1990

Foot and word in prosodic morphology: The Arabic broken plural

John J McCarthy
Alan Prince

JOHN J. MCCARTHY AND ALAN S. PRINCE

FOOT AND WORD IN PROSODIC MORPHOLOGY: THE ARABIC BROKEN PLURAL*

This article proposes a theory of PROSODIC DOMAIN CIRCUMSCRIPTION, by means of which rules sensitive to morphological domain may be restricted to a prosodically characterized (sub-)domain in a word or stem. The theory is illustrated primarily by a comprehensive analysis of the Arabic broken plural: it is further supported by analysis of a number of processes from other languages, yielding a formal typology of domain-circumscription effects. The results obtained here depend on, and therefore confirm, two central principles of Prosodic Morphology: (1) the Prosodic Morphology Hypothesis, which requires that templates be expressed in prosodic, not segmental terms; and (2) the Template Satisfaction Condition, which requires that all elements in templates are satisfied obligatorily.

1. INTRODUCTION

The study of the relationship between morphology and phonology has played an important role in recent linguistic investigations. On the one hand, work in the theory of Lexical Phonology speaks to the problem of phonological rule application in the course of a morphological derivation. On the other, the body of research on templatic morphology shows the essential role played by phonological structure in capturing morphological regularities.

The theory of Prosodic Morphology developed in McCarthy and Prince (1986, 1988, forthcoming a, b) advances several proposals about the basic character of phonological structure and its consequences for morphology. Three fundamental theses are:

(i) **Prosodic Morphology Hypothesis.** Templates are defined in terms of the authentic units of prosody: mora (μ), syllable (σ), foot (F), prosodic word (W), and so on.

(ii) **Template Satisfaction Condition.** Satisfaction of templatic constraints is obligatory and is determined by the principles of prosody, both universal and language-specific.

(iii) **Prosodic Circumscription of Domains.** The domain to which

* We are indebted to Mark Aronoff, A. R. Ayoub, M. G. Carter, Morris Halle, Michael Kenstowicz, Armin Mester, and three anonymous reviewers for valuable comments on this article. Special thanks go to Linda Lombardi for a very close reading of the entire manuscript that contributed much to the content and exposition.

Natural Language and Linguistic Theory **8:** 209–283, 1990.

> morphological operations apply may be circumscribed by prosodic criteria as well as by the more familiar morphological ones. In particular, the **minimal word** within a domain may be selected as the locus of morphological transformation in lieu of the whole domain.

We will elaborate considerably on these principles below, focusing particularly on prosodic circumscription.

New developments in linguistic theory often bring illumination to longstanding descriptive problems, while at the same time exposing new difficulties at a more subtle and abstract level. So it is with the Arabic broken plural. In traditional accounts like that of Wright (1971: 191–233), plural formation has all the properties of a poorly understood or perhaps even chaotic process, requiring a dense taxonomy of 31 plural types, each corresponding to as many as 11 singular types. The theory of CV-based templatic morphology has made considerable inroads into this apparent complexity, isolating a small set of formational processes and unifying a fair number of traditionally distinct patterns under a single template (McCarthy 1979, 1981). But, as Hammond (1988) has observed, the standard conception of templatic morphology brings with it a major new liability, the problem of transferring various characteristics from the singular to the broken plural. This fundamental problem turns out to be intractable in CV-template theories, including (as we will show) the one proposed by Hammond.

Prosodic Morphology offers a new perspective on the problem, and it is a goal of this article to demonstrate how the principal features of the broken plural phenomenon follow directly from its characterization in prosodic terms. In particular, it will emerge that the correct analysis of the transfer problem goes hand in hand with a wide generalization over productive plural types. We propose that the central plural-forming strategy of the language parses out an initial **minimal word** from the base – a prosodically circumscribed domain – and maps the contents of that minimal word onto an iambic foot. The broken plural, then, makes a full, systematic use of the categories and operations provided by the theory of Prosodic Morphology, providing a particularly interesting test case and a robust new source of evidence for the theory.

To secure our empirical claims, we have collected all nouns forming broken plurals in the first half of Wehr (1971), the authoritative English-language dictionary of Modern Standard or Literary Arabic. The data base contains a total of about 3500 singular/plural pairs, when doublets are considered, and should be more than adequate for establishing the actual role and weight of the various patterns. Although most reference grammars, like Wright (1971), deal with Classical Arabic, while Wehr records the contemporary literary language, the differences between these two sources of evidence are negligible; our experience is that the correspondence is nearly exact except for certain very rare plural patterns that have fallen into disuse. Furthermore, our investigation has been much aided by the penetrating and exhaustive treatment of this problem by Levy (1971).

This article will touch on virtually all the broken plural phenomena in Arabic and give a full account of the dominant regularities of the system, aiming to achieve a match between theory and observation that improves significantly on previous work. The article is organized as follows. Section 2 lays out the basic facts of the broken plural and closely related diminutive systems and presents our analysis of them informally. Section 3 develops the formal theory of prosodic circumscription and applies it to the descriptive problems of the Arabic plural and diminutive in all their detail. Section 4 reviews the shortcomings of previous approaches, focusing on that of Hammond (1988). Section 5 treats issues that are ancillary to the main thrust of our analysis; the conclusion briefly summarizes the results.

2. The Broken Plural and Diminutive in Outline

2.1. *The Large-scale Structure of the Arabic Plural*

Traditional grammars of Arabic distinguish between two modes of plural formation, the broken plural and the sound plural. The broken plural primarily involves internal modification of the singular stem, as in *nafs/nufuus* 'soul/pl.' or *jundub/janaadib* 'locust/pl.'; the sound plural is formed by suffixation of masculine +*uun* or feminine +*aat* to a usually unchanged stem, as in (1)[1,2]

[1] We will use the following transcription for the Arabic consonants. *ħ* and ʕ are pharyngeals and *ṭ*, *ḍ*, *ṣ*, *ẓ* denote the emphatic (pharyngealized) consonants.

	t		k	q		ħ
b	d	j			ʔ	
f	θ,s	š				
				x	ħ	
	ð,z			g	ʕ	
	ṭ,ḍ					
	ṣ,ẓ					
m	l,r,n					
w		y				

[2] In feminines CVCC-*at* the sound plural /CVCC-aat/ typically shows an epenthetic vowel

(1)	Singular	Plural	
	ʕuθmaan	ʕuθmaan+uun	'Othman (man's name)'
	šuwayʕir	šuwayʕir+uun	'poet (dim.)'
	kaatib	kaatib+uun	'writing (participle)'
	hind	hind+aat	'Hind (woman's name)'
	ramaḍaan	ramaḍaan+aat	'Ramadan (a month)'
	kaatib+at	kaatib+aat	'writing (fem. participle)'
	taʕriif	taʕriif+aat	'definition (nominalization)'

Although the term "sound plural" suggests normality – and indeed its form is entirely predictable from gender and other grammatical information – the sound plural is in no way the regular or usual mode of pluralization. Essentially all canonically-shaped lexical nouns of Arabic take broken plurals, including many loans, even very recent ones: *film/ʔaflaam* 'film'; *bank/bunuuk* 'bank'; *balyuun/balaayiin* 'billion'. The sound plural is systematically found only with members of the following short list: proper names; transparently derived nouns or adjectives such as participles, deverbals, and diminutives (Levy 1971); noncanonical or unassimilated loans (*tilifuun/tilifuun + aat*); and the names of the letters of the alphabet, which are mostly noncanonical. Surprisingly, the regular plural and past tense suffixes /-z, -d/ of English fall under grossly similar restrictions, even though their range of applicability appears to be vastly wider than the sound plural's. English words transparently derived from other categories always take the regular suffixes, even if they qualify phonologically for subregular morphology. Thus, as Kiparsky (1973) has noted, we have (underived) *leaf/leaves*, with the fricative-voicing subregularity, but in names we find only the regular suffix: the *Toronto Maple Leafs*; *two letter f's*/*[evz]. In the verbal system, subregularities apply to simple verbal stems and their verbal derivatives but not to derivatives of adjectives or nouns. The *ing/ang-ung* pattern, for example, is quite productive among pure verbs, but does not extend beyond them: *he rang the bell/ringed the camp with artillery*, the latter with denominal $[[\text{ring}]_N]_V$. (For recent discussion, see Pinker and Prince (1988).) The main difference is that the subregularities of English do not span much of the input space (and they do it in a largely sporadic fashion), while broken plurals are formed on literally every canonical noun type in Arabic. Both languages organize

in the CC-cluster which is either *a* or a copy of the stem vowel. For example, *kisr-at* 'fragment' pluralizes as either *kisir-aat* or *kisar-aat*. Epenthesis is blocked when CC is a geminate or when the cluster begins with glides *w*, *y*. Only in stems C*a*CC is epenthesis obligatory.

their morphologies into a special case/general case structure, suffixing by default when the other competing inflectional modes are inapplicable, and both languages require that the input to the specialized system meet standards of canonicality, phonological (stem-shape patterns) and morphological (nonderived status). In Arabic the "special case" system is fully articulated and relatively few items escape it to end up with the default "sound" suffix. For the lexicon as a whole, then, broken plural formation is by far the norm rather than the exception.

Although broken plurals are non-exceptional, they present a formal diversity that, when taken head-on, is daunting. A forest, however, supervenes upon the trees. To see it, we need to keep one eye on the prosodic structure of the plural patterns and the other on their actual lexical distribution. Wright's 31 types can be divided into just 4 shape-defined categories:

(2) Wright's Broken Plural Patterns

a.	*Iambic*	b.	*Trochaic*	c.	*Monosyllabic*	d.	*Other*
5.	CiCaaC	1.	CuCaC	2.	CuCC	7.	CuC_iC_iaC
6.	CuCuuC	4.	CiCaC	@12.	CiCC + at	8.	CuC_iC_iaaC
23.	CaCaaC	#28.	CaCaC	18.	CiCC + aan		
*14.	/CaCaaC/	11.	CiCaC + at	19.	CuCC + aan		
+24.	CaCaaC + /ay/	$13.	/CaCuC/	+22.	CaCC + /ay/		
#25.	CaCiiC	3.	CuCuC	#29.	CaCC		
#26.	CuCuuC + at	9.	CaCaC + at				
#27.	CiCaaC + at	10.	CuCaC + at				
16.	CawaaCiC	20.	CuCaC + aaʔ				
17.	CaCaaʔiC	&15.	/CaCiC/ + at				
Q1.	CaCaaCiC	&21.	/CaCiC/ + aaʔ				
Q2.	CaCaaCiiC						

Sigla:

#	Rare according to Wright
*	Metathesizes to ʔaCCaaC
$	Metathesizes to ʔaCCuC
&	Metathesizes to ʔaCCiC
+	Underlying /ay/ to [aa] by regular glide phonology
@	Usually has CiCC + aan doublet, according to Wright

The one analytical decision deserving notice is the representation of plurals pronounced [ʔaCC . . .] as underlying /CaC . . ./. Following Levy (1971), we posit a rule of stem-initial *Ca* Metathesis, which clarifies the structural affinity of these forms. Though morphologically governed, *Ca* Metathesis has fair generality and is active in several aspects of nominal morphology; see 5.4 below for further discussion.

The patterns are named for their characteristic prosody; all the forms in (2a) begin with the iambic foot CvCvv+; the forms in (2b) are all

CvCvC, which is the typical quantitative trochee, a foot of two moras with final consonant extrametricality (justified below); CvCC is the only productive monosyllabic canon in the language (McCarthy and Prince forthcoming a). The four classes are unequal in importance: patterns (2c) and (2d) are of limited interest; the trochaic pattern (2b) has some generality; and the iambic pattern (2a) is truly productive. It is the iambic pattern, therefore, that deserves serious explication. To support this evaluation, we now offer the central findings of our lexical survey, working upward through the scale of productivity.

The most narrowly restricted plural canons are the C*u*C$_i$C$_i$*a*(*a*)C forms (2d), which arise only from lexicalized active participles C*aa*C*i*C (these also form plurals in the iambic and trochaic patterns). Two typical examples: *kaafil/kuffal* 'breadwinner'; *kaafir/kuffaar* 'infidel'. Forms C*u*C$_i$C$_i$ *a*(*a*)C account for about 30% of the masculine plurals (69/245) from C*aa*C*i*C and virtually none of the feminines. It is therefore appropriate to posit, with McCarthy (1983), a rule limited to masculine C*aa*C*i*C that spreads the medial consonant of the singular backwards to close the first syllable, usurping its second mora. Schematically, the spreading rule changes CvvC$_i$vC to CvC$_i$C$_i$vC. The vowel melody of the singular is replaced by /u_a/ which also appears in other plurals of similar semantic classes. Unpredictably, the vowel of the second syllable is often lengthened (70% of our sample, 49/69).

The monosyllabic plural pattern (2c), usually suffixed with +*at*, +*ay*, or +*aan*, is widespread, but at very low levels of frequency in all classes, indicating nonproductivity. Overall, it accounts for only 4% (95/2694) of the triliteral broken plurals in our sample.[3] It is tempting to treat this pattern as root-and-template morphology, but the evidence is not compelling. The predominance of suffixing in this class suggests a process of presuffixal ellipsis, which would then be exceptionless before the suffix +*aan* when it appears in broken plurals. Alternatively, one could regard these suffixes as choosing a monosyllabic stem canon, along the lines of Yokuts (Archangeli 1983) or Southern Sierra Miwok (Smith and Hermans 1982; Smith 1985).

The one populated nonsuffixing form, C*u*CC, is almost entirely limited to deverbal adjectives of color or bodily defect, a semantic class that also plays a role in the Arabic verbal system. Although in the masculine

---

[3] Quadriliteral nouns cannot form monosyllabic broken plurals. This follows from the principle of melody conservation (McCarthy and Prince 1986) and the canons of Arabic syllable structure – a single quadriconsonantal syllable (even with final consonant extrametricality) is prosodically impossible.

singular these adjectives normally have the pattern /CaCaC/ (surface ʔ*a*CC*a*C), the feminine singular has the same canonical monosyllable CvCC shape as the plural: m. sg. *ʔaħmar*, f. sg. *ħamr + aaʔ*, pl. *ħumr* 'red'. Therefore the plural can be formed by imposing /u/ vocalism on the independently required monosyllabic stem allomorph.[4]

The trochaic pattern (2b) is widespread and assumes an important if not exclusive role in 3 distinct lexical niches:

Feminines C*i*CC+*at* and C*u*CC+*at* form plurals C*i*C*a*C and C*u*C*a*C, respectively, about 75% of the time (138/184). Examples: *rukb+at/rukab* 'knee', *ħikm + at/ħikam* 'maxim'.

Unsuffixed forms CvCvvC, with five distinct vocalizations, take the trochaic pattern as the modal (commonest) plural, in the range of 50% in the large and various C*a*C*ii*C form-class (134/265) to 61% in C*a*C*uu*C (17/29). The choice among the various trochaic plural patterns of these nouns is partly determined on semantic grounds; we take this up in Section 5.3. Examples: *waziir/wuzar + aaʔ* 'vizier', *kitaab/kutub* 'book', *janaab/ʔajnib+at* (from /janib+at/) 'wing'.

The lexicalized participles C*aa*C*i*C take the trochaic plural (C*a*C*a*C+*at* or C*u*C*a*C+*at*, depending on the phonology of the root) at a rate of 22% (54/245). This puts the trochaic plural about on a par with the C*u*C_iC_i*a*(*a*)C form just discussed (28%, 69/245) and the iambic pattern C*awaa*C*i*C (26%, 65/245), with 30 nouns taking more than one of these plural patterns. Examples: *ṭaalib/ṭalab + at* 'seeker; student', *saaqiy/suqay+at* 'cupbearer'.

For these cases it appears that a root-and-template approach is appropriate. The template is the disyllabic quantitative trochee (McCarthy and Prince 1986, forthcoming b; Hayes 1987), comprising two moras and two syllables, with an extrametrical final consonant (that is, CvCv(C)). The choice of vocalism is predictable in the feminines: it is just that of the singular, with the vowel /a/ supplied to head the second syllable.

The iambic pattern is one that truly dominates the lexicon (examples below in (4)). Triliteral nouns are built on four major stem canons: CvCC, CvCvC, CvCvvC, and CvvCvC; gender suffixation splits each category in

---

4 Wright (1971: p. 200) reports that CvC*aa*C singulars from medial *w* roots also form monosyllabic plurals: *nawaar*, *nuur* (from /nuwr/) 'a middle-aged, married woman'. Our lexical material contains only four nouns behaving in this way; the remaining twenty or so Cv*waa*C singulars form expected iambic or trochaic plurals.

two, with $+\emptyset$ for formal masculines and *+at* for feminines. Of the eight major canonical classes thus distinguished, the iambic plural is overwhelmingly favored in four, which include all possible stem shapes:

83% of the populous unsuffixed monosyllabic noun class CvCC form iambic plurals (567/681). Of the remainder, 66% have an iambic plural doublet (75/114).

81% of simple unsuffixed disyllables CvCvC take iambic plurals (174/213). Of the remainder, 54% have the iambic plural as one of several options (21/39).

88% of feminines CvCvvC + *at* take iambic plurals (148/168) (including a predictable alternant with roots ending in a high glide). In all of the remainder the iambic plural is one option, usually in competition with the feminine sound plural. Therefore the iambic plural has complete coverage in this class.

97% of feminines CvvCvC + *at* take iambic plurals (87/89).

Furthermore, the iambic plural is important in three of the remaining four triliteral classes – CvCC+*at*, CvCvC+*at*, and CvvCvC – where it competes with the trochaic pattern. Only in the class of unsuffixed CvCvvC forms is it insignificant, at 8% (35/447).

Finally, the iambic pattern is exclusively used for pluralization in the very large "quadriliteral" CvXCv(v)C class (which includes true quadriliteral roots, bi- or triliterals with a derivational prefix, and CvvCvvC nouns). This class comprises between 1/4 and 1/3 of the nouns in our sample, with 817 members altogether.

The results of our investigations are summarized by the following table:

(3)

Stem	$+\emptyset$ (masculine)	+at (feminine)
CvCC	!	√
CvCvC	!	√
CvCvvC	X	!
CvvCvC	√	!
CvXCv(v)C	!!	!!

Sigla:

!! All have iambic form as plural
! Greater than 90% have iambic form as a plural
√ Iambic plural is significant competitor (20%–50% total)
X Iambic plural insignificant (less than 10%)

The facts are clear: the challenge is to substantiate the informal notion that a single pattern unites all the classes grouped under the iambic rubric.

2.2. *A Overview of the Iambic Plural System*

In (4) we display examples of broken plurals from all the classes where the iambic plural is the dominant or sole form:

(4) Iambic Broken Plurals

	Singular	*Plural*	
a.	CvCC		
	nafs	nufuus	'soul'
	qidħ	qidaaħ	'arrow'
	ħukm	/ħakaam/	'judgment' [ˀaħkaam]
b.	CvCvC		
	ˀasad	ˀusuud	'lion'
	rajul	rijaal	'man'
	ʕinab	/ʕanaab/	'grape' [ˀaʕnaab]
c.	CvCvvC + at		
	saħaab + at	saħaaˀib	'cloud'
	jaziir + at	jazaaˀir	'island'
	kariim + at	karaaˀim	'noble'
	ħaluub + at	ħalaaˀib	'milch-camel'
d.	CvvCvC + at		
	faakih + at	fawaakih	'fruit'
	ˀaanis + at	ˀawaanis	'cheerful'
e.	CvvCv(v)C		
	xaatam	xawaatim	'signet-ring'
	jaamuus	jawaamiis	'buffalo'
f.	CvCCv(v)C		
	jundub	janaadib	'locust'
	sulṭaan	salaaṭiin	'sultan'

The key invariant uniting these patterns is the initial iambic sequence CvCvv+. (As noted above, superficially exceptional forms like [ˀaħkaam] and [ˀaʕnaab] are metathetic.) Although the defining iambic sequence has a clearly templatic character, the familiar resources of root-and-template morphology are quite inadequate to the task of representing it. The fault lies not in the notion of template but in its presumed dependence on the consonantal *root*; for the iambic plural systematically reflects aspects of the singular that the consonantal root does not determine. (Although Arabic templatic morphology is chiefly root-based, stem-based templatic formations, other than the broken plural, are also known; see McCarthy (1979) and Bat-El (1989).)

Perhaps the most familiar of the non-root properties "transferred" from singular to plural (Hammond 1988) is the final-syllable vowel length in forms where the first syllable is heavy:

(5)

	Root	*Singular*	*Plural*	
a.	/jndb/	jundub	janaadib	'locust'
	/sl$\d{t}$n/	sul$\d{t}$aan	salaa$\d{t}$iin	'sultan'
b.	/xtm/	xaatam	xawaatim	'signet-ring'
	/jms/	jaamuus	jawaamiis	'buffalo'

The length of the vowel in the final syllable of the singular is carried over to the final syllable of the broken plural. This transfer of vowel length occurs only in singulars with an initial heavy syllable.

Beyond obvious transfer, there are two other notable ways in which the shape of the plural depends directly on the prosody of the input stem. First, the number of syllables in the plural depends on the number of moras in the base. Bimoraic stems form disyllabic plurals; longer stems form trisyllabic plurals:

(6)

Root	*Singular*	*Plural*	*Prosody of base*	
/nfs/	nafs	nufuus	bimoraic	'soul'
/ʕnb/	ʕinab	/ʕanaab/	bimoraic	'grape'
/jndb/	jundub	janaadib	trimoraic	'locust'
/sħb/	saħaab + at	saħaab	trimoraic	'cloud'

Second, triconsonantal singulars with a long vowel require a default consonant *w*, realized as ˀ under certain phonological conditions. The position of the long vowel in the base determines the position of the default consonant in the plural. A long vowel in the first syllable leads to second-syllable insertion; a second-syllable long vowel leads to third-syllable insertion. Thus, a singular CvvCv(v)C corresponds to a plural Cv*w*vvCv(v)C; a singular CvCvvC corresponds to a plural CvCvv*w*vC:

(7)

Root	*Singular*	*Plural*	
/xtm/	x**aa**tam	xa**w**aatim	'signet ring'
/jms/	j**aa**muus	ja**w**aamiis	'buffalo'
/sħb/	saħ**aa**b + at	/saħaa**w**ib/	'cloud' [saħaaˀib]

All of these properties of canonical form are carried over in one way or another from singular to plural, despite the fact that the root itself contains no information about canonical form. Equally absent from the root is information about idiosyncratic patterns of consonant spreading, which hold of both singular and plural:

(8)	*Root*	*Singular*	*Plural*	
	/nwr/	nuwwaar	nawaawiir	'white flowers'
	/jlb/	jilbaab	jalaabiib	'a type of garment'
	/tn/	tinniin	tanaaniin	'sea monster'

Similarly, the fact that a root is reduplicated rather than spread, though not encoded in the root itself, is rigorously carried over to the plural:

(9)	*Root*	*Singular*	*Plural*	
	/zl/	zalzal + at	zalaazil	'earthquake'
	/jd/	judjud	jadaajid	'cricket (zool.)'

This kind of information – vowel quantity, number of syllables, consonant spreading, or reduplicated status – is exactly what the root abstracts away from. In the true root-and-template derivational morphology of the noun and verb, only the root consonantism carries over from one form to another in a prosodically diverse set like {*kitaab* 'book', *kaatib* 'writer', *katab* 'wrote', *kattab* 'caused to write', *kaatab* 'corresponded', (*ʔi*)*staktab* 'dictated', ...}. The broken plural, then, cannot be obtained with the ordinary resources of root-and-template morphology.

The category root is also morphologically inappropriate as the basis of broken-plural formation, since some derivational affixes are transferred intact:

(10)	*Affix*	*Root*	*Singular*	*Plural*	
	/m+/	/rħl/	marħal + at	maraaħil	'stage'
	/m+/	/ftħ/	miftaaħ	mafaatiiħ	'key'
	/ʔ+/	/mθl/	ʔumθul + at	ʔamaaθil	'example'
	/t+/	/qdr/	taqdiir	taqaadiir	'calculation'
	/y+/	/nbʕ/	yanbuuʕ	yanaabiiʕ	'spring'
	/+aan/	/slṭ/	sulṭaan	salaaṭiin	'sultan'

Besides these affixes, the quasi-phonological ʔ obtained by *Ca* Metathesis (see section 5.4) can participate in broken plural formation under the right conditions. The elative adjective *ʔakbar* (from /kabar/ by Metathesis) 'greater, greatest' has two plurals: sound *ʔakbar* + *uun* in a purely adjectival sense and broken *ʔakaabir* for the lexicalized nominal 'grandees'.

A final argument for the impossibility of obtaining the broken plural from root-to-template mapping comes from the unusual phenomenon in the Classical language of the "plural-of-the-plural", in which a plural is formed from a stem that is itself a broken plural. (According to Wright (1971: 232), the plural-of-the-plural can be used when "the objects denoted are at least nine in number, or when their number is indefinite.") Consider these examples:

(11) Plural of the Plural

	Root	Sg.	Pl.	Pl./Pl.	Pl./Pl./Pl.	
a.	klb	kalb	ʔaklub /kalub/	ʔakaalib		'dog'
b.	frq	firq + at	firaq	ʔafraaq /faraaq/	ʔafaariiq	'sect'

The immediate plural of *kalb* is /kalub/ which metathesizes to *ʔaklub*. The consonant ʔ, inserted to fill the empty onset created by C*a* metathesis, is treated on a par with any other stem consonant when plural formation reapplies. Similarly with *firq*, where the doubly derived plural-of-the-plural-of-the-plural *ʔafaariiq* takes as input the metathesized plural-of-the-plural *ʔafraaq*. Final-syllable vowel length in the trisyllabic plurals is transferred, of course, from the (already plural) base – and this vowel length comes from the prior pluralization process, not from the singular, much less from the root.

These observations establish that the iambic plural is related directly to the actual stem from which it is formed, not to the root of that stem. But the relationship cannot be treated as simple holistic accommodation to a template. Under the Template Satisfaction Condition (TSC), formulated earlier in (ii), all templatic constraints are held to be obligatory; in Arabic root-and-template morphology this is demonstrably true, even in CV-based theories, where the TSC cannot be imposed as a universal.[5] Yet the iambic plurals include both two- and three-syllable forms with diverse patterns of vowel length; no single template can obligatorily include them both.

The problem is that there are two distinct systems of invariance: within the plural, the iambic invariant; and between the singular and plural, the various "transferred" structural properties. Only the within-plural

[5] For example, the realization of the Arabic CVCVC template with the biliteral root /sm/ by spreading the final consonant to yield *samam* shows that the TSC must be obligatory for templatic systems in CV-based theories. But the realization of the Tagalog CVCCV reduplicative template in *balik-balik* 'come back (continuative)' shows that the TSC also cannot be universal in such theories, since a medial C remains unfilled.

invariant is templatic; the singular-plural invariance is more like what happens in straightforward affixation.

Under the TSC, a template can only include those elements that are required in every expression of the morpheme at hand. It follows that the plural template must be the canonical iambic foot (here realized necessarily as CvCvv, as we explain later), which directly expresses the plural invariant and includes no mention of irrelevant "optional" material. In order to generalize over mono- and disyllabic input, we must apply the template not to the entire stem *but only to the first two moras of the stem*. From this limitation, all transfer effects will follow.

The formal details are pursued below, but even without them, it is possible to see how the analysis achieves its results:

Consider first a monosyllable like *nafs*, pl. *nufuus* 'soul'. Its first two moras are *naf*. Mapping them onto the iambic template in such a way as to satisfy the basic syllabic requirement of Arabic – syllables must have onsets –, while replacing the vowel melody with /u/ produces *nufuu*. In the context of the base form, this gives *nufuus*.

In a bimoraic disyllable like *ʔasad* 'lion', the first two moras are *ʔasa*. The mapping proceeds as with *nafs*, yielding *ʔusuu* to satisfy the template. *ʔusuud* is the complete form in context.

For *jundub* 'locust', the first two moras are *jun*. Of the plural vowel melody /a_i/, the /a/ spreads over the template, yielding *janaa*. The prosodically unaffected portion of the word picks up the /i/ giving *dib*. Taken together, we have *janaadib*.

For *sulṭaan* 'sultan', the first two moras are *sul*. Proceeding exactly as with *jundub*, the templatic segment *sul* emerges as *salaa*, and the extra-templatic final syllable becomes *ṭiin* absorbing the /i/ of the plural melody /a_i/ to give the plural *salaaṭiin*.

The two-mora limitation isolates the substring that changes, leaving prosodically unchanged the part that transfers intact. Discussion of formal matters, additional exemplification, and independent support are all dealt with below in section 3.

2.3 *Further Evidence of Iambicity*

Central to our account is the observation that the iambic plural is the only broadly-based, productive mode of plural formation in the language. Thus far, this point has been argued on the basis of evidence from the population statistics of the lexicon. Three further arguments establish the productivity of the iambic pattern.

The diminutive is an entirely productive, almost perfectly regular derivational process that is canonically nearly identical to the iambic plural. Compare the examples in (12):

(12) Diminutives

	Noun	*Plural*	*Diminutive*	
a.	CvCC			
	nafs	nufuus	nufays + at	'soul'
	qidħ	qidaaħ	qudayħ	'arrow'
	ħukm	/ħakaam/	ħukaym	'judgment'
b.	CvCvC			
	ʔasad	ʔusuud	ʔusayd	'lion'
	rajul	rijaal	rujayl	'man'
	ʕinab	/ʕanaab/	ʕunayb	'grape'
c.	CvCvvC + at			
	saħaab + at	saħaaʔib	suħayyib	'cloud'
	jaziir + at	jazaaʔir	juzayyir	'island'
	kariim + at	karaaʔim	kurayyim	'noble'
	ħaluub + at	ħalaaʔib	ħulayyib	'milch-camel'
d.	CvvCvC + at			
	faakih + at	fawaakih	fuwaykih	'fruit'
	ʔaanis + at	ʔawaanis	ʔuwaynis	'cheerful'
e.	CvvCv(v)C			
	xaatam	xawaatim	xuwaytim	'signet-ring'
	jaamuus	jawaamiis	juwaymiis	'buffalo'
f.	CvCCv(v)C			
	jundub	janaadib	junaydib	'locust'
	sulṭaan	salaaṭiin	sulayṭiin	'sultan'

The diminutive differs from the iambic plural only in the vocalization of the first and second (therefore templatic) syllables: diminutive $u+ay$ contrasts with plural $a+i$. All other aspects of plural formation are replicated in the diminutive as well – in particular, the plural data in (5-10), demonstrating transfer of various stem properties, are paralleled exactly in the diminutive, showing that the diminutive too is incompatible with simple mapping of root to template.

The particular significance of the diminutive is that it adopts the iambic mode of formation virtually without exception. Even individual nouns or noun patterns that never take iambic plurals do form diminutives on the iambic pattern

(13)	*Noun*	*Plural*	*Diminutive*	
	hind	hind + aat	hunayd + at	'woman's name'
	kaafil	kuffal	kuwayfil	'breadwinner'
	kaafir	kuffaar	kuwayfir	'infidel'
	rukb + at	rukab	rukayb + at	'knee'
	ħikm + at	ħikam	ħukaym + at	'maxim'
	waziir	wuzar + aaʔ	wuzayyir	'vizier'
	kitaab	kutub	kutayyib	'book'

The vitality of the iambic plural pattern is therefore confirmed: identical morphological resources are exercised in the indisputably productive and general diminutive, even in cases where iambicity is not an option in the plural.

A second structural argument for the core status of iambic pluralization comes from the plural-of-the-plural phenomenon. The plural-of-the-plural, if broken, is normally iambic, even when the base noun itself takes a non-iambic plural, and, strikingly, even when the simple plural is of a shape that would normally resist the iambic plural. The following examples illustrate these points:

(14)	*Singular*	*Pl.*	*Pl./Pl.*	
	ḍilaʕ	ʔaḍluʕ	ʔaḍaaliʕ	'rib'
	ʕuqaab	ʕuqb + aan	ʕaqaabiin	'eagle'
	jamal	jimaal	jamaaʔil	'he-camel'

ḍilaʕ and *ʕuqaab* form trochaic (metathesizing) and monosyllabic plurals, respectively, yet they assimilate to the iambic pattern in the plural-of-the-plural. The simple plural *jimaal is* a masculine noun CvCvvC, a type that very rarely takes an iambic plural. Nevertheless, its plural-of-the-plural is iambic. This shows that the iambic plural has regular, default status within the domain of broken pluralization, in much the same way that external suffixation has default status within the whole category of nouns. Deverbal nouns bypass the entire broken-plural system (just as denominal verbs bypass the ablaut system of English: *grandstanded*); similarly, broken plurals – as stems – bypass the lexically-restricted trochaic and monosyllabic subsystems of pluralization. The iambic plural is unambiguously the productive pattern, in that it alone applies to derived forms.

The final argument for the centrality and productivity of iambic pluralization comes from the treatment of loanwords. We have not studied the treatment of loans systematically in our Modern Standard Arabic lexical material, although it is immediately apparent that loanwords always form iambic plurals even when the singulars are canonically consistent with other modes of pluralization. Smeaton (1973) reports the results of a comprehensive study of loans in a conservative Saudi Bedouin dialect. The broken plurals of loans are always formed on the iambic pattern, even though other options are available in this dialect. In the following examples, we have changed Smeaton's transcription to conform to our own conventions and have abstracted away from the superficial phonology of the dialect:

(15)	*Singular*	*Plural*	
	baṣṣ	buṣuuṣ	'bus'
	bulṭ	buluuṭ	'bolt'
	rayl	ʔaryaal	'rail'
	huub (/huwb/)	ʔahwaab	'hub'
	ħafiiz	ħafaayiz	'office'
	dariiš + at	daraayiš	'window (Persian)'
	baarii	bawaarii	'bar'
	muuṭur	mawaaṭir	'motor vehicle'
	xaašuug + at	xawaašiig	'spoon (Persian)'
	bansil	banaasil	'pencil'
	bistin	basaatin	'piston'
	čingaal	čanaagiil	'fork (Persian)'
	dirwaaz + at	daraawiiz	'gate (Persian)'

The complete generality of the iambic plural is apparent from these examples, which range over all the stem canons.

3. The Iambic Plural And Diminutive in Detail

Our purpose now is to provide a comprehensive formal specification of the mechanisms of iambic plural, and by extension diminutive, formation. We first take up the major theoretical prerequisites to the analysis, prosodic specification of a domain and the overwriting of one vowel melody by another. We then apply these results to the detailed analysis of the iambic plural. Finally, we provide independent support within the language for two important features of the description, the bimoraic minimal word and the iambic template.

3.1. *Prosodic Specification of the Morphological Base*

Morphological processes usually operate on a base that is defined in purely morphological terms: e.g., suffix the plural morpheme to a noun *stem*. Even most kinds of reduplication and templatic morphology, though formerly regarded as "base-dependent" (Aronoff 1976), exhibit only morphological base dependency: copy and associate the *stem* melody; link the Arabic *root* to a template. Phonological processes in the lexicon also typically take units of morphological structure as the domain in which they operate, as with cyclic and edge-dependent rules. In certain circumstances, however, rules deviate systematically from this normal state of affairs by calling on phonological criteria to further delineate the base or the domain to which they apply. In this section we develop a theory of phonological circumscription of the morphological base, offering a formal account sufficiently detailed to support the analysis of Arabic.[6]

Rules of phrasal phonology provide a clear case of phonological circumscription, because their domains are obtained from a combination of grammatical and prosodic information (Selkirk 1984 *inter alia*). Within the lexicon – and therefore nearer to present concerns– a key instance is the use of extrametricality to redefine the location of an edge for purposes of a given rule or set of rules. According to the usual understanding, a single phonological constituent – segment, mora, syllable, foot – may be designated as extrametrical at an edge (Hayes 1982; Harris 1983), subtracting it from the morphological domain that the rules reckon with. Extrametricality has been extensively studied in phonology; we claim that it can be seen as well in the most commonly encountered variety of infixing morphology, turning prefixes and suffixes into infixes that stand one unit from an edge (McCarthy and Prince 1986, forthcoming b).

As a first step toward explicating the general phenomenon of prosodic circumscription of bases, let us develop an account of extrametricality. We need to clarify what it means to have an operation, either phonological or morphological, apply under an extrametricality constraint. Our tactic will be to characterize the analysis of the base implicit in the notion of extrametricality, and then to show how this analysis, when made explicit, allows us to generalize the ordinary notion of "operation on a base".

First, the implicit analysis of the base. We need to recognize a function which, given a base B, will pick out the part of B that is characterized by the extrametricality constraint. The function will return the designated

[6] Some aspects of our approach to formalizing the theory of prosodic specification have been influenced by Hoeksema's notion of a "head operation" (Hoeksema 1985; see also Pollard 1984, Bach 1979: p. 516; 1984).

constituent C that sits at the edge E of the base B. Call the function $\Phi\langle C, E\rangle$; let us write the result of applying the function to a base B as $B{:}\Phi\langle C, E\rangle$, or $B{:}\Phi$ for short. Thus, when a final syllable is rendered extrametrical in a given base B, $B{:}\Phi$ is the actual final syllable of B.

The function Φ induces a factoring on the base B, dividing it into two parts: one is $B{:}\Phi$, the part characterized by the constraint $\langle C, E\rangle$; the other is the residue, the part of B outside $B{:}\Phi$, which we will write B/Φ, recruiting Ajdukiewicz's familiar "fractional" notation for our purposes.[7] Defining an operator '*' that gives the relation holding between the two tween the two factors in B, usually left- or right-concatenation, we write:

(16) Factoring Imposed by Phonological Constraint

$$B = B{:}\Phi \, * \, B/\Phi$$

This simply gives us a way of showing how the criterion $\langle C, E\rangle$ divides up the base. In Latin stress, for example, where a final syllable is extrametrical, we have $\Phi\langle\sigma, \text{Right}\rangle$; if B is the word *spatula*, $B{:}\Phi = la$ and $B/\Phi =$ *spatu*. In this case, the operator '*' concatenates $B{:}\Phi$ to the right of B/Φ.

With this notion of factoring in hand, we can 'lift' the prosodic constraint from the level of the base to the level of operations applying to the base. Let $O(X)$ be a phonological or morphological operation defined on a base X. We define O/Φ – the same operation, but conditioned by the extrametricality of $\langle C, E\rangle$ – in the following way:

(17) Definition of Operation Applying under Extrametricality

$$O/\Phi\,(B) = B{:}\Phi \, * \, O(B/\Phi)$$

To apply O to B under extrametricality is just to apply O to B/Φ, concatenating the result with $B{:}\Phi$ in the same way ('*') that the residue B/Φ concatenates with $B{:}\Phi$ in the base B. In this way the operation O/Φ inherits everything that linguistic theory tells us about O, except its domain of application.

For Latin stress, the operation O is right-to-left assignment of bimoraic feet. Plugging our example word into (17), and using brackets to indicate

---

[7] The notation $B{:}\Phi$ abbreviates what would be written as $\Phi(C, E, B)$ in standard functional notation, which we eschew in aid of perspicuity.

Our use of the Ajdukiewicz notation should not be confused with that of Categorial Grammar, where it is used to define syntactical categories. (We are grateful to Diana Archangeli for recommending this clarification.) Categorial formalizations of some of the Broselow and McCarthy (1984) proposals can be found in Schmerling (1983) and Hoeksema and Janda (1988), the latter providing an unusually rich framework for morphological description.

foot-constituency, we have (18):

$$\begin{array}{llll} (18) & \mathrm{O}/\Phi\ (\text{spatula}) & = \text{spatula}{:}\Phi & *\ \mathrm{O}\ (\text{spatula}/\Phi) \\ & & = \text{la} & *\ \mathrm{O}\ (\text{spatu}) \\ & & = \text{la} & *\ \ [\text{spátu}] \\ & & = [\text{spátu}]\text{la} & \end{array}$$

Notice that the display in (18) is not a "derivation" in the usual sense, but simply a sequence of clarifications through which the meaning of each expression is brought out.

In the realm of morphological operations, a parallel is provided by the kind of infixation found with the Tagalog morpheme *-um-*, which marks actor focus in a certain class of verbs. The infix appears after the first consonant of the base: from *tawag* 'call', the perfective *tumawag* is formed. This, we suggest, is a case of prefixation under initial consonant extrametricality. Indeed, if there is no initial consonant, *um is* just a prefix. Since the extrametrical element is the word-initial consonant, the function Φ has its parameters set to ⟨Consonant, Left⟩. By equation (16), we have, for the factoring of the base:

$$(19) \qquad \text{tawag} = \text{tawag}{:}\Phi * \text{tawag}/\Phi = \text{t} * \text{awag}$$

Let UM- be the operation of prefixing *um* to a base. The operation UM/Φ then applies according to the definition (17):

$$\begin{array}{llll} (20) & \text{UM-}/\Phi\ (\text{tawag}) & = \text{tawag}{:}\Phi & * \text{UM-}(\text{tawag}/\Phi) \\ & & = \text{t} & * \text{UM-}(\text{awag}) \\ & & = \text{t} & * \text{um} \quad \text{awag} \\ & & = \text{tumawag} & \end{array}$$

Initial-consonant extrametricality therefore excludes the first C from the domain to which morphological concatenation applies, as desired.

Extrametricality is subtractive, identifying the domain of a rule with the residue left when a constituent is factored out. Less familiar, but no less widespread, is the use of a prosodic constraint to positively identify a rule's domain. This is fundamental to the way that iambic plural and diminutive morphology is imposed in Arabic. Here we will establish the basic formal and typological characteristics of the phenomenon.

One straightforward example, whose relevance to the Arabic case can hardly be missed, comes from Ulwa (Southern Sumu), a language of Nicaragua. Information on this language is due to Hale and Lacayo Blanco (1988); its theoretical significance was first noted by Bromberger and Halle (1988). Ulwa marks possession in nouns by a suffix or by an infix, depending on the phonology of the base (Hale and Lacayo Blanco 1988, vii). Possession by the 3rd person singular is exemplified in (21); the other person/number affixes are placed in the same position.[8]

(21) Ulwa Construct State

	Base	*Possessed*	
a.	al	al-ka	'man'
	bas	bas-ka	'hair'
	kii	kii-ka	'stone'
b.	sana	sana-ka	'deer'
	amak	amak-ka	'bee'
	sapaa	sapaa-ka	'forehead'
c.	suulu	suu-ka-lu	'dog'
	kuhbil	kuh-ka-bil	'knife'
	baskarna	bas-ka-karna	'comb'
d.	siwanak	siwa-ka-nak	'root'
	anaalaaka	anaa-ka-laaka	'chin'
	karasmak	karas-ka-mak	'knee'

The various possessive markers go right after the first iambic foot of the base. They appear as suffixes when the entire base is a single iamb, either monosyllabic (21a) or disyllabic (21b), and as infixes when the base is longer than one foot (21c, d). As usual, the iambic foot consists of a light syllable followed by a heavy syllable, two light syllables, or a single heavy syllable. (Since monomoraic feet are prohibited in the unmarked case, it is to expected that there are no monomoraic bases, a prediction borne out in the lexical material collected by Hale and Lacayo Blanco.) Stress in the language reflects the iambic foot pattern: it falls on the first syllalable if heavy, and on the second syllable when the first is light; the only complication is that it shifts to the initial syllable in disyllables. The possessive suffixes like *-ka*, then, when applied to longer stems, concatenate to the leftmost iambic foot in the base rather than to the whole base

[8] The infix *-ka-* also appears in nouns specified by a demonstrative, as in *aaka al-ka* 'this man'. For this reason, Hale and Lacayo Blanco call the *ka*-marked form the "construct state". The other possessive suffixes are *-ki* '1 sg.', *-ma-* '2 sg.', *-ni-* '1 incl.', *-kina-* '1 excl.', *-mana-* '2 pl.', and *-kana-* '3 pl.'. We have found a few exceptional affix-placements, typically suffixation to the whole base: *kruhbu-ka* 'ocelot', *paapangh-ka* 'father', *ta-ka-pas* 'mouth', *Ulwah-ka* 'Ulwa', *kubalamh-ka*, *kubalamhlamh-ka* 'butterfly'. Thanks to Ken Hale for supplying us with additional material on this language.

as a morphological entity. The prosodic criterion that factors the base is $\langle F_I, \text{Left}\rangle$, where F_I is the iambic foot. The function $\Phi\langle F_I, L\rangle$ will pick out the initial iambic sequence of the base, giving the factoring of 'knee' in (22) by definition (16):

(22) $$\begin{aligned} \text{karasmak} &= \text{karasmak}{:}\Phi * \text{karasmak}/\Phi \\ &= \text{karas} \qquad\quad * \ \text{mak} \end{aligned}$$

Here the operation must apply to the prosodically characterized segment B:Φ rather than to its residue. The formal definition exactly parallels that of O/Φ. We define an operation O:Φ, one that applies to a prosodic domain within a base, as follows:

(23) Definition of Operation Applying under
Positive Prosodic Circumscription
$O{:}\Phi\ (B) = O(B{:}\Phi) \ * \ B/\Phi$

To apply an operation O under prosodic circumscription to a base B is, by definition, to apply that operation to B:Φ, the segment of the base characterized by the prosodic constraint, joining the result with the residue B/Φ in the same way ('*') that B:Φ and B/Φ are joined in B itself. As with O/Φ, the operation O:Φ inherits all the formal characteristics of the unqualified operation O.

For the Ulwa possessive/construct, the operation is "suffix *-ka/kina/ni/*etc.", limited to the leftmost foot. Writing -KA for the operation, we have, for the word 'knee' factored in (22), the following application:

(24) $$\begin{aligned} \text{-KA}{:}\Phi\ (\text{karasmak}) &= \text{-KA}(\text{karasmak}{:}\Phi) * \text{karasmak}/\Phi \\ &= \text{-KA}(\text{karas}) \qquad\quad * \text{mak} \\ &= \text{karaska} \qquad\qquad\quad * \text{mak} \\ &= \text{karaskamak} \end{aligned}$$

Since prosodic circumscription is formally related to extrametricality, it is natural to ask whether its effects can also be seen in phonology. Stress theory, the homeland from which extrametricality emerged, provides a plausible candidate: the appearance of special prosody at the edges of words, typically encoded in non-iterative rules that deploy a single foot either initially or finally. If an ordinary iterative stress rule is subjected to prosodic circumscription, then it can appear to be non-iterative, since its actual domain is so narrow (cf. van der Hulst 1984: 165). Thus, a language which has only penultimate stress in the lexical phonology may be specifically limiting an ordinary iterative foot-building process to the rightmost disyllabic segment of the word. English provides another kind of example, for in that language quantity-sensitivity shows up only at the edges of the stress-domain (cf. Hayes 1982), plausibly a restriction of quantity-sensitive rules to a circumscribed subdomain chosen at edges. Commonly encountered general restrictions on the location of stress with respect to an edge, such as the familiar Dreisilbengesetz (the limitation of stress to one of the last three syllables, as in e.g., Modern Greek), may reflect prosodic circumscription as well as pure prosody. However, since our dominant immediate concern is with the morphology of circumscription, we will not explore the phonological question here.

Affix-placement effects quite similar to those in Ulwa have been noted in reduplication. In Samoan and Chamorro, reduplication inserts a light syllable in prestress position. Stress is typically penultimate in Chamorro, invariably so in Samoan (which treats each vowel as a separate syllable phonologically); for trisyllabic and longer bases, this leads to infixation:

(25) Reduplicative Affixation to a Prosodically Delimited Domain

a. Chamorro Continuative[9]

sága	*sa*[saga]	'stay'
égga	*e*[ʔegga]	'watch'
hugándo	hu*ga*[gando]	'play'
bidan + ñiha	bidan + *ñi*[ñiha]	'what they did; their doing'
inaligáo + ña	inali + *ga*[gao + ña]	'what he looked for; his searching'

b. Samoan Plural

táa	*ta*[taa]	'strike'
nófo	*no*[nofo]	'sit'
alófa	a*lo*[lofa]	'love'
galúe	ga*lu*[lue]	'work'
savali	sa*va*[vali]	'walk'

Following the lead of Broselow (1983: 338) and Broselow and McCarthy (1984), we analyze this as prefixation to the rightmost, main-stressed foot of the word. In Samoan, where all syllables are (C)V, the reduplicative prefix can just be a syllable; Chamorro, with a richer syllable inventory, requires the prefix to be specified as light. The Salishan languages show a range of similar phenomena (McCarthy and Prince forthcoming b). In all such cases, the sometime infix is a prefix to a prosodically characterized base; any residual material outside the actual base B:Φ is outside the scope of the operation, leading to infixation whenever the residue B/Φ is non-null.

Writing Σ- for the operation of prefixing a syllable template, the Samoan

[9] Thanks to Sandra Chung for help with the glosses.

infixing case comes about from the following applications of definitions (16) and (23):

(26) $$\begin{aligned} \Sigma\text{-}{:}\Phi\,(\text{alofa}) &= \Sigma\text{-}(\text{alofa}{:}\Phi) * \text{alofa}/\Phi \\ &= \Sigma\text{-}(\text{lofa}) \quad * \text{a} \\ &= \text{lolofa} \quad * \text{a} \\ &= \text{alolofa} \end{aligned}$$

The Ulwa, Samoan, and Chamorro cases display a property that is entirely characteristic of prosodic circumscription of the base: the prosodic criterion always selects the **minimal base** of the language. In the examples just discussed, the minimal base is descriptively coextensive with the foot. This is no accident. The prosodic hierarchy, as a principle of representational well-formedness, guarantees that words are made of feet, feet of syllables, syllables of moras. The minimal expansion of the category word, which we will denote by W_{min}, therefore consists of a single foot. With this in mind, we propose that positive prosodic specification of the base to which a rule applies is limited by the following constraint:

(27) W_{min} Constraint. Positive prosodic circumscription of a base may only appeal to the category Minimal Word. That is, in $O{:}\Phi\langle C, E\rangle$, $C = W_{min}$.

Considerations of prosodic theory lead to further conclusions about the nature of the W_{min} category. From numerous empirical studies, we know that the lower limit on the size of an unmarked stressed foot is 2 moras. Quantity- sensitive systems, iambic or trochaic, strongly avoid monomoraic feet, and quantity-insensitive systems, where each syllable may be regarded as monomoraic, strongly avoid monosyllabic feet.[10] The minimal unmarked foot, then, is 2 moras. The logic of markedness entails that simple reference to a unit calls the unmarked instance of that unit. Prosodic constraints are therefore framed in terms of unmarked units, and we deduce that whenever grammar invokes the minimal word condition, it

[10] In Hayes (1987), this observation is elevated to a principle absolutely excluding monomoraic stressed feet. This move, although appealing, encounters various empirical problems. On the one hand, it prohibits observed patterns like that of Maranungku (Hayes 1987), a left-to-right quantity-insensitive trochaic system where final odd-numbered syllables are reported to be stressed (e.g., [langka][rate][ti]). On the other hand, it excludes derivation-internal monomoraic feet which are later modified, like those of English (Hayes 1982), Cairene Arabic (McCarthy 1979), or Yup'ik (Hewitt 1989). It may well be that monomoraic feet have to be admitted as a marked option, with surface realization as stressed or unstressed governed by principles of eurhythmy such as those discussed in Prince (1983) and Hayes (1985).

sets 2 moras as the lower limit on word size. This consequence appears to be well-supported in a wide variety of languages, ranging from Estonian (Prince 1980) to Japanese (Itô 1988) to Arabic, as we show below (section 3.4).

Minimalization may extend over more than one level in the prosodic hierarchy. The minimal word must be a single foot; but the foot itself may or may not also be minimalized. This distinction will manifest itself whenever there are several possible expansions to the foot. Hypocoristics and truncated vocatives typically use a W_{min} template; crosslinguistic variation in the gross syllabic structure of such forms is due not only to the familiar differences in foot-types, but also to the degree of minimalization. In English, for example, hypocoristics are formed on the model of the monosyllabic word (McCarthy and Prince 1986, forthcoming b). The template's single bimoraic syllable is the minimal **foot** as well as the minimal word, so that minimalization extends all the way down the prosodic hierarchy. In Yup'ik (Woodbury 1985), on the other hand, the proximal vocative template is a single iambic foot, which may be freely taken to be one or two syllables; here the word is minimalized only to the foot level. (Examples of the English and Yup'ik phenomena appear below in (58).) Arabic will not provide us with examples that distinguish one-level from all-level minimalization, since the minimal word is fixed at two moras.

In the cases examined so far, affixation to a prosodically specified base leads to infixation, an eye-catching order effect. In reduplication, where the copying component of the operation is inherently sensitive to base phonology, prosodic circumscription can to lead to subtler, but ultimately no less striking consequences. Important evidence bearing on the nature of prosodic specification was first noted by Nash (1979; 1980: 144) in a discussion of the problem of Yidin^y reduplication. The key data are exemplified here:

(28) Yidin^y Reduplication[11]

Singular	*Plural*	
mulari	mula + mulari	'initiated man'
jugarba	jugar + jugarba-n	'have an unsettled mind'
gindalba	gindal + gindalba	'lizard species'
kalampaRa	kala + kalampaRa	'March fly'

Yidin^y reduplication exhibits a special kind of dependence on the form of the base. The base *mulari* is syllabified as /mu.la.ri/; the affiliation of

[11] The symbol *R* spells the retroflex continuant. All data are from Dixon (1977), a major descriptive and analytic study of the language.

r with the third syllable of the base is what prevents it from being copied – compare *jugar + jugarba-n*, whose base is /ju.gar.ba/. The base *gindalba* is syllabified as /gi.ndal.ba/; the *l* belongs to the second syllable of the base and therefore it can be copied. The base *kalampaRa* is syllabified as /ka.la.mpa.Ra/ because all homorganic nasal-stop clusters are tautosylla bic, or perhaps not clusters at all, as Nash (1979) argues; this form therefore reduplicates like *mulari*.

The curious property of Yidin^y reduplication is the way that the syllabification of the base is carried over, as if the initial disyllabic sequence were copied whole. A large amount of descriptive research has failed to turn up a reduplicative process that unambiguously *copies* a single syllable (Moravcsik 1978). There are no known cases where, under the same rule, a form like *ta.ka* reduplicates as *ta + taka* and a form like *tak.ta* as *tak + takta*. The actual possibilities are *ta-* for both, with the prefix in the shape of a light syllable, or *tak-* for both, with the prefix specified as a heavy syllable (or just a syllable). Indeed, this very finding lies behind the development of the templatic approach to reduplication (Marantz 1982). In templatic reduplication, the syllabic character of the affix determines all structural properties of the result, whereas the syllabic character of the base contributes nothing. Yidin^y provides us with a striking counterexample to the empirical generalization. Why then should reduplication appear to copy *two* syllables but never just *one*?

Our explanation is that the disyllabic foot of Yidin^y is the actual domain of reduplication.[12] The foot is quite literally the minimal base of Yidin^y, since no stem may be monosyllabic. The Yidin^y reduplicative prefix attaches to the minimal base within the actual base, reduplicating the minimal base just as if it were an authentic morphological unit. Only material contained in the minimal base – the first two syllables of the stem – is available for copying. Normal reduplication, without base specification, quite freely changes syllabic affiliations and would derive forms like **mular + mulari*. Indeed, in the remotely related language Lardil (Wilkinson 1986), which uses a similar form of prefix, we find exactly that: *parel + pareli*, **pare + pareli*. Only the prosodic characterization of the base as minimal, specific to the grammar of Yidin^y, prevents this.

Since the whole of the minimal base is reduplicated, the Yidin^y reduplicative affix can be regarded as W, or total reduplication, a form of com-

[12] Nash (1979; cf. Nash (1980:139)) proposes that the actual Yidin^y operation is a foot-copying transformation, an impossibility in our theory, although foot reduplication by autosegmental spreading had been proposed in McCarthy (1979). Nash (1980: p. 144) alludes to an analysis somewhat closer to the one developed here.

pounding. Writing W- for the operation of prefixing a word template, a typical Yidin^y form is derived as follows:

(29) W-:Φ (mulari) = W-(mulari:Φ) * mulari/Φ
= W-(mula) * ri
= mulamula * ri
= mulamulari

In Lardil, it is the prefix, not the domain, that is characterized as W_{min}. The largest segment of the base melody *pareli* that satisfies the minimal word requirement is *parel.*[13] Given that the unmarked word-shape of the language is exactly the foot-long W_{min}, it should follow from markedness theory that unadorned reference to W- is implicitly reference to W_{min}. Thus, both Lardil and Yidin^y can have the prefix W-, understood via markedness to call W_{min}-. The two languages differ only in that Yidin^y adopts prosodic circumscription as well.

The minimal base has also been found to serve as the locus for a kind of templatic or superpositional morphology much more obviously like the Arabic broken plural. The Cupeño habilitative, studied by Hill (1970) and McCarthy (1984), provides a clear example. The facts are in (30), where square brackets indicate foot-edges:

(30) Cupeño Habilitative

	Verb Stem	*Habilitative*	
a.	[čál]	čáʔaʔal	'husk'
	[tə́w]	tə́ʔəʔəw	'see'
	hə[lʸə́p]	həlʸə́ʔəʔəp	'hiccup'
	kə[láw]	kəláʔaʔaw	'gather wood'
b.	[páčik]	páčiʔik	'leach acorns'
	[čáŋnəw]	čáŋnəʔəw	'be angry'
c.	[pínəʔwəx]	pínəʔwəx	'sing enemy songs'
	[xáləyəw]	xə́ləyəw	'fall'
d.	čí	číʔ	'gather'
	hú	húʔ	'fart'
	ʔáyu	ʔáyu	'want'

Descriptively, the generalization is that, for consonant-final words, the

[13] McCarthy and Prince (1986) show that satisfaction of templatic constraints is typically maximal in that the affix subsumes the largest substring of the base melody meeting the templatic requirement.

habilitative is formed by expanding the stress foot to three syllables in length. The location of stress in stems is lexically determined, according to Hill (1970). If the foot is monosyllabic (30a), two syllables are added; if disyllabic (30b), one syllable is added; and if already trisyllabic (30c), the habilitative is the same as the base. The data in (30d) show that vowel-final words are not subject to this process.

McCarthy (1984) analyzes this phenomenon as accommodation of the base to a trisyllabic foot template.[14] But the base to which this process applies is not purely morphological in character; material outside the actual stress-foot of the word is irrelevant, so that *kə[láw]*, with final stress, goes to *kəláʔaʔaw*, not **kəláʔaw*. The foot-maximizing template is superimposed on the foot – the minimal base – within the morphological stem. The template-mapping operation therefore applies under $\Phi\langle C, E\rangle$, where C is the minimal word category, as predicted by constraint (27).

The Cupeño habilitative also supplies evidence on how the function $\Phi\langle W_{min}, E\rangle$ interprets its factoring mission. The process resembles Arabic iambic plural and diminutive formation in that it selects a minimal-word-sized segment as the domain of templatic morphology. In Cupeño (or Yidinʸ, for that matter), the selected segment is an actual foot of the word, but in Arabic the surface stressing is irrelevant to the Φ-parse of a form – only the moraic structure counts. There are two possible sources for this difference. First, the edge specification in Φ may be opposite to that of foot formation, so that there is no foot at the edge where Φ seeks one. Second, the morphological processes depending on Φ may apply to representations on which feet have not yet been imposed. The Cupeño habilitative is clearly formed after stress has been assigned; the Arabic processes plausibly apply in strata before the relatively superficial determination of word stress. Significantly, no known processes of Classical or Standard Arabic phonology are stress-dependent; and in modern Arabic dialects, with an essentially identical grammar of the iambic plural, stress-sensitive phonological rules and indeed stress itself are applied to the output of broken plural formation. Thus, the ordering of iambic pluralization prior to stress assignment accords fully with the facts. Either condition – specification of the opposite edge, or ordering morphological processes before foot assignment – leads to the following conclusion: in the presence of relevant prosodic structure, Φ will select the actual element

---

[14] Why is the trisyllable recognized as a legitimate category? The Cupeño template may be the *maximal* prosodic unit of the language: one (necessarily binary) foot+one extrametrical syllable, which itself ends in an extrametrical consonant.

that satisfies its parameters. In the absence of relevant structure, Φ is free to parse the desired constituent from the underspecified base.

Circumscription is an activist strategy that allows the longer bases in a language to be treated on the model of the minimal base, preserving the key phonological and morphological properties that happen to coincide in prosodically minimal forms. An equally significant effect of prosodic circumscription is precisely to disallow generalization beyond the minimal: to demand the coincidence of foot and word as a precondition for rule application. In English, for example, irregular verb allomorphy is entirely restricted to monosyllabic stems, the fully minimal word structure of the language. Comparative and superlative *-er* and *-est* are used on one-foot adjectives: *bigger*, *stupider*, *happier*, but *?intenser*, **auguster*, **intelligenter.*[15] In Arabic, only minimal (bimoraic) nouns take special melodies in the broken plural (see section 5.2 below). In Kinande, trisyllabic forms simply do not reduplicate (Mutaka and Hyman 1987); in the Northern Karanga dialect of Shona, trisyllabic stems reduplicate postlexically, showing different tonology from the canonical disyllabic forms (Hewitt and Prince 1989). Examples can be multiplied ad lib. Once again the distinction is between the minimal word and all others, but here the prosodic criterion serves to segregate the two classes, not to join them.

Such cases require that $B = B{:}\Phi$ in order for the operation to proceed. No active parse is allowed, or, to put it another way, the parsing operation must be vacuous. We can incorporate this idea into the formal theory by admitting a variant definition of Φ that turns it into a **partial** function, one that does not return a value for every member of the set of bases.[16] Let us designate this variant as Φ'.

(31) Definition of Partial Function Φ'

$$B{:}\Phi' = B \quad \text{if } B = B{:}\Phi$$
$$\text{else, undefined.}$$

The prosodically restricted operation $O{:}\Phi$ depends on the success of the function Φ, and $O{:}\Phi'$ is therefore undefined when Φ' is. An operation applying under Φ' applies only to minimal words.

The segregating effects of prosodic circumscription can be subtler than total blockage; it can also determine choice among allomorphs. One large

[15] For the comparative and superlative of English, minimalization extends down one level – W_{min} is a single foot. For the deadjectival suffix *-en*, which attaches only to monosyllables (*whiten* vs. **liciten*), minimalization goes down two levels – W_{min} is a single foot which is itself minimal (a stressed monosyllable).

[16] We note an abstract similarity between the feature-changing/feature-filling distinction and the distinction between the total function and the partial function interpretations of Φ.

class of cases is represented by an idiosyncratic alternation in the Dyirbal ergative suffix (Dixon 1972). The ergative suffix is +*ŋgu* with disyllabic bases but +*gu* with longer ones:

(32)	*Root*	*Ergative*	
a.	/yara/	yara + ŋgu	'man'
b.	/yamani/	yamani + gu	'rainbow'

In this and many other languages, one allomorph of the affix attaches to a base of minimal prosodic size – in Dyirbal, disyllabic.[17] The other allomorph is applied in the default case, affecting all other bases. The morpheme -*ŋgu* suffixes under Φ'; prosodically unrestricted -*gu* is limited only by the Elsewhere Condition.[18]

A related phenomenon appears frequently in the reduplicative morphology of a number of Austronesian languages. In Cebuano, reduplication acts in a radically different way depending on whether the base is minimal (disyllabic) or not (Wolf 1966:562–3), as the following data show:

(33) Cebuano Disyllabic Reduplication

a. *Minimal base*

sulti	sulti + sulti	'talk'/continuative
balik	balik + balik	'come back'
higdaʔ	higdaʔ + higdaʔ	'lie'

b. *Nonminimal base*

balibad	bulu + balibad	'refuse offering'
paŋutana	pulu + paŋutana	'ask question'
panaŋhid	pulu + panaŋhid	'ask to leave'
pahulay	pulu + pahulay	'rest'

---

[17] The ergative suffix of Dyirbal looks as if it is in a compensatory relationship with the base: the suffix is shorter with longer bases. Armin Mester has proposed that this observation derives from a templatic requirement that the result of affixation be a prosodic compound – which must contain two (at least minimal) words, thus at least four moras. The rule of allomorphy, then, picks the shortest affix allomorph that satisfies this requirement. The assumption is that +*ŋgu* be bimoraic, +*gu* monomoraic. This is an elegant solution, and an approach of this formal character will certainly illuminate other cases of compensation, as shown in the discussions of Ponapean and Sanskrit in McCarthy and Prince (1986), but it encounters difficulties with Dyirbal or the related phenomena in other Australian languages. Quite generally in these languages only vowels can be moras; further, in some of them, nasal-stop sequences are probably onsets or single segments (Nash 1979).

[18] If the phonological similarity between the allomorphs is linguistically significant, the process can be divided into two parts: suffixation of a nasal element under Φ' and general suffixation of -*gu*. Or the morpheme could be taken to be /-ŋgu/, and the allomorphy process would be DELETE-ŋ/Φ – that is, deletion would only apply when the *ŋ* sits in B/Φ, outside the minimal word.

With disyllabic bases, reduplication is total; with polysyllabic bases, the reduplicative prefix is *Culu*. Less radical versions of this choice of reduplicative affix appear in Tagalog and Makassarese (Aronoff *et al.* 1987, citing Carrier-Duncan (1984) for Tagalog), languages in which the minimal base is also disyllabic:

(34) Tagalog and Makassarese Disyllabic Reduplication

a. Tagalog

i. *Minimal Base*

mag-li:nis	mag-li:nis + li:nis	'clean/clean a little'
mag-walis	mag-walis + walis	'sweep/sweep a little'
pantay	pantay + pantay	'level/quite level'

ii. *Nonminimal Base*

tahi:mik	/tahiʔ/ + tahi:mik	'quiet/rather quiet'
baluktot	/baluʔ/ + baluktot	'bent/variously bent'
kalansiŋ	/kalaʔ/ + kalansiŋ	'jingle of coins/id.'

b. Makassarese

i. *Minimal Base*

bállaʔ	ballaʔ + bállaʔ	'house/little house'
gólla	golla + gólla	'sugar/sweets'
táu	tau + táu	'person/doll'

ii. *Nonminimal Base*

kalúarak	kaluʔ + kalúaraʔ	'ant/lots of little ants'
manára	manaʔ + manára	'tower/sort of tower'
baláo	balaʔ + baláo	'rat/toy rat'

As in Cebuano, minimal bases reduplicate totally. Supraminimal bases also take a different form: they have disyllabic reduplication with final ʔ. The ʔ is realized as vowel length in Tagalog by an independently motivated rule of syllable-final ʔ-deletion with compensatory lengthening. In Makassarese, the ʔ assimilates to a following voiceless consonant but is otherwise phonetically apparent.

All three of these cases could be treated as suppletive allomorphy, requiring one morpheme (total reduplication) under Φ', and another elsewhere.

However, the minimal and supraminimal allomorphs are conspicuously similar, sharing disyllabicity and differing only where the supraminimal allomorph has some form of additional melodic specification. As in templatic morphology proper, a distinction must be made between the structural template and any melodies that are associated with it. The reduplicative template itself is always W-, the unmarked, therefore minimal word of the language. One allomorph is further distinguished by an accompanying purely melodic morpheme, /ʔ/ in Tagalog and Makassarese, /ulu/ in Cebuano. The simple allomorph, with no melodic content, is attached under Φ', applying only to minimal words; the more complex allomorph attaches elsewhere. The melodic material associates from right-to-left with the reduplicative template (as an autosegmental suffix), displacing any competing segmental material from the base. In Tagalog, a stem like *baluktot* would lead to a W- prefix *baluk-*, but syllabic integration of the ʔ melodic suffix supplants the final consonant, giving *baluʔ* and eventually *balu:*. In Cebuano, the melody *ulu* usurps all syllabic positions except the absolute initial one.

Makassarese has a couple of further wrinkles that are worthy of notice. First, the morphological base to which the rule applies is the root, even though reduplicative affixation takes place at the level of the stem, which may include a suffix. Aronoff et al. (1987) point to the following contrast, involving the transitivizing suffix *-i*:

(35) Morphologically Complex Forms in Makassarese

Root	*Stem*	*Reduplication*	
lómpo	lompói	lompo-lom.pó.i	*'big/make big/make somewhat big'*
gássiŋ	gassíŋi	gassiʔ-gas.sí.ŋi	*'strong/make str./make sw. str.'*

Note that the mere presence of *-i* is not enough to invoke the supraminimal allomorph. Although both roots are disyllabic, they choose different reduplication patterns when suffixed by *-i*, because the suffix affects their syllabification differently. The root *lompo* passes the Φ' test, $B = B{:}\Phi$, because in *lompoi* the initial disyllabic sequence $B{:}\Phi =$ *lompo*. The root *gassiŋ* fails the Φ' test when suffixed by *-i*, because the initial disyllabic element $B{:}\Phi =$ *gassi* and this is not equal to the root. The base B relevant to the formula $B = B{:}\Phi$ is therefore the root within the stem, but it is the phonology of the stem that determines the Φ-parse.[19]

Second, stress in Makassarese is penultimate, as is typically the case where W_{min} is disyllabic, but the actual surface stressing of the word does not affect the Φ-parse, which seeks out the leftmost W_{min}, not the rightmost, where the stress is to be found. This is consistent with our interpretation of the parsing power of Φ, since the edge specifications on stress

[19] Thus Makassarese reduplication is, morphologically speaking, a head-operation in the sense of Hoeksema (1985), as noted in Aronoff (1988). Although it must apply at the level of the stem, it applies only to the root inside the stem – the head of the stem. On top of that, it applies to the root as prosodically circumscribed entity.

and on Φ are not identical.[20] But, in accordance with our earlier dictum that "in the presence of relevant prosodic structure, Φ will select the actual element that satisfies its parameters", the pre-existing *syllabic* structure of the base is respected, yielding the contrast noted in (35).

The requirement $B = B{:}\Phi$ segregates the set of bases into two contrasting classes by limiting the definition of Φ to just one of the classes. There is another situation where the character of the base can disturb the smooth functioning of $\Phi\langle C, E\rangle$: when the base fails to contain anything meeting the description 'constituent C at edge E', so that $B{:}\Phi$ is empty. This may not require special comment in the case of $O{:}\Phi$, an operation applying to a circumscribed base, since bases are typically minimal or bigger and, further, many operations will themselves be undefined on the empty string. Linda Lombardi has pointed out to us, however, that a condition $B{:}\Phi \neq e$ (that is, the parsed-out constituent must be non-null) will have an important effect on O/Φ, operations applying to the residue left by extrametricality: it will force extrametricality to be obligatory. Recall the definition of O/Φ, repeated here:

(36) Definition of Operation Applying under Extrametricality
$O/\Phi\,(B) = B{:}\Phi * O(B/\Phi)$

If $B{:}\Phi$ may not be null, then the element $\langle C, E\rangle$ mentioned in Φ must be present for the operation to proceed.

Obligatory extrametricality is not likely to be obvious in run-of-the-mill stress systems, because the prosodic hierarchy demands that stress rules succeed over the whole vocabulary. But we do find a restriction of this sort in the Arabic lexical requirement that all stems end in consonants and that those consonants be extrametrical. Each lexical entry is subject to an identity rule, $I/\Phi''\langle\text{Consonant, Right}\rangle$, where Φ'' denotes the variety of Φ that is undefined when $B{:}\Phi = e$.

More striking evidence for obligatory extrametricality comes from the Cupeño habilitative. The forms above in (30d) show that a final consonant must be present for template mapping to proceed at all. (30a, b) show that the final consonant does not itself participate in left-to-right template

[20] It is not clear, from our limited current knowledge of the language, whether reduplication must follow stress assignment. Forms ending in *l.r.s* are subject to a rule of epenthesis that follows stress assignment, giving rise to surface antepenultimate stress: /tetter/→*tétter*→*téttere?*. These forms count as supraminimal for reduplication: *tette?-téttere?* 'rather quickly'. This would seem to establish the order stress<epenthesis<reduplication. However, the mere fact of epenthesis seems to indicate that *l.r.s* are extrasyllabic word-finally (Aronoff et al., 1987) consequently, $B{:}\Phi \neq B$ for such forms at any point, since e.g., *tette.r*:Φ = *tette*.

mapping; instead, it is tacked on to the end of the trisyllabic foot template. In other words, the final consonant is extrametrical with respect to template mapping, the treatment of vowel-final stems showing that this extrametricality is obligatory.

The analysis of Cupeño thus involves *both* extrametricality and positive prosodic circumscription. A certain subtlety of attack is required, but the means are at hand. Let T be the operation of mapping to the trisyllabic template. In the simple cases, where the base *is* a foot, we want it to apply in the mode $T/\Phi''\langle \text{Cons, Right}\rangle$ – ignore the final consonant, map the residue to the template – as the following derivations show:

(37) Cupeño Template Mapping – Easy Cases

$$
\begin{array}{llll}
\text{(a)} & T/\Phi''(\text{čal}) & = T(\text{čal}/\Phi'') & *\,\text{čal}{:}\Phi'' \\
 & & = T(\text{ča}) & *\,\text{l} \\
 & & = \text{čaʔaʔa} & *\,\text{l} \\
 & & = \text{čaʔaʔal} & \\
 & & & \\
\text{(b)} & T/\Phi''(\text{ʔayu}) & = T(\text{ʔayu}) & *\,\text{ʔayu}{:}\Phi'' \\
 & & = T(\text{ʔayu}) & *\,\text{undefined} \\
 & & = \text{undefined} &
\end{array}
$$

It is the operation T/Φ'', and not just simple T, that must be restricted to W_{min} domain. We therefore write $(T/\Phi''){:}\Phi$ to represent the whole process of Cupeño habilitative formation, composing the two forms of phonological specification. Applying the complex operation calls on each of the definitions used in the theory of prosodic circumscription:

(38) Cupeño Template Mapping in Full

$$
\begin{array}{lllll}
(T/\Phi''){:}\Phi\ (\text{kəláw}) & = T/\Phi''\ (\text{kəláw}{:}\Phi) & & *\,\text{kəláw}/\Phi & \text{by defn. (23)} \\
 & = T/\Phi''\ (\text{láw}) & & *\,\text{kə} & \\
 & = \{T(\text{láw}/\Phi'') & *\,\text{láw}/\Phi''\} & *\,\text{kə} & \text{by defn. (17)} \\
 & = \{T(\text{lá}) & *\,\text{w}\quad\} & *\,\text{kə} & \\
 & = \{\text{láʔaʔa} & *\,\text{w}\quad\} & *\,\text{kə} & \\
 & = \text{láʔaʔaw} & & *\,\text{kə} & \\
 & = \text{kəláʔaʔaw} & & &
\end{array}
$$

If the operation $T{:}\Phi$ were restricted by necessary extrametricality, yielding $(T{:}\Phi)/\Phi''$, the same output would result. For this kind of case, functional composition commutes.

To sum up: the theory of prosodic specification developed here is built on the function $\Phi\langle C, E\rangle$ that returns a constituent C at the edge E of a base. The theory's descriptive capacity devolves from two main cross-classifying parameters involving the function Φ.

(1) An operation may be specified as O:Φ, applying to the section of the base picked out by Φ, which we write as B:Φ, just as if that section were a true morphological domain. An operation may be specified as O/Φ, in which case it applies to the residue B/Φ that is left when B:Φ is ignored in B.

(2) The function Φ may be total, always returning a value. The function Φ may be partial, defined on only a subset of possible bases. The partial function gives rise to allomorphy or to simple blockage of a rule. Furthermore, Φ may be partial in either of two ways, which are intrinsically relevant to different cases. It may be defined only when B = B:Φ, limiting O:Φ to minimal bases, or it may be defined only when B:Φ ≠ *e*, limiting O/Φ to words which actually have the element C present at edge E of B.

This gives a taxonomy of four cells, into which we insert the examples that have been discussed:

(39) Taxonomy of Prosodic Specification.

a. Φ *is Total*

(i) *O:Φ*. Apply an operation to prosodically circumscribed base.
Ulwa *-ka* suffixing infixation (21)
Non-iterative Stressing
Chamorro Continuative prefixing infixation (25a)
Samoan Plural prefixing infixation (25b)
Yidin^y Reduplication (28)
Cupeño Habilitative (30)
Arabic iambic broken plurals

(ii) *O/Φ*. Apply an operation to the residue of a constituent.
Stress with Extrametricality (18)
Tagalog *-um-* prefixing infixation (20)

b. *Φ is Partial.*

(i) *O:Φ′*. Apply an operation to the minimal base; do something different to the others, or do nothing to them.
English Irregular Verb Morphology
English *-er/-est* suffixation
Arabic Plural Melody assignment (section 5.2)
Dyirbal Ergative Allomorphy (32)
Cebuano Plural Reduplication (33)
Tagalog Disyllabic Reduplication (34a)
Makassarese Reduplication (34b)

(ii) O/Φ'. Apply an operation to the residue if a constituent C is present at edge E; else do something different to the others, or nothing.
Lexical Final C requirement
Cupeño Habilitative (37, 38)

As a final observation, we note that'the basic ideas presented here can be given an alternative theoretical development in terms of the effect of Φ on morphological *domains* rather than *operations*. Instead of defining O:Φ and O/Φ, we could define D:Φ and D/Φ, where D represents the morphological category that B belongs to. Under this approach, D:Φ would have B:Φ as an actual domain within D, and rules looking to apply to D would apply instead to B:Φ. Similarly, D/Φ would have B/Φ marked as the actual domain to which rules applied. Definitions would proceed as as in (40):

(40) Domain Re-structuring Theory of Prosodic Specification
(a) $[_D B]{:}\Phi = [_D [_D B{:}\Phi] * B/\Phi]$
(b) $[_D B]/\Phi = [_D B{:}\Phi * [_D B/\Phi]]$

These definitions create a Chomsky-adjoined type of structure $[_D \ldots D \ldots]$, where the contents of the inner D are prosodically specified. The theory of rule application would then contain the following specification, determining how operations apply to such a structure:

(41) Rule Application in a Domain-Restructuring Theory
$O([_D \ldots D \ldots]) = [_D \ldots O(D) \ldots]$

The effect of definition (41) could perhaps be derived from considerations of cyclicity. This approach more closely resembles the characterization of extrametricality in Selkirk (1984) and the quasi-formal treatment of both circumscription and extrametricality phenomena in McCarthy & Prince (1987a, b). One problem that asserts itself immediately is how the **blocking** effect of specification would be handled; if Φ is partial, restructuring would surely fail, but the original domain structure would still invite ordinary rule application. Another fundamental issue is whether prosodic specification turns out to delimit individual operations (as assumed above) or whole sets of operations that belong to a single region of the grammar, in which case it is more properly attributed to the structure of the representations that those operations access. Recent argumentation bearing on this matter is found in Inkelas (1989). For present purposes, it is unnecessary to answer these questions definitively, and we will proceed with the operational approach, which brings the surprisingly wide range of specification effects under a single general theory.

3.2 *Melodic Overwriting*

The earliest CV-theory treatments of templatic and reduplicative morphology (McCarthy 1979, Marantz 1982) recognize a special melody/template relation called prespecification, whereby invariant prior linking of a melodic element to a templatic position overrides or supplants productive, rule-governed linking of a melodic element to the same position. For example, Marantz analyzes the *Ci* reduplication of Yoruba (*lo*, *lilo* 'to go/going') with a CV prefix whose V is prelinked to the invariant *i*.

There is considerable evidence, discussed in McCarthy and Prince (1986, forthcoming b), that the phenomenon of melodic invariance in reduplicative affixes cannot be reduced to prespecification. This evidence comes in part from so-called echo words, a type of total word reduplication in which some systematic change is effected in one copy. Echo word formation seems to be nearly universal; it is found in English (*table* + *shamble*) or, with more instructive results, in the Dravidian language Kolami (Emeneau 1955):

(42)	pal	pal + gil	'tooth'
	kota	kota + gita	'bring it!'
	iir	iir + giir	'water'
	maasur	maasur + giisur	'men'
	saa	saa + gii	'go (cont. ger.)'

The entire word is reduplicated with the initial CV of the second copy fixed at *gi*. The sequence *gi* appears even when the original is vowel-initial, and the vowel *i* occupies both moras of an original long vowel.

This widespread phenomenon is incompatible with templatic prespecification. Since the reduplication copies the whole word rather than some substring of it, the reduplicative affix (in this case presumably a suffix) must be the prosodic word W (Marantz 1982). (Indeed, it might be better to say that there is no reduplicative affix at all, and the basic process is the purely morphological one of compounding the base with itself – tautologous compounding (McCarthy and Prince 1988).) To what, then, would the melodic invariant *gi* be prelinked in the reduplicative affix, as prelinking theory requires? The grammar does not enumerate the terminal elements of the reduplicative affix W (or some X = N, A, V, P) – it cannot, since W has infinitely many terminal elements – yet it is exactly to those terminal elements that the melodic invariant *gi* would have to be prelinked. Needless to say, this problem exists independently of the choice of terminal elements: syllables, moras, onsets, nuclei, or segments all are unboundedly many in W. Moreover, even if we somehow did manage to enumerate the terminal elements of W, we would then need to prespecify templatic VV as *i* to obtain *maasur* + *giisur*, but this would wrongly predict long *ii* in all cases, yielding **kota* + *giita*. Clearly some other process is at work here.

Following McCarthy (1979: p. 319) and McCarthy and Prince (1986), we propose that we are witnessing here the same kind of melody-to-template mapping that is visible in standard root-and-template morphology. The melody *gi* has an autonomous status as a purely melodic entity with its own autosegmental plane, just like *ktb* or *a_i* in the Arabic verbal system; the difference is that *ktb* and *a_i* are mapped to empty templatic slots in a "feature-filling" fashion, whereas the melody *gi* is applied in a "feature-changing" manner, overwriting the original melodic material of the base.

The echo morphology of Kolami, then, consists of tautologous compounding, plus the melodic echo morpheme *gi*, along with the information that this melody links to the second member of the compound. The base itself supplies the array of prosodic positions that the melody anchors to. Coming in on its own plane, with free access to the prosodic positions of the base, the melodic morpheme associates in the usual left-to-right fashion, delinking the base phonemes as it goes. As with feature-filling association in Arabic, the vocalic melodeme must link to both vocalic moras in a heavy syllable, so that we obtain *maasur* + *giisur* rather than **maasur* + *giasur*. From this interpretation of melody-to-template mapping, which is inevitable in the context of recent rule typology, melodic invariance follows without prespecification. Within the theory of Prosodic Morphology, there is the further prediction that prosodically null positions like the onset may be supplied by melodic overwriting, so that *iir* + *giir* is possible, while prosodically genuine positions – like a long vowel or a moraic coda consonant – cannot be an invariant part of echo formation. Only templates, not melodies, can supply invariant prosody. Thus, we predict the non-existence of an echo-word system that takes arbitrarily long input and that specifies both the quality and the quantity of some segment in the output (e.g., an echo-word system with *kota* → *kota-giita* and *koota* → *koota-giita* or one with *kota* → *kota-gita* and *koota* → *koota-gita*). So far as we know, this prediction is borne out.

Melodic overwriting is important to determining the vocalism of Arabic iambic plurals and diminutives. Although the portion parsed by the bimoraic minimal word and assigned to the iambic template could receive vocalism in one of several ways, the residue cannot. Its vocalism is regularly overwritten by /i/ in both iambic plural and diminutive. This overwriting shows the same preservation of vowel length observed in Kolami: contrast the plurals *janaadib* and *salaaṭiin*.

3.3. *Synthesis and Exemplification*

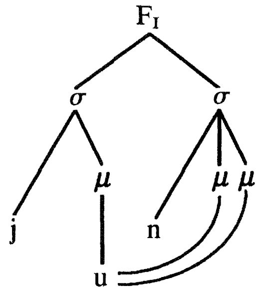
We now have the theoretical resources to deal with the details of the Arabic iambic plural and diminutive. The prosodic hierarchy tells us that the minimal word is a single foot; as we show below in Section 3.4, the stress foot of Arabic is a moraic trochee ($\mu\mu$), and a great deal of evidence converges on this bimoraic sequence as the minimal word of the language. The operation of template mapping in the broken plural and diminutive exhibits "positive prosodic circumscription", and we have observed that the prosodic constituent that specifies such a domain is without exception the minimal word (see (27)). The general left-to-right bias in Arabic melody mapping (McCarthy 1981) suggests that the parsing of the base by the minimal word should be in that direction, and this is just what is required. The criterion of phonological specification is thus $\Phi\langle W_{min}, L\rangle$, with the template-mapping operation applied according to the definition in (23).

The iambic foot, identified by metrical theory as the disyllabic sequence light-heavy, is the sole invariant in the plurals and diminutives. (It is also, as we show below in section 3.5, an important feature of Arabic prosody in other domains.) Thus, the iambic foot is the template to which the contents of the prosodically circumscribed domain B:Φ will be mapped. Material outside the minimal-word domain – in the residue B/Φ – will be unaffected by this template-mapping morphology.

In the plural, the melody /a_i/ is introduced and is subject to a rule spreading the /a/ across the template. When the residue B/Φ becomes accessible after template mapping, the /i/ links if possible (that is, if the residue contains any metrical moras), overwriting the preexisting vocalic melody. Otherwise /i/ deletes by Stray Erasure. In the diminutive, an autosegmental suffix *y* fills the last mora of the template. The diminutive melody /u_a_i/ then associates one-to-one and left-to-right, as usual, leaving the /i/ unassociated. As in the plural, the /i/ of the melody associates to the residue whenever possible. In our representations, the morphemic vowels of the plural and diminutive are maintained on different tiers from the stem melody, consistent with their separate morphological function and with phenomena like the spreading of /a/ across consonants in the iambic plural.

Let us begin with the heavily-populated class of quadriliteral plurals, represented by *jundub/janaadib* and *sulṭaan/salaaṭiin*. These plurals are derived from their corresponding singular stems as follows (F_{QT} denotes a quantitative or moraic trochee; F_I an iamb). The prosodic criterion $\Phi\langle W_{min}=F_{QT}, L\rangle$, interpreted of course in the total-function mode, parses out the contents of the first two moras of these forms. Suppose B=*jundub* 'locust'; then B:Φ=*jun*. This must be mapped to F_I. The mapping proceeds as in (43). (Further details of moraic notation are explored below in section 3.4.)

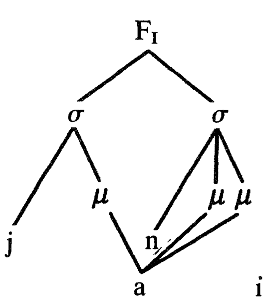
(43)



We show the vowel associated and spread, but it is important to note that there can be no direct evidence for this, since it is overwritten by the *a* of the /a_i/ plural melody. The consonantal association is a straightforward filling of the only obligatorily consonantal positions, the syllable onsets. Spreading to fill the second mora of the heavy syllable is not an option, which accords with recent findings that spreading is not automatic (Pulleyblank 1986). In fact, it is plausible that spreading of consonants to fill empty positions is the default only at the earliest level of the morphology, when verbal derivation and the lexical shapes of nouns are determined. At later strata, as we will see, epenthetic consonants are supplied to fill empty onsets: *w* at the level at which broken plurals are formed, ʔ postlexically.

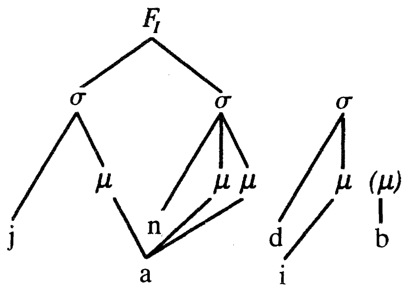
To the form in (43) the templatic melody /a_i/ associates by spreading the /a/, causing the delinking of the stem-vowel; this is portrayed in (44):

(44)



Now the residue B/Φ=*dub* once again becomes accessible and is subject to melodic overwriting by the as-yet unassociated /i/, which displaces the stem vowel, yielding *dib*. The definition of prosodic circumscription in (23), it will be recalled, says O:Φ (B) = O(B:Φ)*B/Φ, where '*' is the relationship holding between B:Φ and B/Φ in B. Since *jun*=B:Φ concatenates to the left of *dub* = B/Φ, the transform of *jun* left-concatenates to *dub*, giving as output *janaadib*, when the final vowel is overwritten:

(45)



Pluralization of *sulṭaan* proceeds in a parallel fashion. The only difference of note follows from the different structure of the final syllable *ṭaan*=B/Φ. When the melody /i/ links, it occupies both vocalic moras (as always in Arabic), yielding *ṭiin*, for a plural *salaaṭiin*.

The corresponding diminutives are derived identically, modulo the differences in the vowel melody and the autosegmental suffix *y*.

Singulars with a medial geminate stand in a striking relationship to the Φ-parse. Consider these examples:

(46)

Sg.	*B:Φ*	*B/Φ*	*Plural*	*Diminutive*	
nuwwaar	nuw	waar	nawaawiir	nuwaywiir	'blossom'
jilbaab	jil	baab	jalaabiib	julaybiib	'type of garment'
tinniin	tin	niin	tanaaniin	tunayniin	'sea monster'

Consider B = *nuwwaar*. Since B:Φ = *nuw* and B/Φ = *waar*, the doubly-associated melodeme *w* will be both inside and outside the domain of template-mapping. From the standpoint of the template-mapping operation O:Φ, which sees only the contents of its prosodically characterized domain B:Φ, the melodic element *w* is subject to the usual melody-to-template mapping, which links it to the onset of the second syllable of the iambic foot. The residue B/Φ, outside the scope of the minimal word, must remain unaffected by the template-mapping operation: it starts out and remains *waar*. One melodic element *w* has a bivalent character imposed on it by the logic of domain characterization. This analysis, therefore, solves the problem that the broken plural presents for geminate integrity (Kenstowicz and Pyle 1973; Steriade 1982), the observation that geminates cannot be split up by rules of epenthesis.

A second interesting effect arises when the W_{min} domain contains insufficient consonantism to fill the two onsets of the iambic foot (recall that spreading is not permitted at this stage of the morphology). Onsets are obligatory, of course, and must therefore be supplied. These are the relevant data:

(47)	*Sg.*	*B:Φ*	*B/Φ*	*Pl.*	*Dim.*	
	xaatam	xaa	tam	xawaatim	xuwaytim	'signet ring'
	jaamus	jaa	muus	jawaamiis	juwaymiis	'buffalo'

Here B:Φ = *xaa*, *jaa* and *O(B:Φ)* = *xawaa*, *jawaa*. Since the 2nd mora of B:Φ is vocalic, the domain contains only a single consonant. By left-to-right mapping, it ends up filling the first onset of the (bisyllabic) template F_I; the second onset cannot be filled from the contents of B:Φ. Consequently, a *w* is inserted to fill the obligatory onset position (McCarthy 1979, 1983).

(48) Consonantal Default Rule

$\emptyset \rightarrow w$, when required by syllabic well-formedness.

Melody-mapping has precedence over the default rule (48).[21]

Perhaps the most interesting effect of Φ-parsing arises in stem forms CvCvvC:

(49)	*Sg.*	*B:Φ*	*B/Φ*	*Pl.*	*Dim.*	
	saħaab-at	saħa	ab	saħaaʔib	suħayyib	'cloud'
	jaziir-at	jazi	ir	jazaaʔir	juzayyir	'island'
	ħaluub-at	ħalu	ub	ħalaaʔib	ħulayyib	'milch-camel'

In all the other forms we have seen so far, B:Φ happens to coincide with a constituent of the base – the first syllable. Here B:Φ = CvCv actually cuts half-way into the second syllable. Furthermore, the residue B/Φ = vC is not itself a syllable in the base. These facts demonstrate clearly that the parse is accomplished on moraic structure, since Φ⟨W_{min}, Left⟩ in Arabic

[21] Independent evidence for the Consonantal Default Rule can be found in the derivatives of sub-minimal bases (like *ʔabawaan*) discussed below in section 3.4, where *w* is inserted to fill-in a position for which no root consonants are available. The default *w* is also found in forms like *ħamraaʔu* 'red (f. nom.)', where it surfaces as *ʔ* by a general phonological rule. When the case-marking desinence *-u* is absent phrase-finally, the default *w* is no longer required to fill the onset and so it too is absent: *ħamraa*.

plurals is a factoring on as-yet stressless representations at the edge opposite that of the stress rule and W_{min} is defined in terms of moras, not syllables. Thus, the Φ-parse respects neither foot structure (there is none) nor syllabic structure (because W_{min} is bimoraic, Φ seeks moras, not syllables) in the form to which it is applied. The Arabic case is thus to be contrasted with Yidin^{y} (28) or Cupeño (30), where Φ takes a pre-existing foot, and with Makassarese (35), where Φ seeking a disyllabic W_{min} takes pre-existing syllables but disregards a stress foot deployed at the opposite end of the word. The treatment of B:Φ here is similar to that of the truncated vocative, discussed below in Section 3.4. The prosodic requirement *Vocative* = $W_{min} + v$ is met in such forms as [*ma.ji*]*i* because the first one-and-a-half syllables are analyzed as $W_{min} = \mu\mu$.

The residue vC can be syllabified by the usual rules that derive syllabic representations from moraic structure, but it can provide no onset for the syllable thus derived. Rule (48) is therefore invoked, giving intermediate forms /*jazaa*wir/ 'pl.' and /*juzay*wir/ 'dim.' Independently motivated rules of glide realization, discussed at length in Brame (1970: 244ff., 273), are responsible for the surface forms *jazaaʔir*, *juzayyir*.

The remaining iambic plurals and diminutives are those from bimoraic singulars, exemplified in the following table:

(50)	*Sg.*	*B:Φ*	*B/Φ*	*Pl.*	*Dim.*	
	nafs	naf	s	nufuus	nufays + at	'soul'
	qidħ	qid	ħ	qidaaħ	qudayħ	'arrow'
	farx	far	x	/faraax/	furayx	'young of bird'
	qadam	qada	m	/qadaam/	qudaym	'foot'

Recall that underlying /faraax/ and /qadaam/ surface as *ʔafraax* and *ʔaqdaam* by C*a* Metathesis. Other than this, derivation is straightforward. The unmarked mode of pluralization for bimoraic singulars is C*a*C*aa*C (see Section 5.2), with the same /a_i/ melody as the longer broken plurals, but with the /i/ lost by Stray Erasure. The diminutives, of course, are unexceptional.

Final consonantal moras are extrasyllabic (see Section 3.4 below). The Φ-parse includes all the metrical contents of the final syllable in CvCvC disyllables like *qadam*. In that way, the distinction between the disyllabic class *qadam* and the monosyllabic class *farx* is entirely neutralized in the diminutive and broken plural – both have just two metrical moras, and so all metrical moras are recruited in the domain of template mapping. Bimoraic bases are mapped exhaustively (up to extrametricality) onto the iambic foot template of the broken plural and diminutive. Disyllabicity of the plural and diminutive in just these cases follows from the lack of intrametrical moras outside the scope of template mapping (that is, the residue only contains an extrasyllabic final consonant in both types). This contrasts with the treatment of *jaziir*, where the stranded mora is vocalic and therefore not licensed outside of a syllable.

Paralleling the "transfer" effect in (46), where a multiply-associated consonant is both inside and outside the domain of template mapping, we find similar effects involving melodic elements that are, in part, linked to extrasyllabic positions. Like *saħaab + at/saħaaʔib* is *jariir + at/jaraaʔir* 'guilt'; with bimoraic singulars we have nouns like *sirr/ʔasraar* 'secret' or *sabab/ʔasbaab* 'rope; reason'. As with *nuwwaar*, the multiply-linked consonant is bivalent, associating to the iambic template within the domain of the prosodically specified base and remaining unaffected in B/Φ outside that base.

We can now sum up. The problem confronting us at the outset was the great superficial diversity in canonical form of the iambic plural and diminutive. The basic assumptions of the theory of Prosodic Morphology dictate the solution. The Template Satisfaction Condition entails that the template can only include the canonical invariants in a prosodically diverse set. The parametric option of prosodic circumscription of domains gives the scope of the canonical invariance. And the Prosodic Morphology Hypothesis requires that the template be construed as an iambic foot.

Yet this solution entails all of the transfer results as well. Transfer of vowel length in the final syllable follows from the fact that it lies outside the domain of melody-to-template mapping. The distribution of the inserted *w* follows from the structure of the templatic domain and its relation to the residue. Preservation of idiosyncratic patterns of consonant association or reduplication follows from the fact that the stem rather than root is the base of pluralization. The treatment of affixes follows in the same fashion. Finally, the distinction between templatic and nontemplatic portions of the plural or diminutive yields a straightforward explanation for the scope of the derived vocalism.

3.4. *The Minimal Word in Arabic*

We have seen that a bimoraic sequence, the moraic or quantitative trochee, is the prosodic constituent subject to template mapping in the iambic plural. The moraic trochee plays an important role in the Arabic system of prosody. Not only is it the basis for stress assignment, but it also conditions a variety of minimal word effects. The prosodic hierarchy asserts that the minimal word is a single metrical foot. Since the stress-foot of Arabic is bimoraic, it follows that the minimal word should also be bimoraic up to extrametricality.

Our investigations begin with the stress system. There is inconsistency 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, the nearly universal norm is represented by the data in (51):

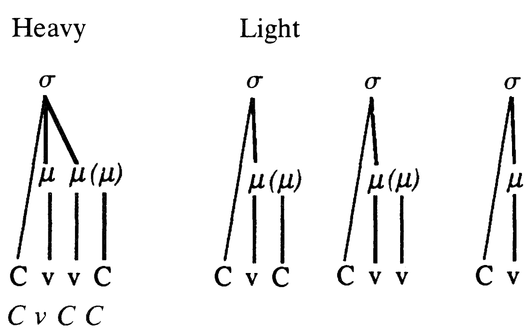
(51) Stress Placement

Final	*Penult*	*Antepenult*
yaquul	yaquulu	kataba
qaanuun	yaqulna	katabat
sirħaan	qaalat	katabuu
ḍarabt	dirham	
tarjamt		

The stress system is obviously weight-sensitive: final syllables are stressed if superheavy CvvC or CvCC; penults are stressed if heavy Cvv or CvC; otherwise the antepenult is stressed.

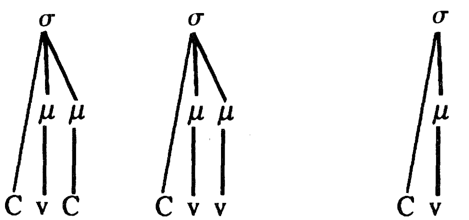
The treatment of syllable-structure in Arabic is as follows. Assume a bimoraic (μμ) representation of heavy syllables and monomoraic (μ) representation of light syllables (Prince 1983, Hyman 1985). The first, or strong, mora can only be a vowel *i*, *a*, or *u*. The second, or weak, mora can be any consonant or the second half of a long vowel. Exactly one consonant must begin any syllable. Superheavy syllables, which are limited to final position, as well as word-final heavy syllables, have a final extraprosodic mora (or syllable (McCarthy and Prince; forthcoming)):

(52) Syllable Weight in Word-final Position



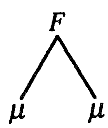
Syllable Weight in Word-medial Position

Heavy Light



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:

(53) Quantitative Trochee



For purposes of stress assignment, final light syllables – whether originally light like *kataba* or light by virtue of extrasyllabicity like *katabat* – are not included in the application of this foot. They are therefore extrametrical, invisible to foot assignment.[22]

By the prosodic hierarchy, in which Word dominates Foot, the minimal base of Arabic must be a single quantitative trochee or, equivalently, two moras. Since final moras are extrasyllabic (and all lexical stems are consonant final), the minimal base must contain the two moras of the foot plus an extrasyllabic one. Therefore the minimal base of Arabic is exactly satisfied by CvCC, CvvC, or CvCvC. We will now proceed to examine this prediction against the facts.

The smallest noun stems that occur with any frequency are CvCC

---

[22] An interesting case is presented by forms like *tárjama* 'he translated', with a heavy antepenult followed by a light syllable (also see Hayes (1987: p. 282)). 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 the foot-assignment must move on to yield $[tar]_{F}ja$.

(appearing 940 times in our lexical material) and CvCvC (appearing 281 times). CvvC stems are independently prohibited at underlying representation, although they do occur at the surface (McCarthy and Prince, forthcoming a). Words smaller than these do occur, but only under very special circumstances. Some candidate words that are too small appear in (54):

(54)	*Non-words*		*Biliterals*		*Imperatives*
	wa	'and'	ʔab	'father'	li (imperative √wly 'be near')
	bi	'in'	bn	'son'	daʕ (imperative √wdʕ 'put down')
	qad	'past'			ktub (imperative √ktb 'write')
	laa	'not'			

All of these forms have at most one mora; in fact, 'son' evidently has no moras at all in underlying representation, since it has no syllables. All of these apparent counterexamples are subject to reanalysis.

Those in the category "non-words" in (54) 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. Cross-linguistically we know that nonlexical vocabulary need not have the prosodic or segmental properties of lexical vocabulary. For example, the only "words" of English beginning with ð are non-lexical: *the*, *this*, *that*, *thou*, *then*, etc. Furthermore, they need not bear a stress – contain a foot – unlike major category words. These non-words are therefore irrelevant to determining the size of the minimal base.

The examples in the second column of (54) are lexical vocabulary items – they are nouns – but they too are not compelling evidence against the bimoraic minimal base. The reason is that they come from a very small, closed class of items that have probably never reflected a productive pattern of the language. In the 1-mora class with *ʔab* are *ħam* 'father-in-law', *ʔax* 'brother', *dam* 'blood', *fam* 'mouth', and *yad* 'hand'. And in the ø-mora class with *bn* are *sm* 'name' and *st* 'anus'.[23] These lists are exhaustive, so the numbers are obviously quite small. Moreover, these words are brought into conformity with the minimality requirement whenever they participate in any of the truly productive morphology of the language. (Itô (1988) has observed that this is a general property of minimality.)

[23] All of the biliterals refer to near kin or body parts which in many languages require a possessive pronoun, as Ellen Woolford has pointed out to us. Since Semitic characteristically indicates pronominal possession by a suffix, at an earlier stage of the language an obligatory possessive suffix may have supplied the additional mora.

ʔab, for instance, is brought to bimoraicity with the help of the default *w* (48) in the dual and nisba (a productive denominal adjective obtained by suffixing *iy*): dual *ʔabawaan* 'two fathers', nisba *ʔabawiy* 'paternal'. It also has this *w*, in its phonologically predictable alternant *ʔ*, in the iambic plural *ʔaabaaʔ+u* (from /ʔabaaw+u/ via /ʔaʔbaaʔ+u/, by *Ca* Metathesis and compensatory lengthening). Finally, it lengthens the case suffix (making the word bimoraic) in the definite singular: *ʔal-ʔabuu* (nom.), *ʔal-ʔabii* (gen.). These observations indicate that the biliterals are in fact exceptional in nearly all respects; it is no surprise, then, that they are exceptions (rather than counterexamples) to the minimality hypothesis.

The examples in the third column of (54) are all imperatives. Traditionally, imperatives are special in two respects, both of which involve a morphological truncation or deletion process. 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. These monomoraic words are not bases, then, but rather are the result of applying later truncating morphology.

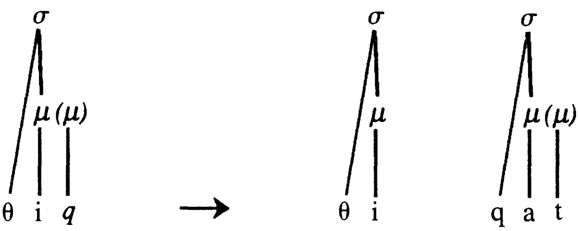
Apart from these basic observations, there are at least four other arguments in support of the bimoraic minimal base in Arabic. First, it is clear that CvC bases like *ʔab* are abnormal even when the root is biconsonantal. Versus the tiny number of words like *ʔab*, our lexical material contains over 150 monosyllables 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 three. 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, there is a morphophonological rule by which many roots whose initial consonant is *w* lose this *w* in the imperfective as well as the masdar (a kind of nominalization), as (55) shows:

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

The problem is why just these masdar forms require the feminine suffix *+at* – the bimoraic minimum accounts for it. 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 (56) shows:

(56)



The traditional idea (Wright 1971: p. 118) that the feminine suffix compensates for the loss of the *w* is expressed formally by the bimoraic minimal word requirement. The addition of a suffix to satisfy minimality may seem odd, but in Arabic this particular suffix has an unusually broad range of morphological functions.[24] In fact, its phonological compensatory function is not even limited to just these masdars. An additional bit of evidence in support of this analysis comes from the denominal adjective (called *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 the default *w*, just as in the case of *ʔab*: perfective *waʕad* 'make a promise', imperfective *ya* + *ʕid*, masdar *ʕid* + *at* 'a promise', nisba of masdar *ʕidaw* + *iy*, 'promissory'.

A third 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 (57):

(57)

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

[24] We are indebted to Michael Kenstowicz for reminding us of the multifunctionality of Arabic *+at*.

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 loanwords 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 in unpublished work Kenstowicz (1981) has argued that vowel-length alternations observed in Lebanese Arabic imperatives like *ktoob* 'write! (m. sg.)', *ktibu* 'write! (pl.)' demonstrate a two-mora minimality requirement with final extrametricality, as we have argued for Classical Arabic.

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 often are based on the minimal word or, equivalently, the foot (McCarthy and Prince 1986, forthcoming b). The apparent truncation in such cases is the result of mapping to a minimal word template.[25] Some examples appear in (58):

(58) a. Yapese (Jensen 1977: 101, 114)

Full Noun	*Vocative*
luˀag	luˀ
bayaad	bay
maŋɛɛfɛɛl	maŋ

b. Central Alaskan Yup'ik Eskimo (Woodbury 1985)

Full Noun	*Proximal Vocative*
Aŋukaɣnaq	Aŋ ~ Aŋuk
Nupiɣak	Nup ~ Nupix ~ Nupik
Cupəɬ:aq	Cup ~ Cupəɬ

[25] Mester (1988) has observed that truncation can also be accounted for by a slight modification of prosodic base specification (in which the residue B/Φ is not restored) as well as by mapping to a template. Mester shows that a variety of truncation phenomena in Japanese require one or the other of these mechanisms as well as both combined.

c. Japanese (Poser 1984)

Name	*Hypocoristic*
midori	mii + tyaN
	mit + tyaN
	mido + tyaN

d. English

Thomas	Tommy
Barbara	Barbie
Algernon	Algie
Benjamin	Benjie
Cynthia	Cindy
Edward	Eddie, *Edwie

In Yapese, the smallest licit independent word is a CvC syllable, and this clearly corresponds to the output of vocative truncation. The Yup'ik case, which is insightfully analyzed in prosodic terms by Woodbury (1985), is a clear example in which the morphology must make reference to an iambic foot. The patterns assumed by proximal vocatives correspond exactly to the complex requirements that the Yup'ik stress system must in any case place on this foot type – it is monosyllabic or disyllabic, it contains at least two moras, it must end in a consonant, and bimoraic syllables are permissible only on the right. Poser (1984) carefully demonstrates that the bases of Japanese hypocoristics with suffixed *tyaN* are composed of one (or two) bimoraic units. Extensive evidence that Japanese has a recurrent bimoraic unit – that is, a foot – appears in Poser (1984), McCarthy and Prince (1986, forthcoming b), Tateishi (1989), and Itô (1988). Finally, the English examples involve a case where the minimal word is coextensive with the syllable (to which the stress-neutral suffix *i*: is added). This discussion is subject to further qualification and interpretation; the truncated words often display idiosyncrasies of mapping or realization of the original segments. But the overall inference can be drawn confidently that hypocoristics or vocatives provide a solid handle on the minimal word.

Arabic has truncated vocatives in classical verse, discussed by Wright (1971:2.88) and Howell (1986:1.1.191-4). Representative data, all proper nouns, appear in (59):

(59)

	Proper Noun	*Poetic Vocative*
a.	CvvCvC nouns	
	maazin	maazi
	maalik	maali
	ʕaamir	ʕaami
	ħaariθ	ħaari
b.	CvCvvC nouns	
	suʕaad	suʕaa
	majiid	majii
	θamuud	θamuu
c.	CvCCvC nouns	
	jaʕfar	jaʕfa
d.	CvCCvvC nouns	
	ʕuθmaan	ʕuθma
	marwaan	marwa
	manṣuur	manṣu
	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.[26]

The most interesting contrast in (59) 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 word is bimoraic, then the truncated vocative is a minimal word followed by a vowel: [*maji*]*i*, [*marw*]*a*. The vowel is not some arbitrary appurtenance to the minimal word 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 ħaariθ* + *u*, *yaa jaʕfar* + *u* (*yaa* is the vocative particle). In fact, the final vowel of the

[26] There is some disagreement in the early sources about the correct treatment of noncanonical noun patterns in the vocative, but in no case is the resulting vocative sub-minimal.

truncated vocative may assume the melody of the nominative case-marking *yaa ħaar + 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. Minus the case suffix, the vocative is a stem, like all stems requiring that any final consonant be extrametrical. We thus have the contrast between the minimal base of the iambic plural or diminutive *[jaʕ]far* and the minimal stem of the vocative *[jaʕf]ar*.

The conclusion is unavoidable, then, that the productive vocabulary of Arabic eschews Cv or CvC bases, minimally requiring CvCC or CvCvC (with moraically equivalent CvvC ruled out by independent considerations). We interpret this as a minimum base size of two moras, which finds independent motivation in the masdars of roots with initial *w*, loan phonology, and truncated vocatives. Moreover, the bimoraic minimal word – a quantitative trochee – is what is predicted by the prosodic hierarchy and the analysis of the Arabic stress system.

3.5 *The Iambic Template*

The invariant canonical shape of the iambic plurals is CvCvv+, and, from the Prosodic Morphology Hypothesis (which holds that templates are specified in prosodic terms) and the Template Satisfaction Condition (which requires that templatic constraints be met obligatorily), it follows that the template of these plurals must be an iambic foot. Recent typological work (Hayes 1985) indicates that the iambic foot is always quantity sensitive. This means that the canonical expansion of an iambic foot is always a sequence of a light syllable followed by a heavy one. From the vantage of universal grammar, the broken plural and diminutive template is sufficiently specified as an iambic foot.

There is good local evidence for this characterization as well. Fleisch (1968: p. 63–67) observes that the iambicity of the broken plural system is paralleled in two other loci in the language. First, he compares the distribution of singular nouns CvvCvC versus CvCvvC. The differences between these are remarkable. Our lexical data yield the following:[27]

[27] The totals given here can be reconciled with those given earlier as follows. Our earlier count of 245 C*aa*C*i*C nouns did not include those with other vocalizations (8) or sound-plural doublets (18). Our earlier count of 447 also did not include 16 sound plural doublets.

(60)

	CvvCvC		CvCvvC	
	C*aa*C*i*C	263	C*a*C*ii*C	265
	C*aa*C*a*C	7	C*i*C*aa*C	106
	C*aa*C*u*C	1	C*a*C*aa*C	37
			C*a*C*uu*C	29
			C*u*C*aa*C	25
			C*i*C*ii*C	1
	Total	271	Total	463

Iambic (CvCvvC) stems are much more common and occur in many more vocalic patterns in a more even distribution than CvvCvC stems. On deeper analysis, the skew turns out to be even worse than this – all C*aa*C*i*C nouns owe their existence to a single morphological process, the formation of the active participle of the Form 1 (underived) verb (e.g., *kaatib* 'writing'). Apart from this single source, there essentially are no CvvCvC stems, while the iambic stems are abundant and diverse.

The explanation for this dramatic skew comes from the way prosody imposes constraints on stem structure. 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 is asymmetrically light-heavy, but the trochaic consists of two equal parts, two moras or two syllables (or perhaps always moras if syllables in quantity-insensitive stress systems are construed as monomoraic). 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 a final extrametrical consonant), but CvvCvC is a bimoraic (trochaic) foot plus something more: $[_F$Cvv]CvC. The desirable equation is then Template = Foot, a clear relative of the minimal word conditions discussed above. More generally, among Arabic nouns, there is a requirement that the stem pattern be exactly measurable in feet, so Template = F^+. (This is simply a special case of the Prosodic Morphology Hypothesis.) Taken with the limitation, noted above, that canonical nouns are not longer than two syllables, this derives the basic nominal stem-forms of the language: one foot – CvCC, CvCvC, CvCvvC; two feet – CvCCvvC, CvvCvvC.[28] The anti-iambic form $_F$[Cvv]CvC fails this requirement and is therefore excluded from the list of nonderived stem types.

[28] See McCarthy and Prince (forthcoming, a) for a detailed account of the noun system in these terms. Note that the canon CvCCvC arises only with quadriliterals (e.g., *jundub*) or templatically prefixed triliterals (e.g., *marħal + at*), with very few exceptions. Since the association pattern in these cases is entirely predictable from the requirement that root consonants must be expressed (melodic conservation), quantity is irrelevant, and the template itself is merely disyllabic.

The morphology shows that CvvCvC is indeed a derived stem type. Since it occurs in the noun system only 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. The finite verbs that are also heavy-light, like Form 3 CvvCvC, are derived as well: 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. Details and justification of these observations are in McCarthy and Prince (forthcoming a).

Fleisch goes on to observe that iambicity plays a role in verse as well. He argues from the statistical work of Vadet (1955), pointing out that the four clearly iambic metres *ṭawiil*, *kaamil*, *waafir*, and *basiiṭ* are used in 93% of a corpus of nearly 2300 classical poems, with eleven other metres dividing up the remaining 7%. To this we can add the structural observation that *all* meters are based on repeating an iambic core – the 'peg' of traditional analysis – which consists of a light syllable followed by a heavy syllable. (See Prince (1989) for some recent discussion.)

The evidence available for stress placement in Arabic indicates trochaic feet, and not iambic ones, though Hayes (1986) finds them in Cyrenaican Bedouin. In different domains of the language, one or both foot types are active. The system of versification is iambic. The major broken plural pattern is iambic as well, but the most important secondary one is trochaic. For the stress system and the minimal word, we also have trochees. And in the characterization of the basic templates of nouns, both trochees and iambs are required (McCarthy and Prince forthcoming, a). For Arabic, it is remarkable confirmation for the Prosodic Morphology Hypothesis that the diversity of phonological, morphological, and poetic phenomena can be shown to depend on just the two quantity-sensitive foot types supplied by phonological theory.

4. Discussion of Earlier Treatments

The first step toward the templatic treatment of Arabic broken plurals and diminutives was the CV template proposed in McCarthy (1979, 1983):

(61) Broken Plural and Diminutive Template

CVCVVCVVC

All of what we have called iambic plurals can be regarded as being constructed on this template. A plural like *salaaṭiin* exhausts the positions in the template, and *jawaamiis* exhausts the positions with the provision of an inserted *w* in the second C position. The plural *janaadib* exploits all of the template except for one V position in the final syllable; the plurals *xawaatim* and *jazaaħir* are formed similarly, but with inserted *w* and ʔ (from underlying /w/) in the second and third C positions respectively. Forms like *nufuus*, *qidaaħ*, *and ʔaħkaam* are formed only on the initial CVCVVC substring of the template.

But these observations about the template in (61) hardly constitute a rule for forming the broken plural – they do not explain why one singular requires a particular substring of the broken plural template and another singular requires a different substring, nor why *xawaatim* and *jazaaʔir* differ in the position where an extra consonant is inserted. These problems are addressed by transderivational constraints in McCarthy (1979) and by positing a rule of infixation in addition to the template in McCarthy (1983). Criticizing the latter analysis, Hammond (1988) proposes that template mapping in Arabic plurals is mediated by an extensive revision of Clements's (1985) mechanism for transfer of segmental quantity and syllabicity in reduplication.

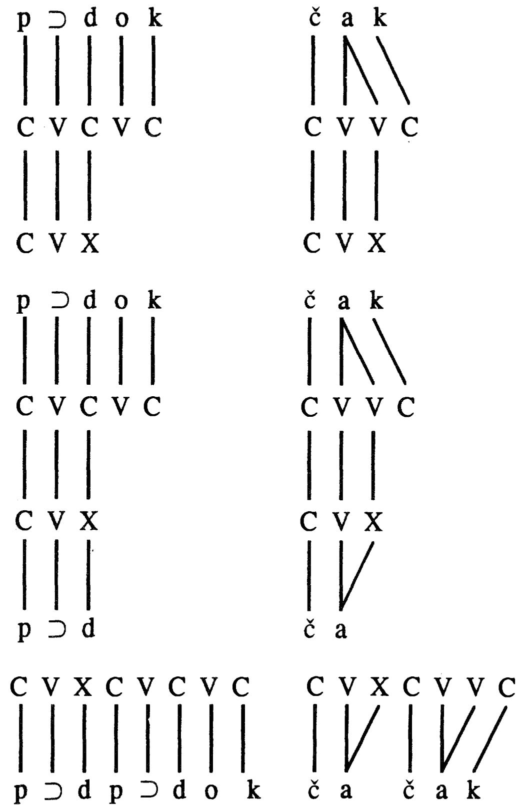
Clements's proposal is an adaptation of Marantz's (1982) idea that reduplication is accomplished by concatenating to a stem an affix composed purely of the skeletal units C and V. In Marantz's account, the affixal skeleton is satisfied by copying the segments of the base (the "phonemic melody") and linking them to the skeletal positions of the affix. In Clements's alternative conception, the reduplicative affix is in fact a suprafix, a skeletal morpheme that is originally parallel to the skeleton of the base rather than concatenated with it. The skeleton of the suprafix is satisfied by first aligning it with the skeleton of the base through association lines, and then by replicating on the parallel skeleton the melodic associations of the base skeleton.

Clements's proposal primarily addresses problems like the following, first recognized by Levin (1983). In Mokilese, prefixing reduplication of a base whose initial syllable is short copies CVC: *pɔd* + *pɔdok*. But prefixing reduplication of a base whose initial syllable is long copies CVV: *čaa* + *čaak*. The difference between long and short vowels is represented purely configurationally: a long vowel is one that is linked to two V positions. If we suppose that the Mokilese affix is CVX (X a variable over C and V), then Clements's theory (originally applied to this example by Levin (1985)) accounts for the transfer of this configurational information from base to affix:

(62) Suprafixation

Transfer

Linearization



In the final step of the derivation, the linear order of suprafix and base is determined, yielding the order of elements actually observed.

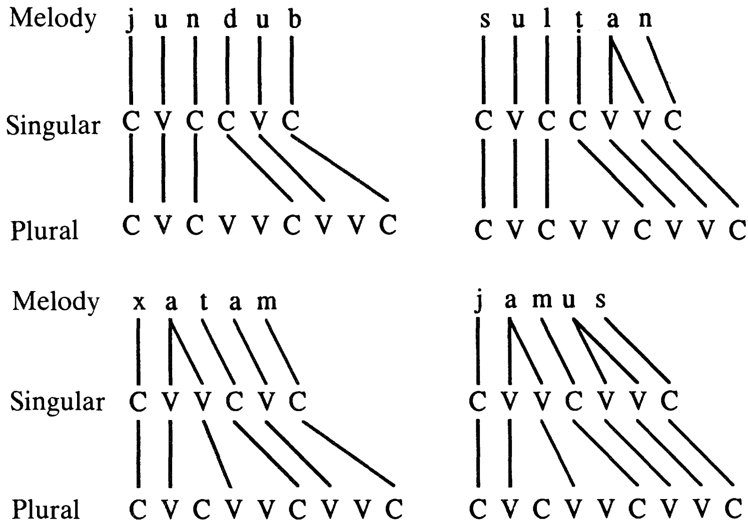
This mechanism is not uncontroversial as an account of transfer phenomena in reduplication (Marantz and McIntyre 1986; McCarthy and Prince 1988), but Hammond argues that, whatever its status in reduplication, it must play a role in accounting for the Arabic broken plural. He initially confines his attention to singular/plural pairings like the following:

(63)

Singular	*Plural*
jundub	janaadib
sulṭaan	salaaṭiin
xaatam	xawaatim
jaamuus	jawaamiis

If the CVCVVCVVC template.proposed in McCarthy (1979, 1983) is suprafixed to the skeleton of the singular and association proceeds left-to-right from the singular skeleton onto the plural skeleton, the following pattern of linking is obtained:

(64)



The fact that the second C position of the plural skeleton in *xawaatim* or *jawaamiis* is empty, to be filled later by an inserted *w*, follows immediately at this point, because association of the singular [CVV... to the plural [CVCV... traps the second C, leaving it unlinked and unlinkable. But the difference in final vowel length between *janaadib* and *xawaatim* on the one hand and *salaaṭiin and jawaamiis* on the other does not. Hammond invokes a special, language-particular rule to account for this distinction, one that shortens the vowel in words like *janaadib* or *xawaatim*:

(65) Vowel Deletion

$$V \rightarrow \emptyset / V\_\_C]$$

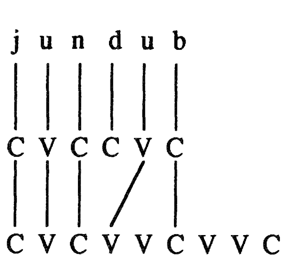
That is, an unlinked vowel in the final syllable only is deleted.

Clearly rule (65) is a major liability of the analysis. Whereas transfer of vowel length follows automatically in reduplication, it requires the intervention of this rule, which essentially stipulates the transfer effect, in the templatic morphology of Arabic. In fact, the grammar would be simpler if vowel length were not transferred at all, so that all broken plurals, regardless of their singulars, invariably had long vowels in the final syllable. This presumably simpler pattern is not met with in any of the Arabic dialects nor in any Semitic language that has retained broken plural formation. Worse yet, rule (65) cannot be eliminated by any straightforward appeal to a more general principle of erasure of unlinked skeletal elements along the lines of Marantz (1982). This sort of generalization immediately runs up against the persistence of the unlinked C's of *xawaatim* or *jawaamiis* or the unlinked medial VV sequences of all the broken plurals in (64).

This problem is of profound importance, since it points to a fundamental failing of the CV-skeletal theory in which this analysis is embedded. Absent a principle like the Template Satisfaction Condition and the more impoverished (moraic) templates that it requires, the CV approach is unable to make a principled distinction between obligatory and optional skeletal elements and is driven to stipulations like (65). An analysis constrained by TSC, like the one proposed here, necessarily restricts the scope of the template to the true canonical invariant, the iambic foot.

It is also worth noting that the mechanism of templatic transfer is quite remote from Clements's conception of reduplicative transfer. The stipulative character of vowel length transfer in the templatic analysis is one symptom of this. Furthermore, while linearization is crucial to reduplication, since reduplicative prefixes and suffixes are what is actually observed, linearization is impossible in the broken plural case. When forming the plural, all traces of the singular must be erased after they have done their work of supplying the phonemic melody to the plural template. Plural forms like **jundubjanaadib* or **janaadibjundub*, where linearization of the base and suprafix has proceeded normally, are quite impossible. Another indication that there is no suprafix comes from the observation of Clements (see also Mester (1986)) that overapplication[29] of phonological rules in reduplicated forms can be accounted for by applying the rule to the shared structure before linearization. Overapplication is never met with in the Arabic case. Finally, association in the Arabic case is crucially from the singular skeleton onto the plural skeleton, driven by the linear sequence of skeletal slots in the singular. In Clements's theory, on the other hand, association between the two skeleta is directional, but first by vowels and then by consonants. This mode of association, which is necessary to account for properties of reduplication in the Clements theory, produces the following result in Arabic: .

(66)



[29] A phonological rule is said to overapply in a reduplicated form when it applies in both original and copy even though its structural description is apparently met in only one.

Here the V-driven association procedure predicts loss of the unlinkable base consonant *d*, which is not only factually wrong but impossible in the general context of Arabic templatic morphology, where the loss of root melody elements is not tolerated. Taken together, these considerations show that the extension of Clements's theory of reduplicative transfer to Arabic broken plurals relies primarily on graphic rather than substantive resemblances. Thus, the Arabic case must stand or fall on its own merits, without regard to any evidence that comes from reduplication.

Hammond claims three other results for his theory of the Arabic broken plural. First, consonant spreading will be transferred from singular to plural, as in the examples *jilbaab*, pl. *jalaabiib* or *nuwwaar*, pl. *nawaawiir* cited earlier. This is certainly the case, but it should be noted that transfer of vowel spreading (that is, long vowels) and transfer of consonant spreading involve different mechanisms in this account. The latter follows directly from the theory; the former requires the intermediation of a special deletion rule.

Second, the transfer account shares with the earlier analyses of McCarthy (1979, 1983) an explanation for the behavior of certain trisyllabic (therefore noncanonical) singulars in broken plural formation. Examples adduced in Hammond (1988) are the following:

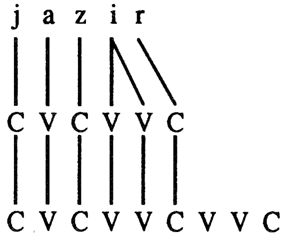
(67)	*Singular*	*Plural*	
	jaħmariš	jaħaamir	'lazy old woman'
	safarjal	safaarij	'quince'
	namuuðaj	namaaðij	'model'

The first two examples are quinqueliteral; they have more consonants than there are slots available in the broken plural template. Any left-to-right template mapping mechanism requires that the last consonant be lost in the plural. The final example has only four consonants, but is impossible as a normal (productive, native) singular noun of the language. It too will align with the plural template in the correct way.

The problem of these noncanonical singulars is examined in detail below (Section 5.1); it emerges that they in no way reflect a regular grammatical process of the language. For now it is enough to observe that broken plural treatment of noncanonical singular nouns is the exception rather than the rule; at all historical stages of the language noncanonical singulars lawfully form only feminine sound plurals.

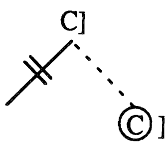
The third argument put forth in support of the transfer analysis comes from another broken plural type, the one applied to nouns like *jaziir+at* to form *jazaaʔir*. These nouns, with short initial and long final syllables, have an inserted /w/ (surface ʔ) in the third C slot of the template. The immediate result of the transfer procedure, however, fills that slot with the final consonant of the root:

(68)



Thus, an additional rule is required to move the association line of the last C of the singular skeleton to the last C of the plural skeleton:

(69) Consonant Reassociation



The circled C in the context of the rule must be unassociated; the rule itself performs two simultaneous transformations, erasing one association line and inserting another.

Again, it is clear that the natural outcome of the analysis is incorrect, and an additional rule without independent support must be stipulated.[30] The grammar of the language would be simpler if rule (69) did not exist at all, presumably yielding **jazaarii* (from */jazaariw/). And again, it should be noted that no Arabic dialect nor any other Semitic language that has retained the broken plural exhibits this simpler grammar.

Of course, one might turn our own question around and ask how the prosodic analysis could be changed in small ways and whether the result is a possible grammar. Clearly the template could be different, but that alone will not distinguish the theories. In fact, it is far more likely that a CVCVVCVVC template would vary from language to language than the iambic template, since the iambic foot is not an arbitrary concatenation of C's and V's but rather one of a small number of specific prosodic

[30] Hammond (1988: 15n.), pursuing a suggestion by Michael Kenstowicz, proposes that rule (69) is independently required to account for the pattern of medial gemination in verb forms like *kattab* 'caused to write'. Aside from technical problems of formulating the unified process, there is little reason to suppose that a rule like (69) is involved in medial gemination. One alternative is to adopt Yip's (1988) edge-in association, which is quite generally compatible with the prosodic account of the broken plural so long as empty onsets take priority. Another view, pursued in McCarthy and Prince (forthcoming, a), generalizes the plural gemination of (2d) (e.g., *saamir*, pl. *summar*, by leftward spreading to a mora) to the verbal cases.

categories and, furthermore, iambicity has deep roots in the language (see Section 3.5). What if the grammar had the iambic template but lacked any prosodic specification of the base? For independent reasons, that is simply not an option. Ordinary morphology is always melody-conserving (McCarthy and Prince 1986, forthcoming b; Yip 1988) – for example, we do not find quadriliteral roots being squeezed into triliteral templates by a mechanism like Stray Erasure. Thus, an iambic template could never apply to quadriliteral nouns without prior prosodic restriction of the base. What if the grammar took the Cebuano or Dyirbal option of limiting the morphological operation to bases that are exhausted (up to extrametricality) by the prosodic constituent? That may be the situation in Biblical Hebrew. In Hebrew, all nouns take the sound plural suffixes, but CVCC nouns additionally have broken plural morphology, so we find *melex* /malk/, pl. *məlāxîm* /malak + iim/. The options afforded by our analysis are therefore excluded by independent principles or actually attested.

Finally, it is important to note that the empirical coverage of the transfer analysis is much less than that of previous accounts like McCarthy (1979, 1983). Many broken plurals were previously analyzed as being built on the CVCVVCVVC template, but they are not accounted for in the transfer treatment. These are the plurals of unsuffixed nouns CVC(V)C (567 in our sample), examples of which appear in (4a, b). In the transfer analysis, there is no natural characterization of this phenomenon. Applying the principles already developed to a singular like *nafs* (pl. *nufuus*) yields the structure in (70):

(70)
```
n a f s
C V C C
C V C V V C V V C
```

Applying Vowel Deletion (65) and Consonant Reassociation (69) produces the following result, which represents the impossible form **nafaaʔis*:

(71)
```
n a f s
C V C C
C V C V V C V C
```

It is clear that several additional ad hoc rules would be needed to obtain the desired CVCVVC canonical pattern of the plural *nufuus*. And again, these additional rules find no support in the cognates in Arabic dialects or other Semitic languages.

Hammond argues that this lack of empirical coverage is not a significant liability of his analysis. He writes:

> First, unlike the cases presented in the text [i.e., the trisyllabic plurals – JJM/ASP], one cannot predict which of the three patterns here a noun in CVCC or CVCVC will assume in the plural. Second, the plural forms here bear no obvious formal similarity to the patterns in the text, e.g. the vocalisms are different and the number of consonants and vowels are different. (Hammond 1988: 267 fn. 16)

These are observations rather than arguments. Linguistic regularities are not based on obvious formal similarities, but on deeper structural principles. The lack of predictability (considerably overstated here – see Section 5.2) reveals nothing except that the vowel melodies are more complex in this case; yet it is the skeleton alone that is at issue. The "three patterns" of disyllabic plurals are canonically identical, modulo the independently motivated rule of C*a* Metathesis. Moreover, the canon of the disyllabic plurals, even in CV templatic theories, is a substring of the canon of the trisyllabic plurals, yet the transfer analysis is unable to capture this important generalization. The differences in numbers of consonants and vowels alluded to in the quotation are simply part of this generalization that must be accounted for.

More importantly, as is shown here and in McCarthy (1979, 1983), Arabic provides us with persuasive evidence (see Section 2.3) that broken plurals like *nufuus* and *salaaṭiin*, despite their supposed lack of obvious formal similarity, are in fact constructed by exactly the same rule. The treatment of loanwords and plurals of plurals, which generalize the iambic pattern of *nufuus* and *salaaṭiin* in the same way, are two sources of evidence. Even more compellingly, diminutive formation, which is entirely regular and productive, demands a unified account of diminutives like *nufays+at* and *sulayṭiin*, which itself entails a unified account of the canonically identical broken plurals *nufuus* and *salaaṭiin*. Indeed, diminutive formation exhibits all of the transfer effects that are adduced in support of Hammond's analysis of the plural. The transfer theory forces an arbitrary, empirically unmotivated distinction between the disyllabic and trisyllabic forms, in both plural and diminutive. This alone is sufficient to disconfirm it.

This analytical failure follows directly from the same intrinsic shortcoming of CV- and X-based theories that leads to the necessity of stipulating vowel-length transfer with rule (65). The CV- or X-based approach cannot be extended to the disyllabic broken plurals like *nufuus* because it is not informed by a theory of obligatory templatic elements like the Template Satisfaction Condition. The CV skeletal approach is forced to relate CVCVVC to CVCVVCVVC by language-particular rules erasing selected unassociated elements, a nearly hopeless undertaking. The TSC, combined with the Prosodic Morphology Hypothesis, forces a very different relation, via prosodic circumscription of domains – the iambic foot, the only canonical invariant consistent with the Prosodic Morphology Hypothesis, is the constant that unifies all of these plural and diminutive types. Moreover, the prosodic analysis, grounded in a restrictive theory, involves only properties that are themselves independently motivated in the grammar of Arabic (the iambic foot, the minimal word, the characterization of syllable weight and extrametricality) or that appear in similar forms in other languages (prosodic specification of the base of a morphological process). The failures of Hammond's analysis are intractable failures of principle, straightforward consequences of attempting templatic morphology on segmental representations; they highlight the analytical junctures where the Prosodic Morphology Hypothesis leads to successful generalization and new understanding.

5. Ancillary Issues

5.1. *Noncanonical Nouns*

A difference between our account and the earlier templatic analyses of McCarthy (1979, 1983) and Hammond (1988) lies in the treatment of noncanonical singular nouns. Arabic places strong restrictions on the shapes that its singular nouns can assume. Noun stems, like all bases, are minimally bimoraic, as we have already argued. No noun stem contains more than two syllables, and every disyllabic noun stem must begin and end in exactly one consonant (peripheral vowels and consonant clusters are prohibited except in monosyllabic nouns, which require CvCC). There are other restrictions which we will not discuss here (v. McCarthy and Prince (forthcoming a)). Singular nouns that deviate from these requirements we will call noncanonical.

Noncanonical nouns have a number of salient characteristics. First, they are themselves never created by any root-based templatic morphology. Second, they do not normally contribute their roots to further derivational processes – for instance, denominal verbs are almost never created from noncanonical nouns.[31] Third, they are always loanwords, and in fact many

[31] Our lexical survey has revealed that the example *magnaṭiis* 'magnet', *magnaṭ* 'to magnetize' cited in McCarthy (1979) is unique.

can be identified synchronically as loans independently of structural noncanonicity because they violate the native restrictions on consonant or vowel cooccurrence. Fourth, with rare exceptions noncanonical loans do not participate in broken plural or diminutive morphology.

For example, in Wehr (1971) we find significant numbers of noncanonical loans like the following (all from various European languages), none of which form broken plurals; they assume the feminine sound plural instead:

(72)	banṭaluun	'pantaloon'
	tarabeeza	'table'
	turumbeeṭ	'trumpet'
	tiligraaf	'telegraph'
	tilifuun	'telephone'
	bansiyuun	'pension'
	fitamiin	'vitamin'
	funugraaf	'phonograph'
	kiluusikl	'kilocycle'
	kurantiin + at	'quarantine'

It is the noncanonicity of these words, not their status as loans, that prevents them from forming broken plurals. We know of three arguments for this conclusion. First, canonical loans readily – in fact, almost obligatorily – form broken plurals (see Section 2.3). Indeed, Smeaton's (1973: p. 83ff.) study of borrowings reveals that the loans which fail to form broken plurals are just those that are noncanonical, often by virtue of having initial clusters: *bruš* 'brush', *draywal* 'drywall', *dram* 'drum', *fyuuz* 'fuse', and *swič* 'switch'. Second, ancient loans, synchronically identifiable as such solely by their noncanonicity, also resist broken plural formation. Witness the following examples from Wright (1971: p. 198), all of which take the feminine sound plural:

(73)	suraadiiq	'canopy'
	biimaaristaan	'hospital'
	šaadurwaan	'fountain'
	ʔagaa	'Agha'
	baašaa	'Pasha'

Third, native noncanonical nouns, although they have very limited distribution, also systematically fail to form broken plurals. The names of the letters of the alphabet are one type we have already mentioned. Another source of native noncanonical nouns is the historical reanalysis of the results of the rule of Identical Consonant Metathesis (Brame 1970, McCarthy 1981, 1986). Details aside, this rule permutes C_iVC_iV sequences to VC_iC_iV, as in /maħlal + un/ → *maħall + un* 'place + nom. indef.'. Many nouns of this type[32] have plural doublets in Wehr (1971), one broken (/maħaalil + un/ → *maħaall + un*) and the other sound (*maħall + aat + un*). The innovating sound plural doublet makes sense only if the noncanonical output of Identical Consonant Metathesis (noncanonical because the stem is disyllabic yet ends in a cluster) is being taken as the input to plural formation. Noncanonicity blocks the broken plural of input /maħall/, and so the sound plural steps in as the default.

Nevertheless, there is a very small number of noncanonical loans that do in fact form broken plurals. Hammond cites three examples that work as predicted in his analysis as well as other CV treatments (McCarthy 1979, 1983):

(74)	*Singular*	*Plural*	
	safarjal	safaarij	'quince'
	jaħmariš	jaħaamir	'lazy old woman'
	namuuðaj	namaaðij	'model'

The prediction that this analysis makes is quite clear: because of left-to-right association, the final consonant should be lost; and because of the mechanism for transfer of vowel length, a long final syllable in the plural is possible if and only if the singular has a long vowel between its third and fourth root consonants.

The lexical material we have collected supplies a total of 13 noncanonical singular nouns out of a sample of altogether 2483 lexical entries. Of these 13, only 1 provably works in the expected way:[33]

(75)	*Singular*	*Plural*	
a.	*Expected*		
	sulaħfaw + at	salaaħif	'turtle'

[32] Other examples in our lexical material include: *misall+at* 'large needle, obelisk'; *misann* 'whetstone'; *mašaqq + at* 'trouble, toil'; *maṣ abb* 'outlet, drain'; *maḍarr + at* 'harm'; *miẓall + at* 'umbrella'.

[33] A less systematic search through the rest of the dictionary produces ten more examples, of which just three work as expected. The data are: *kardinaal*, pl. *karaadil + at* 'cardinal'; *karaxaan + at*, pl. sound or *karaaxiin* 'workshop'; *faramaan*, pl. sound or *faraamiin* 'firman'; *namuuðaaj*, pl. sound and *namaaðij* 'model'; *kustubaan*, pl. *kasaatibiin* 'thimble'; *manjaniiq*, pl. sound and *majaaniq* 'mangonel'; *firdaws*, pl. *faraadiis* 'paradise'; *firʕawn*, pl. *faraaʕin + at* 'Pharaoh'; *faylasuuf*, pl. *falaasif + at* 'philosopher'.

b. *Wrong Consonant Lost*

zanbarak	zanaabik	'(metal) spring'
birðawn	baraaðin	'workhorse'
barnaamaj	baraamij	'program'

c. *Wrong Consonant Lost and Wrong Vowel Length*

xinnawṣ	xanaaniiṣ	'piglet'
ħirðawn	ħaraaðiin	'lizard'
sinnawr	sanaaniir	'cat'
šintiyaan	šanaatiin	'loose trousers'
jirðawn	jaraaðiin	'large rat'
burnayṭ + at	baraaniiṭ	'(European) hat'
ʔusṭuwaan + at	ʔaṣaaṭiin	'celebrity'

d. *Indeterminate*[34]

barahman	baraahim + at	'Brahmin'

These data indicate that the predictions of the analysis are not borne out in the modern literary language recorded by Wehr. Application of broken plural morphology to noncanonical singulars is abnormal in itself, and neither the treatment of excess consonants nor of vowel length are as predicted. Does the analysis fare any better in the classical language?

The classical grammarians and lexicographers supply a few more examples of noncanonical singulars that form broken plurals, but more importantly they provide a detailed discussion of this phenomenon. Howell (1986: 1.3.935ff. and 1.3.1168ff.) summarizes the testimony of a large number of grammarians; the situation was clearly very confused. According to some, the formation of broken plurals or diminutives from noncanonical (quinqueliteral) nouns is simply impossible. Others report unusual formations like pl. *safaarijal-safaarijil*[35] and dim. *sufayrijal-sufayrijil* or even *sufayrijl*. Others record the existence of forms like those in (74), but with significant complications. Ibn 'Aqîl's treatment (Dieterici 1852) is typical in this respect. He observes that the final consonant can generally be lost, as in *safaarij*, but if the penult consonant is "servile" or homorganic with a servile consonant, it may be lost instead:

---

[34] It is impossible to tell whether or not the vowel length of this form is as predicted or not. Plurals of quadriliteral nouns referring to humans normally take the feminine suffix and shorten the stem-final vowel if it is long.

[35] Wehr (1971) also contains an example of this type: *kustubaan*, pl. *kasaatibiin* 'thimble'. In these rare forms the regular process of iambic plural formation has been applied regardless of the noncanonicality of the input.

(76)

Singular	Plural 1	Plural 2	
xadarnaq	xadaariq	xadaarin	'spider'
farazdaq	faraaziq	faraazid	'lump of dough'

Moreover, if any consonant anywhere in the root is servile, it may be lost, as in *barnaamaj* 'program', pl. *baraamij* or *ʔistabraq* 'thick gold brocade', dim. *ʔubayriq*. The servile consonants are those that occur in affixes which attach to template positions (*m*, *t*, *n*, *st*) and the glides *w* and *y*, which are often phonologically unstable. It is not necessary that the servile consonant actually be an instance of a bound morpheme – mere resemblance is enough. For example, *barnaamaj* looks as if it might be derived from Form Q3 of the verb with infixed *n*, and *ʔistabraq* looks like a Form 10 form of the verb with prefixed *st*.[36] These superficially correct but factually wrong morphological analyses are enough to cause the apparently affixal consonants to be lost in favor of preserving the obviously more salient root material intact – in other words, the "roots" /brmj/ and /ʔbrq/ are back-formations or folk etymologies. Therefore the choice of which consonant to drop is made on analogic rather than grammatical grounds. The loss of high glides in broken plural formation can be accounted for by a somewhat different analogy: surface and underlying high glides are in a very opaque relationship to one another, with many underlying ones realized as zero on the surface and with underlying zero sometimes realized as a surface glide.[37]

The conclusion that emerges from this is that the adaptation of noncanonical nouns by loss of consonants in the broken plural is strongly governed by non-grammatical, analogic factors. The loss of consonants is not a response to template satisfaction, which predicts loss of a peripheral consonant only, but rather is a result of enforcing a separate requirement that roots have at most four consonants. The actual practice of obtaining a quadriconsonantal root seems to be largely a matter of analogy.

In this view, the noncanonical singular, in the rare event that it is subject to broken plural formation, does so by essentially analogic means. The

[36] The license for an excess consonant to delete on grounds of homorganicity with a servile consonant presumably depends on the possibility of assimilation. For example, the *d* of *farazdaq* could be analogically treated as the infix *t*, regularly assimilated in voicing to the preceding consonant (cf. *zdaraʕ* 'to sow' from root /zrʕ/ in Form 8 of the verb).

The "forms" of the verb are derivational classes with constant canonical pattern. They are often referred to by a traditional Western numbering system.

[37] Relevant to this is Anderson's (1981: p. 533) observation that spelling pronunciations, another sort of partly metalinguistic activity, always stay within the domain of the existing phonological processes of the language.

root is stripped of apparent nonessentials and treated as if it conformed to some native quadriliteral model. Because this mechanism is outside the formal grammar, it is consistent with our observations about the operation of grammatically false analogies in determining what the root is. It also makes a prediction that the purely formal analyses cannot: a native plural like either *janaadib* or *salaaṭiin* might serve as the basis for the analogy, and so the vowel length in the final syllable of plurals and diminutives from noncanonical nouns should be arbitrary or inconsistent rather than grammatically determined.

This prediction is correct. The lexical data in (75) bear it out, as does the witness of the native grammatical tradition. According to Ibn 'Aqîl and Wright, in both the diminutive and the broken plural of noncanonical nouns the vowel length of the final syllable is essentially arbitrary: for example, safarjal has variant plurals *safaarij* ~ *safaariij* and variant diminutives *sufayrij* ~ *sufayriij*.

Let us now sum up. Formation of broken plurals from noncanonical nouns is itself an abnormal process – such nouns ordinarily form sound plurals. When it does occur, it exhibits conspicuous effects of analogy in determining which consonants to retain and it shows lack of grammatical specification of vowel length in the output. We conclude that this process is entirely analogic and therefore irrelevant to establishing the correctness of a grammatical description. This is hardly surprising: borrowed words, especially those that mark themselves formally as outside the system, are frequently subject to analogy.

5.2 *Melody Selection in CvC(v)C Nouns*

As we have already observed, nouns with singulars CvC(v)C have three different vocalizations imposed on the iambic template, exemplified by the forms in (4a, b). That is, it is just the minimal (bimoraic) singular nouns that form iambic plurals with diverse vocalizations, while longer nouns form iambic plurals exclusively with /a_i/. This sort of complication is exactly what the theory leads us to expect: the phonological and morphological distinctness of template and vowel melody means that they can cover somewhat different domains of the lexicon; and the definition of the partial function Φ' in (31) will split that coverage along the line between minimal and larger bases.

We first examine the diversity of CvC(v)C vocalizations statistically with the following table, based on the lexical data we have collected. A rare, trochaic plural pattern *ʔa*CC*u*C (/CaCuC/) sometimes attested for this class is included for comparison:

(77)

			Plurals		
Singulars	CuCuuC	CiCaaC	/CaCaaC/	/CaCuC/	*Sum*
CaCC	**180**	46	97	21	344
CiCC	43	8	**92**	4	147
CuCC	22	12	**76**	2	112
CvCvC	17	21	**126**	6	170
Sum	**262**	87	**391**	33	773

It is clear that the major competition is between C*u*C*uu*C and /CaCaaC/; the other two patterns are relatively insignificant. (In fact, 24 of the 33 /CaCuC/ plurals are doublets of one of the iambic ones.) Moreover, the clear winner in all categories except for singular C*a*CC is the plural /Ca CaaC/. Even the disyllabic stem CvCvC strongly favors the low-voweled plural, this despite the fact that nearly all the disyllabic stems are vocalized C*a*C*a*C.

More detailed examination of the data reinforces these conclusions, since it shows that many of the apparently exceptional patterns of plural vocalization reflect other subregularities that override the main generalization. Levy (1971: p. 36) observes a kind of dissimilatory effect whereby C*a*CC nouns with middle radical *w* resist the expected C*u*C*uu*C plural – for instance, *lawn* 'kind', pl. *ʔalwaan*, **luwuun*. She also notes that the C*a*C*u*C plural pattern seems to be especially frequent with those rare nouns that are grammatically but not morphologically or semantically feminine. For example, *nafs* 'soul' takes feminine agreement and forms the plural *ʔanfus* as an option to *nufuus*. In addition (Levy 1971: p. 38), the relatively small number of human nouns C*a*CC tend to favor /CaCaaC/ over C*u*C*uu*C: *rabb* 'lord', pl. *ʔarbaab*. Finally, Levy points out that the unexpected C*u*C*uu*C plural for C*i*CC and C*u*CC is a doublet rather than sole plural for two-thirds of the nouns in her comprehensive sample.

When none of these special conditions obtain, however, in the overwhelming majority of cases the regularity is that C*a*CC singulars form C*u*C*uu*C plurals, while all other bimoraic singulars form /CaCaaC/ plurals. Both of these generalizations have connections to phenomena elsewhere in the language. The latter is just exactly the vocalism we find in the trisyllabic plurals like *janaadib*, minus the syllable that lies outside the iambic template. In other words, the quite general vocalism for iambicizing plurals is /a_i/, with /a/ spread onto the template portion and /i/ on any syllable outside the template and otherwise lost by Stray Erasure. This generalization is exceptionless for the trisyllabic plurals, and holds in a majority of cases of disyllabic plurals. A lexical rule of melody selection providing C*a*CC singulars with the /u/ plural melody accounts for most of the rest. These central rules are overlain by various phonologically or grammatically determined subregularities of the sort noted by Levy.

5.3 *The Trochaic Plural*

The only major non-iambic mode of plural formation is the trochaic foot CvCvC, and it appears in quite diverse circumstances. The following data are the result of a purely formal taxonomy of our lexical material; the semantic classifications, where given, are due to Levy (1971):

(78) CVCVC Plurals

Singular	*Plural*	*Frequency*	*Out of*
CiCC + at	CiCaC	138	184
CuCC + at	CuCaC		
CvC(v)C	/CaCuC/	33	897
CvCvvC			
nonhuman	/CaCiC/ + at	88	445
nonhuman	CuCuC	68	
human	CuCaC + aaʔ	69	
human & weak	/CaCiC/ + aaʔ	34	
	total:	259	
CaaCiC	CaCaC + at	26	245
	CuCaC + at	28	

The plural forms given in underlying representation (/. . ./) undergo the rule of C*a* Metathesis (Section 5.4). "Weak" nouns are those whose third radical is a high glide or so-called geminate nouns, with biliteral roots.

These data are obviously orthogonal to our major conclusion about the iambic plural and diminutive; they are included primarily for completeness. Nevertheless, we can observe that the trochaic plural shares with the iambic plural, and indeed with templatic morphology in general, an indifference to the canonical form of its input. Trochaic plurals cross-classify various shapes of singulars, in some cases with significant frequency. From this it can be concluded that the trochaic plural is templatic – there is a trochaic template to which the melody of the singular is applied.

Some evidence for the nature of this template mapping comes from the common trochaic plural of feminine singulars C*i*CC + *at* and C*u*CC + *at*, C*i*C*a*C and C*u*C*a*C respectively. This perseveration of the singular vowel into the plural makes sense if both consonantism and vocalism of the singular are mapped onto the trochaic template in the plural. Since the second syllable of the plural is not supplied with a vowel by template mapping, it receives instead the vowel /a/. Other instances of the trochaic plural receive morphologically-specified melodies by melodic overwriting.

The only other case where the trochaic plural is in the majority is with singular nouns CvCvvC. The apparent diversity of formation mostly yields to the finer classification imposed by Levy (1971). Human nouns take the suffix *-aaʔ*. Their vocalization is normally /u_a/, but they revert to the /a_i/ vocalization of the nonhuman class when the root is biliteral or ends in a high glide. The nouns in this nonhuman class assume the suffix *-at* in the plural. The only remaining nondeterminism is in the choice of plural vocalization and suffixation for the nonhuman nouns; roughly equal numbers of both types are represented.

5.4 *Ca Metathesis*

Levy (1971) is responsible for the observation that surface ʔ*a*CC*aa*C plurals can be derived from underlying C*a*C*aa*C by a rule we have called *Ca* Metathesis. This point is obviously important to the extension of the iambicizing plural to the bimoraic singulars, in parallel with diminutive formation.

Ca Metathesis turns out to play a wide-ranging role in the nominal system; it is in no way restricted to just this one plural pattern. First, it occurs with the trochaic plural pattern, CvCvC: /CaCuC/ → ʔ*a*CC*u*C; /CaCiC + at/ → ʔ*a*CC*i*C + *at*; /CaCiC + aaʔ/ → ʔ*a*CC*i*C + *aa*ʔ. Second, it applies to the productive elative adjective: /kabar/ → *ʔakbar* 'greater; greatest'. Third, it applies to so-called "verbs of surprise": /kaðab/ 'lie' → *ʔakðab* in *maa ʔakðabahu* 'what a liar he is!'. Although these are called verbs, they appear to have the properties of nouns, completely eschewing normal verbal inflection. Fourth, *Ca* Metathesis derives the cardinal number *ʔarbaʕ* 'four' from /rabaʕ/; the root is /rbʕ/ on the evidence of regular formations like *raabiʕ* 'fourth', *murabbaʕ* 'fourfold', and *rubaaʕiy* 'quadriliteral'.

Independent support for the *Ca* Metathesis rule comes from a minor variation on the /C*a*C*aa*C/, ʔ*a*CC*aa*C plural pattern observed by Levy (1971: pp. 90–93, 259). Certain nouns and adjectives with stems CVCC followed by the masculine suffix +*aan* or the feminine suffixes +*aa*ʔ and +/ay/ take /CaCaaC/ plurals plus +/ay/:

(79)	*Singular*	*Plural*	
	sakr + aan	sakaar + /ay/	'drunk'
	ʕaðr + aaʔ	ʕaðaar + /ay/	'virgin'
	waħm + /ay/	waħaam + /ay/	'pica (of pregnant woman)'

These cases exceptionally retain the underlying C*a*C*aa*C pattern.

Although the precise conditions on this rule are not wonderfully transparent, it appears that it is fairly generally applicable to derived disyllables in initial *Ca*, transposing the consonant and vowel and inserting ʔ in the familiar onset-filling fashion. Although ʔ insertion is the normal postlexical mode of supplying an onset in Arabic, the ʔ derived by *Ca* Metathesis is demonstrably different from this, since it appears even when the preceding word ends in a consonant.

6. Conclusion

In this article we have given an account of the productive, general patterns of broken plural and diminutive formation in Arabic. We have shown that these phenomena rely fundamentally on the prosodic circumscription of the morphological base, a notion which we formalized and explored through a diversity of manifestations. We have related this to detailed evidence for the minimal word in Arabic, and we have shown how, combined with an iambic template, the prosodically characterized base yields exactly the distribution of invariants and dependencies that the language actively exploits.

Our proposal has been developed in terms of the theory of Prosodic Morphology, relying on the fundamental tenet that templatic or reduplicative morphology must refer only to the units of the prosodic hierarchy. We have seen that CV theory (or in fact any similar revision of it), even aided by the mechanism of transfer, is unable to express these same generalizations. Our conclusions bear not only on the parochial issues of the broken plural in Arabic but also on the broader topic of the relation between phonology and morphology.

References

Archangeli, D.: 1983, 'The Root CV-Template as a Property of the Affix: Evidence from Yawelmani', *Natural Language and Linguistic Theory* **1**, 348–84.

Anderson, S.: 1981, 'Why Phonology Isn't Natural', *Linguistic Inquiry* **12**, 493–540.

Aronoff, M.: 1976, *Word Formation in Generative Grammar*, MIT Press, Cambridge, Massachusetts.

Aronoff, M.: 1988, 'Head Operations and Strata in Reduplication: a Linear Treatment', *Yearbook of Morphology* **1**, 1-15.

Aronoff, M., A. Arsyad, H. Basri, and E. Broselow: 1987, 'Tier Configuration in Makassarese Reduplication,' in A. Bosch, B. Need, and E. Schiller (eds.), *CLS* 23 *Part Two*: *Parasession on Autosegmental and Metrical Phonology*, Chicago Linguistic Society, Chicago.

Bach, E.: 1979, 'Control in Montague Grammar,' *Linguistic Inquiry* **10**, 515–31.

Bach, E.: 1984, 'Some Generalizations of Categorial Grammars', in F. Landman and F. Veltman (eds.), *Varieties of Formal Semantics*, Foris, Dordrecht, pp. 1–23.

Bat-El, O.: 1989, *Phonology and Word Structure in Modern Hebrew*, unpublished Ph.D. dissertation, UCLA.

Brame, M.: 1970, *Arabic Phonology: Implications for Phonological Theory and Historical Semitic*, unpublished Ph.D. dissertation, MIT.
Bromberger, S. and M. Halle: 1988, 'Conceptual Issues in Morphology', unpublished manuscript, MIT.
Broselow, E.: 1982, 'On Predicting the Interaction of Stress and Epenthesis', *Glossa* **16**, 115–32.
Broselow, E.: 1983, 'Salish Double Reduplications: Subjacency in Morphology,' *Natural Language and Linguistic Theory* **1**, 317–46.
Broselow, E. and J. McCarthy: 1984, 'A Theory of Internal Reduplication', *The Linguistic Review* **3**, 25–88.
Carrier-Duncan, J.: 1984, 'Some Problems with Prosodic Accounts of Reduplication,' in M. Aronoff and R. Oehrle (eds.), *Language Sound Structure*, MIT Press, Cambridge, Massachusetts, pp. 260–86.
Clements, G. N.: 1985, 'The Problem of Transfer in Nonlinear Phonology', *Cornell Working Papers in Linguistics* **5**, 38–73.
Dieterici, F. (tr.): 1852, *Ibn 'Akîl's Commentar zur Alfijja des Ibn Mâlik*, Ferd. Dümmlers Verlags-Buchhandlung, Berlin.
Dixon, R. M. W.: 1972, *The Dyirbal Language of North Queensland*, Cambridge University Press, Cambridge.
Dixon, R. M. W.: 1977, *A Grammar of Yidin^y*, Cambridge University Press, Cambridge.
Emeneau, M.: 1955, *The Kolami Language*, University of California Press, Berkeley.
Fleisch, H.: 1968, *L'arabe classique: Esquisse d'une structure linguistique*, Dar El-Machreq, Beirut.
Hale, K. and A. Lacayo Blanco: 1988, *Vocabulario preliminar del ULWA (sumu meridional)*, Centro de investigaciones y documentacion de la Costa Atlantica, Karawala, Zelaya Sur, Nicaragua, and Center for Cognitive Science, MIT.
Hammond, M.: 1988, 'Templatic Transfer in Arabic Broken Plurals', *Natural Language and Linguistic Theory* **6**, 247–70.
Harris, J.: 1983, *Syllable Structure and Stress in Spanish*, MIT Press, Cambridge, Massachusetts.
Hayes, B.: 1982, 'Extrametricality and English Stress', *Linguistic Inquiry* **13**, 227–76.
———: 1985, 'Iambic and Trochaic Rhythm in Stress Rules', in M. Niepokuj, M. VanClay, V. Nikiforidou, and D. Feder (eds.), *Proceedings of the Berkeley Linguistics Society* **11**, 429–46.
———: 1986, 'Stress Shift and Syncope in Cyrenaican Bedouin Arabic', paper given at the Winter Meeting of the Linguistic Society of America, New York.
———: 1987, 'A Revised Parametric Metrical Theory', in J. McDonough and B. Plunkett (eds.), *NELS* 17, GLSA, Department of Linguistics, University of Massachusetts, Amherst.
Hill, J.: 1970, 'A Peeking Rule in Cupeño', *Linguistic Inquiry* **1**, 534–39.
Hewitt, M.: 1989, 'Quantity Sensitivity in Alutiiq', unpublished manuscript, Brandeis University.
Hewitt, M. and A. Prince: 1989, 'OCP Effects in the N. Karanga (Shona) Verbal Paradigm', paper presented at WCCFL VIII, University of British Columbia, Vancouver.
Hoeksema, J.: 1985, *Categorial Morphology*, Garland Publishing, New York.
Hoeksema, J. and R. Janda: 1988, 'Implications of Process-Morphology for Categorial Grammar', in R. Oehrle, E. Bach, and D. Wheeler (eds.), *Categorial Grammars and Natural Language Structures*, D. Reidel, Dordrecht, pp. 199–247.
Howell, M. S.: 1986 (repr.) *A Grammar of the Classical Arabic Language*, Gian Publishing House, Delhi.
van der Hulst, H.: 1984, *Syllable Structure and Stress in Dutch*, Foris, Dordrecht.
Hyman, L.: 1985, *A Theory of Phonological Weight*, Foris, Dordrecht.
Inkelas, S.: 1989, 'The Representation of Invisibility', manuscript, Stanford University and UCLA.

Itô, J.: 1988, 'Minimality Templates and Derived Forms', lecture, University of Massachusetts, Amherst.

Jensen, J.: 1977, *Yapese Reference Grammar*, University of Hawaii Press, Honolulu.

Kenstowicz, M.: 1981, 'The Metrical Phonology of Arabic Accent', talk presented at the UCLA-USC Lake Arrowhead Conference on Phonological Theory.

Kenstowicz, M. and C. Pyle: 1973, 'On the Phonological Integrity of Geminate Clusters', in M. Kenstowicz and C. Kisseberth (eds.), *Issues in Phonological Theory*, Mouton, The Hague.

Kiparsky, P.: 1973, 'Phonological Representations', in O. Fujimura (ed.), *Three Dimensions of Linguistic Theory*, TEC, Tokyo, pp. 3–136.

Levin, J.: 1983, 'Reduplication and Prosodic Structure', unpublished manuscript, MIT. Cambridge Massachusetts.

———: 1985, *A Metrical Theory of Syllabicity*, unpublished Ph.D. dissertation, MIT, Cambridge, Massachusetts.

Levy, M.: 1971, *The Plural of the Noun in Modern Standard Arabic*, unpublished Ph.D. dissertation, University of Michigan.

Marantz, A.: 1982, 'Re Reduplication', *Linguistic Inquiry* **13**, 435–82.

Marantz, A. and L. McIntyre: 1986, 'Reduplication Revisited', unpublished manuscript, University of North Carolina.

McCarthy, J.: 1979, *Formal Problems in Semitic Phonology and Morphology*, unpublished Ph.D. dissertation, MIT, Cambridge, Massachusetts, distributed by Indiana University Linguistics Club.

———: 1981, 'A Prosodic Theory of Nonconcatenative Morphology', *Linguistic Inquiry* **12**, 373–418.

———: 1983, 'A Prosodic Account of Arabic Broken Plurals', in I. Dihoff (ed.), *Current Trends in African Linguistics* **1**, Foris, Dordrecht, pp. 289–320.

———: 1984, 'Prosodic Organization in Morphology', in M. Aronoff and R. Oehrle (eds.), *Language Sound Structure*, MIT Press, Cambridge, Massachusetts, pp. 299–317.

———: 1986, 'OCP Effects: Gemination and Antigemination', *Linguistic Inquiry* **17**, 207–63.

McCarthy, J. and A. Prince: 1986, 'Prosodic Morphology', manuscript, University of Massachusetts, Amherst, and Brandeis University.

McCarthy, J.: 1987a, 'Prosodic Categories in Infixing Morphology', talk delivered by Prince at MIT, Cambridge, Massachusetts, March 19, 1987.

———: 1987b, 'Categories and Operations in Prosodic Morphology', University of Arizona videotape, October 16, 1987, Tucson.

———: 1988, 'Quantitative Transfer in Reduplicative and Templatic Morphology', in Linguistic Society of Korea (ed.), *Linguistics in the Morning Calm* 2, Hanshin Publishing Co., Seoul, pp. 3–35.

———: forthcoming a, 'Prosodic Constraints on Word Structure', in M. Eid and J. McCarthy (eds.), *Perspectives on Arabic Linguistics: Papers from the Second Symposium on Arabic Linguistics*, J. Benjamins, Amsterdam.

———: forthcoming b, *Prosodic Morphology*, MIT Press, Cambridge, Massachusetts.

Mester, A.: 1986, *Studies in Tier Structure*, unpublished Ph.D. dissertation, University of Massachusetts, Amherst.

Mester, A.: 1988, 'Patterns of Truncation in Japanese', unpublished manuscript, University of Texas, Austin.

Moravcsik, E.: 1978, 'Reduplicative Constructions', in J. Greenberg (ed.), *Universals of Human Language*, Stanford University Press, Stanford, pp. 297–334.

Mutaka, Ng. and L. Hyman: 1987, 'Syllables and Morpheme Integrity in Kinande Reduplication', unpublished manuscript., University of Southern California, Los Angeles.

Nash, D.: 1979, 'Yidin^y Stress: A Metrical Account', *Cunyforum* **7/8**, 112–30.

———: 1980, *Topics in Warlpiri Grammar*, unpublished Ph.D. dissertation, MIT, Cambridge, Massachusetts.

Pinker, S. and A. Prince: 1988, 'On Language and Connectionism: Analysis of a Parallel Distributed Processing Model of Language Acquisition', *Cognition* **28**, 73–194.

Pollard, Carl: 1984, *Generalized Phrase Structure Grammars, Head Grammars, and Natural Language*, unpublished Ph.D. dissertation, Stanford University.

Poser, W.: 1984, 'Hypocoristic Formation in Japanese', in M. Cobler, S. MacKaye, and M. Wescoat (eds.), *Proceedings of the West Coast Conference on Formal Linguistics* 3, Stanford Linguistics Association, Stanford University.

Pulleyblank, D.: 1986, *Tone in Lexical Phonology*, D. Reidel, Dordrecht.

Prince, A.: 1980, 'A Metrical Theory for Estonian Quantity', *Linguistic Inquiry* **11**, 511–62.

———: 1983, 'Relating to the Grid', *Linguistic Inquiry* **14**, 19-100.

———: 1989, 'Metrical Forms', in P. Kiparsky and G. Youmans (eds.), *Rhythm and Meter*, Academic Press, New York, pp. 25–80.

Schmerling, S.: 1983, 'Montague Morphophonemics', in J. Richardson, M. Marx, and A. Chukerman (eds.), *Papers from the Parasession on the Interplay of Phonology, Morphology, and Syntax*, Chicago Linguistic Society, Chicago, pp. 222–37.

Selkirk, E.: 1984, *Phonology and Syntax: The Relation between Sound and Structure*, MIT Press, Cambridge, Massachusetts.

Smeaton, B. H.: 1973, *Lexical Expansion due to Technical Change*, Indiana University Press, Bloomington.

Smith, N.: 1985, 'Spreading, Reduplication, and the Default Option in Miwok Nonconcatenative Morphology', in H. van der Hulst and N. Smith (eds.), *Advances in Nonlinear Phonology*, Foris, Dordrecht, pp. 363–80.

Smith, N. and B. Hermans: 1982, 'Nonconcatenatieve woordvorming in het Sierra Miwok', *Glot* **5**, 263–84.

Steriade, D.: 1982, *Greek Prosodies and the Nature of Syllabification*, unpublished Ph.D dissertation, MIT, Cambridge, Massachusetts.

Tateishi, K. 1989, 'Theoretical Implications of a Japanese Musicians' Language', paper presented at WCCFL VIII, University of British Columbia, Vancouver.

Vadet, J.: 1955, 'Contribution à l'histoire de la métrique arabe', *Arabica* **2**, 313–21.

Wehr, H.: 1971, *Dictionary of Modern Written Arabic*, J. M. Cowan (ed.), Spoken Language Services, Ithaca, New York.

Wilkinson, K.: 1986, 'Syllable Structure and Lardil Phonology', unpublished manuscript, University of Massachusetts, Amherst.

Woodbury, A.: 1985, 'Meaningful Phonological Processes: A Consideration of Central Alaskan Yupik Eskimo Prosody', unpublished manuscript, University of Texas, Austin.

Wolf, J.: 1966, *Beginning Cebuano*, Part 1, Yale University Press, New Haven.

Wright, W.: 1971 (repr.), *A Grammar of the Arabic Language*, Cambridge University Press, Cambridge.

Yip, M.: 1988, 'Template Morphology and the Direction of Association', *NLLT* **6**, 551–77.

Received 4 August 1988
Revised 17 March 1989

McCarthy
Department of Linguistics
University of Massachusetts
Amherst, MA 01003
mccarth@umass.bitnet

Prince
Linguistics and Cognitive Science Program
Brandeis University
Waltham, MA 02254-9110
prince@brandeis.bitnet