the active participle of the CvCvC Form 1 verb, participial CvvCvC can be derived from finite CvCvC by prefixation of a mora, lengthening the initial vowel:

(33) Form 1 Active Pariciple

$$[\, _\sigma$$
$$\mu$$

The finite verbs that are also heavy-light, like Form 3 CvvCvC, are derived as well; as we show below in section 7, they are composed of a heavy syllable base and a light syllable suffix, the latter marking them as finite. In the language as a whole, there is no role for the prosodically incoherent CvvCvC sequence as a primitive, underived template.

Before pursuing these matters, it is worthwhile to develop further the role of iambicity in the system. The contrast between general CvCvvC and restricted CvvCvC nouns lies in the fact that the former can be analyzed by an iambic foot (a light syllable followed by a heavy syllable), but the latter are unanalyzeable with the independently motivated constituents of prosody. But the prosodic analysis has also claimed that the foot type required by the stress system, the minimal word, and the characterization of

McCarthy and Prince: Prosodic Morphology and Templatic Morphology

the bimoraic noun stems is the quantitative trochee (two moras,
either in one heavy syllable or two light syllables). There is no
contradiction here. Universal grammar supplies a small vocabulary
of possible foot types among which languages are free to choose.
Although perhaps ideally a language would refer to a single foot
type in all rules of morphology or prosody, nothing in the theory
requires this. In Arabic, stress and the minimal word rely on the
quantitative trochee, but the broken plural (McCarthy and Prince
1990) and the system of versification are iambic. The basic noun
templates draw from both types: the quantitative trochees are CvCC
and CvCvC, while the iamb is CvCvvC.

        Along similar lines, our investigations have revealed a
hitherto unnoticed fact about "medial geminate quadriliterals",
triliteral nouns like jabbaar (from the root /jbr/) which have the
quadriliteral pattern of rasmaal but with medial gemination. The
basic observation is that CvC_iC_ivC with medial gemination is quite
rare, while CvC_iC_ivvC with medial gemination is common by
comparison. There are two sources of evidence for this. First,
among nouns taking broken plurals in Levy's (1971) comprehensive
study of the Wehr dictionary, there is the following distribution:

(34) Medial Geminate Quadriliterals
CvC_iC_ivC CvC_iC_ivvC

CuC_iC_iaC 6 CaC_iC_iaaC(+at)   60
CuC_iC_iaC+at   2 CaC_iC_iuuC 34
                                CvC_iC_ivvC+at   15

Total   8   Total   109

Clearly the vast majority (93%) of medial geminate quadriliterals
have a long vowel in the second syllable. Second, CvC_iC_ivC is not
used by the derivational morphology of the noun, but CvC_iC_ivvC is
heavily, productively used in the noun of profession or habitual
action: kallaaf 'stablehand', kawwaay 'slanderer'. Since such
nouns do not take broken plurals, they do not bias the statistics
above. In contrast, CvC_iC_ivC plays no role in the derivational
morphology of the noun. (This pattern is important in the verb,
but only as a derived template. See section 7.)

        Again there is a significant skew between two seemingly
equivalent patterns. The explanation is similar to the earlier
one. CvC_iC_ivC is a heavy syllable followed by a light one (as
always, assuming final consonant extrasyllabicity), a prosodically
meaningless configuration. CvC_iC_ivvC is composed of two heavy
syllables, equivalent to a sequence of two minimal words or two
quantitative trochaic feet. Here, then, Template=P^2, or perhaps
P^*, up to the limit of two imposed by locality considerations
(McCarthy and Prince 1986.

is the proper description of CvCC, CvvC, and CvCvvC nouns. CvCvvC nouns are formed by a single iambic foot. A sequence of two trochees yields the CvC_iC_ivvC pattern. Note that two trochees in sequence must each be a single heavy syllable; the usual trochaic option of two light syllables is unavailable because of the Maximal Stem Constraint. For the same reason, two disyllabic iambic feet cannot be concatenated together.[10] The licit basic stem patterns of Arabic nouns are exactly those that can be generated from a vocabulary of iamb and quantitative trochee, the option of a sequence of two of the same constituent, and the bounds set by the minimal word constraint and the Maximal Stem Constraint. The basic templates that the language actually employs are far more narrowly restricted by the Prosodic Morphology Hypothesis than was first suspected.

This analysis fits well with the facts, but leaves a major question unanswered: what about CvCCvC nouns from true quadriliteral roots, like xanjar? Although true quadriliterals CvCCvvC like rasmaal are somewhat more common than true quadriliterals CvCCvC, there is no radical skew between the two types.

The explanation for this is that templates of true quadriliteral nouns are lexically underspecified compared to the templates of triliteral nouns like jabbaar or jaamuus. The templates of the triliterals jabbaar or jaamuus must specify the weight of both syllables -- to ensure gemination or vowel length in the first syllable and vowel length in the second. With a quadriliteral root, though, the template only needs to specify the weight of the second syllable. The first syllable is necessarily heavy or bimoraic, since four consonants must be linked. There is no option in the language for linking a quadriliteral root onto a skeleton with just three available positions for consonants; for that reason, in McCarthy and Prince (1986: 66, 105) we proposed a general principle of melodic conservation, requiring that all root segments be linked to the skeleton. Moreover, even the fact that the skeleton is disyllabic is predictable for quadriliteral nouns; no other configuration is possible that conserves the melody. In other words, given a quadriliteral root, it is sufficient to know only whether or not it contains a light syllable. The base of the quadriliteral nouns as simply light syllable vs. heavy syllable; the rest of the template can be supplied by rule.[11]

---

[10] The minimal expansion of an iambic foot is a single heavy syllable. This is then identical to a quantitative trochee.

[11] See Archangeli 1988 for a very interesting approach to a similar problem of skeleton generation.

The complete prosodic analysis of the basic noun stems is
quite different from the first attempt or, indeed, from the
inventory required in CV skeletal theories. It is much more
restricted, focusing as it does on the overall prosodic well-
formedness of the different templates. In the following table, we
use the notation F_{QT} to refer to the quantitative trochaic foot
and F_I to refer to the iambic foot:

(35)

a. CvCC (baHr)

$$F_{QT}$$
$$|$$
$$\sigma$$

b. CvCvC (badal)

$$F_{QT}$$
$$\bigwedge$$
$$\sigma \ \ \sigma$$

c. CvCvvC (ǰadiid)

$$F_I$$

d. CvvCvvC (ǰaamuus)
   CvC_iC_ivvC (ǰabbaar)

$$F_{QT}F_{QT}$$

e. CvC_iC_jvC (xanǰar)

$$\sigma$$
$$|$$
$$\mu$$

f. CvC_iC_jvvC (rasmaal)

$$\sigma$$
$$\bigwedge$$
$$\mu \ \ \mu$$

The two types of bimoraic nouns are distinguished as monosyllabic
versus disyllabic quantitative trochees. Foot theory permits only
one type of iamb, so it is unnecessary to indicate that the iambic
type is disyllabic. The Maximal Stem Constraint limits stems to
two syllables, so it is unnecessary to say that the two
quantitative trochees concatenated together in a single stem are
each monosyllabic. The quadriliteral nouns, constrained by
melodic conservation, specify only the weight of a single
syllable, from which the full skeleton can be unambiguously
determined with general conditions of prosodic well-formedness.

Let us now sum up the results to this point. The same
notions of mora, syllable, and extrasyllabicity that function in
Arabic phonology also characterize the basic noun templates. In
the course of demonstrating this, two constraints have been
presented -- the bimoraic minimal word and the disyllabic maximal
stem. Together with the Prosodic Morphology Hypothesis, these
constraints have explained the central regularities in the
formation of the basic stem types.

7. Templatic Morphology in the Arabic Verb

The classic example of templatic morphology is presented by
the derivational system -- the conjugations -- of the Arabic verb.
The abundant evidence for the templatic character of the Arabic
conjugations essentially reduces to the observation that the shape

or canonical pattern of an Arabic verb stem is severely restricted
by both its morphological class membership and overall patterns of
the language. In comparison with a language like English, for
example, this is a truly remarkable situation.

(36) presents the stem patterns of the Arabic verb along the
same lines adopted earlier in (27), using the model roots /sm/
'poison', /fʕl/ 'do', and /dHrj/ 'roll' to represent all
biliterals, triliterals, and quadriliterals respectively. The
numerals preceding the forms are the designations of the Western
system of classification; the parenthesized numbers following the
forms give an exact indication of the frequency of these types in
Wehr's (1971) dictionary; those stem patterns with a count of zero
are known only from the Classical literature.

(36) Basic Verb Stem Patterns

Biliteral /sm/

1 samam (270)	2 sammam (127)	7 nsamam (31)	10 stasmam (34)
	3 saamam (18)	8 stamam (68)	11 smaamam (0)
	4 ʔasmam (78)	9 smamam (0)	12-15 ?

Triliteral /fʕl/

1 faʕal (2299)	2 faʕʕal (1271)	7 nfaʕal (229)	10 stafʕal (355)
	3 faaʕal (445)	8 ftaʕal (553)	11 fʕaalal (2)
	4 ʔafʕal (873)	9 fʕalal (18)	12 fʕawʕal (7)
			13 fʕawwal (0)
			14 fʕanlal (2)
			15 fʕanlay (0)

Quadriliteral /dHrǰ/

| | Q1 daHraǰ (296) | | Q3 dHanraǰ (1) |
| | | | Q4 dHarjaǰ (8) |

Excluded from analysis:
5 tasammam (85)
6 tasaamam (17)
5 tafaʕʕal (940)
6 tafaaʕal (377)
Q2 tadaHraǰ (111)

One remark is in order before we continue. The stems with prefixed
ta are not regarded as basic patterns to be treated in this
analysis. There are several reasons for this difference from
earlier analyses (McCarthy 1981). First, they obviously have a
prefix tV, whereas the other stem patterns cannot be
straightforwardly decomposed into concatenations of a prefix plus
independently occurring base morpheme. Second, statistical
examination of the pairings of different conjugations for a given
root shows, not surprisingly, that Form 5 tafaʕʕal tends to occur
only together with roots having Form 2 faʕʕal, Form 6 with Form 3,

and Form Q2 with Q1. This dependency between different conjugations, which an analysis with a <u>t</u>V prefix predicts, is otherwise unknown in the Arabic verb system. Third, the conjugations with prefixed <u>t</u>V are special in the vocalism that they present in the imperfective active; it is <u>yatafaaʕalu</u> when <u>yatafaaʕilu</u> is expected, based on what happens in the other conjugations. Fourth, just these conjugations take a peculiar form of the masdar that involves no alteration in the canonical shape of the verb: 5 <u>tafaʕʕul</u>, 6 <u>tafaaʕul</u>, Q2 <u>tadaħruj</u>. They thus present other evidence for special treatment.

Unlike the nouns, the verbs in the third and fourth columns of (36) have an initial consonant linked to an incomplete syllable. As we earlier showed, this consonant must be moraic by the Contiguity Constraint. A first pass at the analysis appears in (37), with the columns of (37) corresponding to those of (36):

(37) Moraic Skeleta for Verb Templates
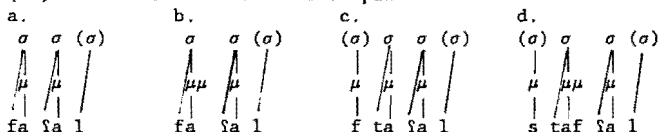
One issue which we must address is the conformity of these templates to the Prosodic Morphology Hypothesis in its strongest sense. Only the template in (37a) corresponds to a prosodic constituent (the quantitative or moraic trochee); the others do not. We will show that none of the verb templates is basic. Rather, all are derived by concatenating templatic morphemes, each of which is itself a prosodic constituent.

Some significant differences between the verb and the noun immediately emerge. First, the verb stem is always disyllabic, whereas nouns come in both monosyllabic (CvCC) and disyllabic (CvCvC, CvCvvC, etc.) flavors. Second, the finite verb stem permits no contrast in the weight of the second syllable -- it is always light, containing just one mora. But there are nouns with one (CvCvC) and nouns with two (CvCvvC) moras in the final syllable.

These observations obviously require some sort of explanation. In fact, a single explanation is possible for both: all finite verb stems have a light syllable templatic suffix. 38 represents this:

(38) Finite Verb Suffix

$$\sigma$$
$$|$$
$$\mu$$

In actual stems, this suffix is followed by the obligatory final (). The finite verb template suffix explains why all verbs have the same final syllable weight: the final syllable of the verb stem is actually a suffix which is constant across all conjugations of the verb. The differences among verb templates of different conjugations are therefore limited to the weight of the first (only) syllable and the presence or absence of an initial (σ). Moreover, this suffix also explains why finite verb stems are necessarily disyllabic. The finite verb suffix is attached to a monosyllabic base that is specified for each conjugation (the character of this base is investigated further below). If the base to which this suffix is attached were disyllabic, then with this suffix the result would exceed the Maximal Stem Constraint. And there must be a base to attach the suffix to, so no verb stem can be monosyllabic either.

There is solid independent motivation for this somewhat surprising result. There is only one pattern of masdar (nominalization) formation in the language that applies in a general way across the different conjugations of the verb. This masdar pattern is exemplified by the following forms. When unbracketed, they are the primary or only means of masdar formation for a particular conjugation (except in Form 1, where fiʕaal is one of about six common options); when bracketed, they are attested but described by Wright (1971) as rare:

(39)

Conjugation	Finite Verb	Masdar
1	faʕal	fiʕaal
2	faʕʕal	[fiʕʕaal]
3	faaʕal	[fiiʕaal]
4	ʔafʕal	ʔafʕaal
7	nfaʕal	nfiʕaal
8	ftaʕal	ftiʕaal
9	fʕalal	fʕilaal
10	stafʕal	stifʕaal
11	fʕaalal	fʕiilaal
12	fʕawʕal	fʕiwʕaal
13	fʕawwal	fʕiwwaal
14	fʕanlal	fʕinlaal
15	fʕanlay	fʕinlaay
Q1	daHraǰ	diHraaǰ
Q3	dHanraǰ	dHinraaǰ
Q4	dHarǰaǰ	dHirǰaaǰ

The cross-categorial generalization is that the masdar is identical to the finite verb except that the vowel melody is [i_a] and the final syllable contains a long vowel. This difference in final vowel length between the masdar and the finite verb is straightforwardly accounted for by setting up a different suffix for non-finite verbs:

McCarthy and Prince: Prosodic Morphology and Templatic Morphology

(40) Non-finite Verb Suffix

σ
/\
μ μ

Indeed, other processes for forming non-finite verbs or nouns from
finite verbs very often seem to involve this suffix: passive
participle 1 _mafʕuul_; masdar 2 _tafʕiil_; noun of instrument 1
mifʕaal.

The canonical form of the stem of every conjugation includes
the finite verb suffix. Therefore all differences in the
canonical form of different conjugations reside in the
monosyllabic base obtained by stripping off this suffix. The
bases, corresponding to the columns of (36), are:

(41) Base Templates
a.   b.   c. (σ) σ   d. (σ) σ
σ   σ | |
| /\   μ   /\
μ   μμ   μμ
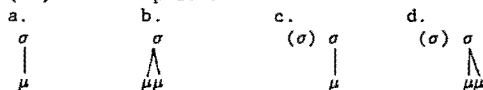
The bases are necessarily monosyllabic, for reasons already
outlined. All possibilities occur within the range delimited by
two options: (a) whether the syllable is heavy or light; and (b)
whether or not there is an initial (σ), the extrasyllabic mora.

The option for the initial syllable to be heavy or light is,
of course, expected under prosodic morphological theory. The
initial (σ), then, is clearly what requires our attention now.

Moore (to appear) has gone on to argue on phonological
grounds that an initial extrasyllabic mora (what we analyze as
(σ)) is a separate morpheme. Morphological considerations lead to
the same conclusion. The role of the initial incomplete syllable
(σ) within the morphological system as a whole is rather closely
circumscribed. Most conspicuously, it is impossible in nouns
(except for obviously deverbal ones like the masdars). This
observation follows straightforwardly if (σ) is a prefix of the
verbal system only, therefore unavailable in nouns. In that case,
the base templates of the various verbal conjugations would reduce
to just two possibilities, a light monosyllable and a heavy
monosyllable. We can even make a stab at the meaning of this
morpheme. Consider the set of conjugations that have this
putative prefix: 7-15, Q3, Q4. What these all have in common is
intransitivity or, more correctly, a reduction or minimization of
the valence of the underlying verb. Form 7 is usually described
as a passive or middle: _kasar_ 'break (tr.)', _nkasar_ 'break
(intr.)'. Form 8 is also a kind of reflexive or middle: _faraq_
'divide (tr.)', _ftaraq_ 'divide (intr.)'. Form 10 is yet another
reflexive or middle, related in meaning to Form 4: _ʔaslam_ 'give

up', staslam 'give oneself up'. Forms 9 and 11 describe the state

of being of color or hearing a bodily defect: Sfarar 'be yellow',
Hwalal 'have a squint'. The rare forms 12-15 and Q3-Q4 all
describe states as well.

The statistics of the distribution of conjugations among
different root types also support the analysis of (σ) as a
separate morpheme. Roots tend not to occur in both Form 7 and
Forms 8 or 10: out of 3062 biliteral and triliteral roots, only 69
occur in both 7 and 8 and only 29 occur in both 7 and 10. (The
scarcity of roots that take both 7 and 10 is significant at the
.05 level.) This is plausibly analyzed as a blocking effect
(Aronoff 1976): forms 7, 8, and 10 are functionally similar and
share the formal property of the prefixed (σ).

No doubt a more precise characterization of the semantics of
the different conjugations could better pin-down the meaning of
(σ), but it is sufficient for our purposes to recognize that the
different conjugations with initial (σ) have enough in common to
warrant setting it up as a prefix:

(42) Detransitivizing Verbal Prefix

(σ)

This leaves only one unexplained source of differentiation
in canonical pattern among the conjugations, whether the single
syllable of the base is monomoraic (light) or bimoraic (heavy).
The conjugations with a light-syllable base are 1, 7, 8, and 9.
What these have in common is that all are plausibly related to
Form 1. Form 7 is a kind of passive of 1, Form 8 is the reflexive
of 1, and form 9 could be regarded as the stative of 1. At this
point the analysis is somewhat subtle and conjectural, but
nevertheless it is worth pursuing this point to its logical
conclusion. The monomoraic syllable is the base of Form 1 and
closely related conjugations, while the bimoraic syllable is a
kind of default base, appearing with all other conjugations, a set
of derivational patterns that appear to have nothing in common:

(43) Bases of Verbal Derivational System

a. Form 1 Base   b. Default Base

   σ σ
   | /\
   μ μ μ

Further evidence for the default status of the bimoraic base comes
from the treatment of recently borrowed verbs -- they are always
in Form 2, with the bimoraic base: barrak 'park a car', dayyaT 'go
on a date' (cf. Smeaton 1973).

support for this idea. The posited relation between Forms 1, 7,
and 8 (9 is too uncommon for meaningful statistical analysis) is
confirmed by an authentic tendency for roots to take Form 7 or
Form 8 only if they also take Form 1. Of 260 roots in Form 7, 249
take Form 1 as well. Of the remaining 11 roots, 7 occur only in
Form 7. (These could plausibly be analyzed as instances of the
"missing base" phenomenon, like English uncanny/*canny.) The
connection is less striking between Forms 1 and 8. Of 621 roots
occurring in Form 8, 581 occur in Form 1 as well. Of the
remaining 40, 12 appear only in Form 8.

        Let us now sum up. The canonical patterns of the various
conjugations of the Arabic verb can be analyzed into a set of
morphological constituents. Conjugations marked by intransitivity
have a prefix (σ); those not so marked lack this prefix. Form 1
and its close relatives 7, 8, and 9 have a monomoraic monosyllabic
base; other conjugations have a default, bimoraic monosyllabic
base. All finite verb stems have a light-syllable suffix $[\mu]_\sigma$;
the most general pattern of nonfinite verb (masdar) formation has
a heavy-syllable suffix $[\mu\mu]_\sigma$. Thus, the moraic skeleta in 37 are
decomposed as follows (to which final (σ) is obligatorily added):

(44)

a.	b.	c.	d.
$\sigma + \sigma$	$\sigma + \sigma$	$(\sigma) + \sigma + \sigma$	$(\sigma) + \sigma + \sigma$
$\mu \quad \mu$	$\mu\mu \quad \mu$	$\mu \quad \mu$	$\mu\mu \quad \mu$
Base+Sfx	Base+Sfx	Pfx+Base+Sfx	Pfx+Base+Sfx

The units which make up the verb stem templates -- prefix, base,
and suffix -- each individually conforms to the Prosodic
Morphology Hypothesis in its strongest sense. The concatenations
of these morphemes do not, but this is what we expect; the
Hypothesis governs only the shape of basic templates, not what the
syntax or morphology do with them.

        The Prosodic Morphology Hypothesis has led to a much deeper
understanding of the internal structure of Arabic verb templates
than previously. Looking at the templates as a sequence of
prosodic units rather than a concatenation of CV segments reveals
significant internal regularities: all verb templates are
decomposable into a sequence of prosodic units with distinct
morphological functions.

<u>8. Template Satisfaction</u>

        It is now appropriate to turn to the question of template
satisfaction: how are the root and skeleton associated with one
another? This is of particular importance since moraic theory,

unlike its CV theory predecessor, is unable to distinguish between the two types of heavy syllables Cvv and CvC.[12] It will emerge that the ability to make this distinction is a liability rather than an advantage of the CV theory, since Arabic grammar does not actively exploit this putative skeletal distinction.

During the following discussion, it is important to keep in mind that terms like "CV theory" or "CV skeleton" are being used loosely, to refer to a whole family of phonology theories with the following properties. First, they must have segment-sized skeletal elements, unlike the moraic elements of prosodic theory. Second, they must be capable in principle of distinguishing Cvv from CvC syllables. This is obviously true of CV theory proper, but it is also true of those theories that can distinguish them by differences in syllabic structure. In particular, those theories with undifferentiated skeletal elements that nevertheless posit a branching syllabic nucleus for Cvv syllables but a branching rhyme (and a non-branching nucleus) for CvC syllables will meet this criterion (Levin 1983, 1985; Lowenstamm and Kaye 1986).

There are several basic observations about root/skeleton association in Arabic. First, syllable onsets are obligatorily filled. Second, all stems must be consonant-final. In our terms, all stems must end in an incomplete syllable (σ) by rule 19. Third, association of root with skeleton has a left-right asymmetry: there are biliteral verbs samam or nouns samm, but no biliterals *sasam or *sasm. Fourth, as is universally the case in prosodic morphology, maximization of melodic association takes absolute precedence over other considerations (McCarthy and Prince 1986: 66, 105). In particular, root consonants must be conserved; there are no cases (except for sporadic lexical exceptions) where a root consonant is lost by Stray Erasure (McCarthy 1979) because there is no templatic position available for it to occupy.

The basis has now been established for addressing the problem of the moraic nondistinctness of Cvv and CvC heavy syllables. Consider the monosyllabic stems like baĦr or barr vs. baab. From (45), it looks as if the CV skeleton can distinguish these two types but the moraic skeleton cannot:
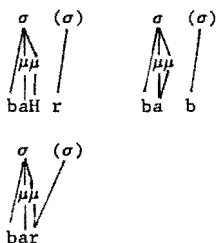
---

[12] See also Levin 1983, 1985 and Lowenstamm and Kaye 1986 for other approaches to eliminating the Cvv/CvC distinction.
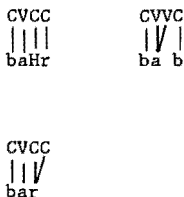
(45) Monosyllabic Stems: Moraic and CV Skeleton Comparison

Moraic Skeleton   CV Skeleton

But the CV skeleton is making a distinction that the Arabic
language really doesn't make. While words like triliteral baHr or
biliteral barr are extremely common and fully integrated into the
morphological system, with over one thousand examples in the
lexical material we have examined, words like baab are quite
special. It is arguably the case that all such words are derived
by regular phonological rules from underlying disyllables /CawaC/
or /CayaC/, as evidenced by singular/plural alternations like
baab/?abwaab 'door'. There may be a few nouns which present no
independent evidence for the underlying disyllable (like baaz,
plural biizaan 'bustard'), but they constitute a tiny minority of
irregular lexical expections (less than 1% of our data).

       So CvCC is the obligatory treatment of the monosyllabic noun
stem. This is one part of a more pervasive regularity: observe
that no disyllabic noun stem ever ends in CvCC, but many end in
CvvC. (Again, there are a few lexical exceptions, like dimašq
'Damascus'.) There is, then, no lexical distinction between the
two types of heavy final syllables; stem-finally, a bimoraic
syllable is necessarily CvC in monosyllables and Cvv in
disyllables. (Likewise, the masdars in (39), necessarily
disyllabic, also have Cvv final syllables.) The CvC/Cvv split can
be intepreted along minimal/super-minimal lines. Minimal words
with a final heavy syllable are necessarily monosyllabic; super-
minimal words with a final heavy syllable are necessarily
disyllabic. Many cases with this sort of segregation of the
lexicon are discussed in McCarthy and Prince (1990), where a
theory of the phenomenon is elaborated. The general idea is that
the behavior of minimal words constitutes a special case to which
super-minimal behavior is the default. The following rule
characterizes the generalization in these terms:

(46) Final Mora Association
μ] if minimal, otherwise μ]
⋮   ⋮
C   v

This rule does not need to be explicitly limited to the second
mora of a heavy syllable; its inapplicabilty to light syllables is
guaranteed by the requirement (in the Contiguity Condition) that
all nonperipheral syllables have vowels.

    Within the noun, there is only one other locus where an
apparent CvC/Cvv distinction is made: CvCCvvC medial geminates
like jabbaar versus CvvCvvC nouns like jaamuus. The number of
root consonants and the prosodic skeleta are identical in both
cases; how then to account for the apparent contrast between a
closed and open heavy initial syllable?

    As was already observed, medial gemination in the noun is
not limited to underived nouns like jabbaar, but also applies
productively in the noun of profession or habitual action and
semi-productively in one type of broken plural. Nouns like
jaamuus, on the other hand, are rather rare and this pattern is
not used in any systematic way by the morphology. In fact, Levy
1971 refers to it as only "semi-canonical". It is therefore
tempting to suggest that cases like jabbaar are the norm and that
jaamuus is lexically marked. (A few roots occur in both forms:
ballaaʕ+at 'sink, drain', baaluuʕ+at 'sewer, sink, drain'.)
Unfortunately, this explanation, whatever its merits for the noun,
is clearly not generalizeable to exactly the same problem in the
verb. In particular, Forms 2 and 3 (faʕʕal and faaʕal) are both
built on the heavy syllable templatic base, one with medial
gemination and one with vowel length.

    Medial gemination presents a problem for the otherwise
fairly straightforward extension of autosegmental phonology as a
theory of tone to prosodic morphology like that of Arabic. In
autosegmental tonal phonology (see especially Clements and Ford
1979 for the most striking evidence of this regularity), the
normal mode of association is one-to-one and left-to-right or
right-to-left. When the root /fʕl/ is associated with the CV
skeleton CVCCVC in this way, the result is incorrect, as (47)
shows:

(47)
CVCCVC ─ *faʕlal
 | //
 fʕl

Associating in the other direction also fails, producing the
impossible form *fafʕal. In the original treatment (McCarthy
1979, 1981), cases like this were dealt with by adding an
additional rule that produces the desired medial geminate by

36

reassociation. Alternative approaches to this problem have
subsequently been proposed: Levin (1983), Broselow (1984),
Angoujard (1984), Farley (1987), Farwaneh (to appear), Yip (1988),
Hoberman (1988), and no doubt others as well.

It is obviously impossible to review all of these proposals
here. Rather, we will focus on the most important observation
about medial gemination: it is always grammatically controlled.
That is, medial gemination appears under several different
conditions in the verb and noun which all must be specified
grammatically: nouns of occupation, plurals of lexicalized active
participles (see (48) below), Form 2 of the verb. There is just
one exception to this: the relatively unusual (about 2% of all
nouns) underived nouns like ǰabbaar. In contrast, final doubling
of a consonant is, in most cases, phonologically controlled,
appearing systematically with biliteral roots: nouns barr, sabab,
ǰadiid, baarir, ǰaaruur; verbs Halal, Haaǰaǰ, Hallal, ʔaHlal,
nHalal, Htalal, staHlal.

The broken plurals with medial gemination provide a further
clue about how this grammatical control is exercised. Arabic
active participles, like other productively derived nouns, do not
normally form broken plurals. Nevertheless, when they become
lexicalized (Levy 1971), as evidenced by some degree of
specialization of meaning, they may form broken plurals according
to one of the following patterns (the numbers represent the
frequency of each type in our sample):

(48)
a. fuʕʕal type
baahil   buhhal   21   'free'

b. fuʕʕaal type
ʔaabiq   ʔubbaaq   50   'fugitive'

The pattern of vocalization, [u_a], appears elsewhere in the
language with the same function, marking the broken plurals of
human nouns: waziir 'vizier', pl. wuzar-aaʔ.

Although most Arabic broken plurals are templatic (McCarthy
1983, McCarthy and Prince 1990) -- that is, they are formed on
different skeleta from their corresponding singulars -- these
plurals are not. For one thing, they obviously bear a close
resemblance in canonical form to their singulars. For another,
unlike the templatic broken plurals, these are formed from
singulars of invariant shape. The whole point of templatic
morphology is to satisfy a criterion not met here -- the
independence of the canonical form of input and output.

Assembling these observations into an analysis leads to the
following (cf. McCarthy 1983: 312-313). Substitution of medial
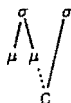gemination for vowel length is the primary mechanism relating the

plural to the singular. In addition, lexically specified words of
this type undergo an additional rule lengthening the final vowel.
The two rules are formalized in (49a); sample derivations appear
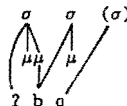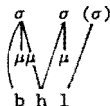in (49b):

(49)
a.

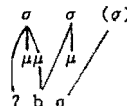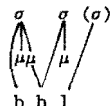Medial Gemination   Vowel Lengthening (in some words)

        σ   σ   σ]
        /\   /   ⋮
       μ μ /   μ
           ⋮/
           C

b.

Singular

              σ   σ (σ)   σ   σ (σ)
             /\ /\ /   /\ /\ /
            (μμ \/ μ /   (μμ \/ μ /
              \/   \/
              b h l   ʔ b q

Medial Gemination

              σ   σ (σ)   σ   σ (σ)
             /\ /\ /   /\ /\ /
            (μμ \/ μ /   (μμ \/ μ /
              \/   \/
              b h l   ʔ b q

Vowel Lengthening   DNA

                                            σ   σ (σ)
                                           /\ /\ /
                                          (μμ \/ μμ /
                                            \/
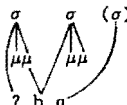   ʔ b q

[u_a] Vocalism   buhhal   ʔubbaaq

     The discussion of this type of broken plural formation
yields two results that are of great importance to the treatment
of medial gemination. First, it shows that the distinction
between the two types of heavy syllables that a CV skeleton can
make is, if anything, an impediment to the analysis of medial
gemination. The rule of Medial Gemination in (49a) places crucial
reliance on the moraic equivalence of Cvv and CvC heavy syllables.
Second, at least in this case, medial geminates are derived by
rule from representations without geminates. Arabic must contain a
rule creating geminates by adding an association line from an
onset consonant onto a preceding mora.

     Generalizing from this one case where the source of medial
gemination is demonstrably an association rule, it is plausible
that all instances of medial gemination are derived by applying
the association rule 49a. This essentially moves the problem of
gemination outside the scope of skeletal theory; it is enough if

McCarthy and Prince: Prosodic Morphology and Templatic Morphology

the skeleton provides the mora to which the geminated consonant
will be associated by rule. There is, then, no need for the
greater expressive power of the CV skeleton with its distinction
between the two types of heavy syllables. This also explains the
original observation that medial gemination is always
grammatically controlled: it must be grammatically controlled
because it is derived by a grammatically conditioned rule of
association. The relatively uncommon words like <u>jabbaar</u> are
derived by lexically governed applications of this rule.

    With medial gemination done by rule, the association of
consonantal root to template in the noun system can be entirely
determined from the following constraints:

(i)   Final Incompleteness (19), the requirement that all stems
end in an incomplete syllable (that is, a consonant, by the Onset
Rule).

(ii)   Final Mora Association (46), the requirement that final
consonant clusters appear in all and only monosyllables.

(iii) The Onset Rule, the requirement that all syllables begin
with a consonant.

(iv)   Melodic conservation, the requirement that all root
consonants be linked.

(v)   Left-to-right association.

These principles correctly generate all patterns of association
observed in noun stems. Some of them are independently motivated,
and may in fact be universal; (i) and (ii) are presumably
language-particular, but they capture significant generalizations
that have been mostly overlooked in previous accounts.

    We will take only a cursory look at consonant association in
the verb system, which still presents certain problems. In most
cases, the properties of the verb system can be accounted for in
exactly the same way as the noun:

1. Form 2 (<u>sammam</u>, <u>faʕʕal</u>) vs. Form 3 (<u>saamam</u>, <u>faaʕal</u>). Medial
gemination in Form 2 is the result of a grammatically-conditioned
rule of association. Form 3, then, is derived by simple filling
of the obligatory consonantal positions (onsets) in (37b), to
which Form 2 adds the application of the Medial Gemination Rule.

2. All cases like Form 1 (<u>samam</u>, <u>faʕal</u>), Form 7 (<u>nsamam</u>, <u>nfaʕal</u>),
Form 8 (<u>stamam</u>, <u>ftaʕal</u>), Form 9 (<u>smamam</u>, <u>fʕalal</u>), Form 11
(<u>smaamam</u>, <u>fʕaalal</u>), and the rare Forms 12 (<u>fʕawʕal</u>, 13 (<u>fʕawwal</u>),
14 (<u>fʕanlal</u>), 15 (<u>fʕanlay</u>), and Q3 (<u>dHanraj</u>) involve root
associations only to obligatorily consonantal positions (onsets
and an initial incomplete syllable), sometimes in competition with

a consonantal affix whose position is specified by the morphology.

(The affixes are presorted to the template along the lines of McCarthy 1979, 1981.)

3. Association in Form Q1 daHraj follows from melodic conservation.

What remains after these cases have been put aside are Forms 4 and 10 with biliteral roots, ʔasmam and stasmam. Conservation of melodic elements cannot in general account for the fact that the first syllable in these stems is closed, because this condition is satisfied in the biliterals by *ʔaasam and *staasam. Perhaps these are ill-formed because no root consonants are linked with the heavy-syllable "base" of the template (see (43)).

In general, association of root to skeleton in the verb reduces to filling of obligatorily consonantal positions exclusively, subject only to melodic conservation. Positioning of non-root templatic consonants like the t infix of Form 8 must be stipulated in Prosodic theory, as in any other. Medial gemination is the result of a grammatically conditioned rule that also applies in other morphological constructions.

9. Conclusion

We have argued that templatic constraints on word structure should be characterized in prosodic terms -- that is, in terms of notions like minimal word, foot, syllable and mora. In particular, we have seen that basic, underived templates of Arabic must be analyzeable in prosodic terms, as required by the Prosodic Morphology Hypothesis. Taken together with the treatment of the broken plural in McCarthy and Prince (1990), this material provides a comprehensive analysis of Arabic templatic morphology within prosodic theory.

Bibliography

Abu-Mansour, Mahasen. 1987. A nonlinear analysis of Arabic syllabic phonology, with special reference to Makkan. University of Florida dissertation.

Angoujard, J.-P. 1984. Aspects d'une micro-prosodie (le modele arabe). University of Paris VIII dissertation. [Not seen.]

Archangeli, Diana. 1986. Extrametricality and the percolation convention. University of Arizona, Tucson, manuscript.
_____. 1988. Prosodic theory and Yawelmani templates. University of Arizona, Tucson, manuscript.

Aronoff, Mark. 1976. Word formation in generative grammar. Cambridge, Mass.: MIT Press.

Brame, Michael. 1970. Arabic phonology: implications for phonological theory and historical Semitic. MIT Doctoral dissertation.

Broselow, Ellen. 1976. The phonology of Egyptian Arabic.
     University of Massachusetts, Amherst dissertation.
_____. 1982. On predicting the interaction of stress and
     epenthesis. Glossa 16, 115-132.
_____. 1984. Amharic, automatic spreading, and the obligatory
     contour principle. SUNY Stony Brook MS.
Chomsky, Noam and Morris Halle. 1968. The sound pattern of
     English. New York: Harper and Row.
Clements, G. N. and Kevin Ford. 1979. Kikuyu tone shift and its
     synchronic consequences. Linguistic Inquiry 10, 179-210.
Davis, Stuart. 1988. Syllable onsets as a factor in stress rules.
     Phonology 5, 1-19.
Farley, Anne. 1987. The 'doubled' verb of Standard Arabic. CLS 23
     Part Two: Parasession on Autosegmental and Metrical
     Phonology, ed. by Anna Bosch, Barabara Need, and Eric
     Schiller, 106-115. Chicago Linguistic Society, Chicago.
Farwaneh, Samira. To appear. Well-formed associations in Arabic:
     rule or condition. Perspectives on Arabic Linguistics II,
     ed. by Mushira Eid and John McCarthy. John Benjamins,
     Amsterdam.
Goldsmith, John. 1976. Autosegmental phonology. MIT Doctoral
     dissertation.
Halle, Morris and Jean-Roger Vergnaud. 1987. An essay on stress.
     Cambridge, Mass.: MIT Press.
Harris, James. 1983. Syllable structure and stress in Spanish.
     Cambridge, MA: MIT Press.
Hayes, Bruce. 1982. Extrametricality and English stress.
     Linguistic Inquiry 13, 227-276.
_____. 1985. Iambic and trochaic rhythm in stress rules.
     Proceedings of the Berkeley Linguistics Society 11, ed. by
     M. Niepokuj, M. VanClay, V. Nikiforidou, and D. Feder.
     Berkeley, CA: Berkeley Linguistics Society.
_____. 1987. A revised parametric metrical theory. NELS 17, ed. by
     Joyce McDonough and Bernadette Plunkett, 274-289. Amherst,
     MA: Department of Linguistics, University of Massachusetts.
_____. 1989. Compensatory lengthening in moraic phonology.
     Linguistic Inquiry 20, 253-306.
Hoberman, Robert. 1988. Local and long-distance spreading in
     Semitic morphology. Natural Language and Linguistic Theory
     6, 541-550.
Howell, Mortimer Sloper. 1986 repr. A grammar of the Classical
     Arabic language. Delhi: Gian Publishing House.
Hyman, Larry. 1984. A theory of phonological weight. Dordrecht:
     Foris.
Itô, Junko. 1982. The syllable structure of Russian. University of
     Massachusetts, Amherst MS.
_____. 1986. Syllable theory in prosodic phonology. University of
     Massachusetts, Amherst dissertation.
_____. 1989. A prosodic theory of epenthesis. Natural Language and
     Linguistic Theory 7, 217-260.

Kenstowicz, Michael. 1981. The metrical phonology of Arabic Conference on Phonological Theory.

Levin, Juliette. 1983. Reduplication and prosodic structure. MIT MS.

_____. 1985. A metrical theory of syllabicity. MIT dissertation.

Levy, Mary. 1971. The plural of the noun in Modern Standard Arabic. University of Michigan dissertation.

Lowenstamm, Jean and Jonathan Kaye. 1986. Compensatory lengthening in Tiberian Hebrew. Studies in Compensatory Lengthening, ed. by Leo Wetzels and Engin Sezer, 97-132. Dordrecht: Foris.

McCarthy, John. 1979. Formal problems in Semitic phonology and morphology. MIT dissertation.

_____. 1981. A prosodic theory of nonconcatenative morphology. Linguistic Inquiry 12, 373-418.

_____. 1983. A prosodic account of Arabic broken plurals. Current Trends in African Linguistics 1, ed. by I. Dihoff, 289-320. Dordrecht: Foris.

_____ and Alan Prince. 1986. Prosodic morphology. University of Massachusetts, Amherst, and Brandeis University MS.

_____ and _____. 1988. Quantitative transfer in reduplicative and templatic morphology. Linguistics in the Morning Calm 2, ed. by Linguistic Society of Korea, 3-35. Seoul: Hanshin Publishing Co.

_____ and _____. 1990. Foot and word in prosodic morphology: the Arabic broken plural. Natural Language and Linguistic Theory 8, 209-284.

_____ and _____. Forthcoming. Prosodic morphology. Cambridge, MA: MIT Press.

Moore, John. To appear. Doubled verbs and vowel association in Standard Arabic. Perspectives on Arabic Linguistics II, ed. by Mushira Eid and John McCarthy. John Benjamins, Amsterdam.

Prentice, D. J. 1971. The Murut languages of Sabah. Canberra: Australian National University.

Prince, Alan. 1983. Relating to the Grid. Linguistic Inquiry 14, 19-100.

Selkirk, Elisabeth. 1980. The role of prosodic categories in English word stress. Linguistic Inquiry 11, 563-605.

_____. 1981. Epenthesis and degenerate syllables in Cairene Arabic. Theoretical Issues in the Grammar of Semitic Languages (MIT Working Papers in Linguistics 3, ed. by Hagit Borer and Youssef Aoun, 209-232. Cambridge, MA: Department of Linguistics and Philosophy, MIT.

Smeaton, B. H.. 1973. Lexical expansion due to technical change. Bloomington, Indiana: Indiana University Press.

Strehlow, T. 1942. Aranda phonetics. Oceania 12, 255-302.

Wehr, Hans. 1971. A dictionary of Modern Written Arabic. Ed. J. M. Cowan, Ithaca, NY: Spoken Language Services.

Yip, Moira. 1988. Template morphology and the direction of association. Natural Language and Linguistic Theory 6, 551-578.