Contrastive Structure in the Underlying Representation of Long Vowels in Mayo*

Larry Hagberg
University of Arizona
Decermber 21, 1989

0. Introduction.

Several structural configurations for vowel length have been proposed since the introduction of moraic theory. These proposals may be divided into two categories: Hyman (1985), McCarthy and Prince (1986) and Hayes (1989) all claim that underlying vowel length must be represented on the moraic tier, so that a long vowel has one more underlying mora than a short vowel, but both types of vowels crucially have only one root node.[1] I refer to this as the Moraic Tier (MT) approach. A crucial assumption of the MT approach is that an underlyingly long vowel is a single vowel rather than a sequence of two short vowels. In contrast, Selkirk (1988) claims that moras are not present in underlying representation. Instead, vowel length is represented on the root tier so that a long vowel crucially has two root nodes while a short vowel has only one root node. This view, which I refer to as the Root Tier (RT) approach, is essentially a return to the Structuralist view of vowel length as a sequence of two identical vowels.[2] Under the RT approach, all moras are assigned in the course of the derivation, while the MT approach allows for both underlying and derived moras. The MT approach and the RT approach are represented schematically in figure A, where μ represents a mora, R represents a root node and P stands for 'place', representing everything that is subordinate to the root.

In this paper I examine the empirical claims of each of these views as applied to Mayo, a Uto-Aztecan language of northwestern Mexico. From the

<u>Figure A</u> <u>Moraic Tier</u>: <u>Root Tier</u>:

```
μ   μ          R   R
 \ /            \ /
  R              P
```

Hyman 1985, McCarthy & Prince 1987 Selkirk (1988)

distribution of stress I demonstrate that Mayo has two types of underlying vowel length which contrast in the manner in which they syllabify. One type of long vowel spans two syllables in phonetic representation (PR), while the other type is entirely contained within a single syllable. This contrast cannot be accounted for if only the MT approach to vowel length is assumed, nor can it be accounted for if only the RT approach is assumed. Rather, I show that both approaches are needed in order to account for the distribution of stress in words with underlying vowel length. The theoretical implication is that a language may represent temporal contrasts on either the root tier or the moraic tier, or both; Mayo is an example of the latter.

My central claim, then, is that Mayo exhibits two different structural representations of underlying vowel length. This means that terms such as *vowel length* and *long vowel* are ambiguous, so throughout this paper I use these terms only in a descriptive sense. Where I wish to formally distinguish between the two types of length, I refer to either the MT approach or the RT approach.

The paper is organized as follows: In section one I illustrate lengthening of underlyingly short vowels and consonants in Mayo and propose two prosodic rules to explain these phenomena. In section two I discuss the interaction of underlying vowel length with metrical structure and the two rules proposed in section one. I show that, although every long vowel is bimoraic at PR, in some stems containing a long vowel these two moras are tautosyllabic while in other stems they are heterosyllabic. I then use syllabification theory to explore the possible underlying representations that might account for this surface contrast. I conclude that both the MT approach and the RT approach are needed in order to account for all the surface manifestations of length that occur in Mayo.

1.0. Derived length.

Length occurs in both vowels and consonants in Mayo, but in most cases it is derived rather than underlying. Hagberg (1988) describes two parallel connections in Mayo between metrical structure and derived length. On the one hand, there is a correlation between the distribution of derived vowel length and the distribution of second syllable stress. On the other hand, there is a similar correlation between the distribution of derived consonant length and the distribution of first syllable stress. I discuss these two sets of phenomena in sections 1.1 and 1.2, respectively.

1.1. Derived vowel length and second syllable stress.

Phonetically long vowels occur quite frequently in Mayo, but in many instances the length is predictable from the environment. In this subsection I propose two prosodic rules and demonstrate how they account for vowel length in most of the places where it occurs. Section two discusses those cases which are not accounted for by these rules.

Mayo has no instances of a monomoraic word in phonetic representation, but some words can be shown to be monomoraic in underlying representation. For example, the word for 'no' has the form [kaa] when it occurs as a word by itself and the form [ka] whenever anything is attached to it:[3]

(1) Kaá kó'okore. *He is not sick.*
no be sick

(2) ka-tím kó'okore. *They are not sick.*
no-they be sick

(3) ká-k waánte *He doesn't feel any pain.*
no-location feel pain

This alternation in vowel length occurs in the word for 'me' in the same environments; it surfaces as [nee] when it occurs as a word by itself and as [ne] when it is followed by an enclitic:

(4) neé bítcha. *He sees me.*
me sees

(5) ne-chím bítcha. *They see me.*
me-they see

There are two possible approaches to analyzing the length alternations in (1) through (5): Either an underlyingly short vowel gets lengthened, or else an underlyingly long vowel gets shortened. In what follows I consider first the former approach, then the latter, showing that it has to be the case that the vowel length in (1) through (5) is derived, not underlying.

If the vowel length in (1) through (5) is derived, then some kind of rule is needed which, in pre-theoretic terms, lengthens a word-final short vowel when it is the only vowel in the word. This is formalized in (6).

6. Vowel Lengthening: V ⟶ V: / $[(C)\_\_\_]_W$

Although Vowel Lengthening accounts for the data presented thus far, it fails to capture a significant generalization about monosyllabic words in Mayo: words with the phonetic form [CV:] are attested, as are words with the form [CVC], but [CV:C] is unattested in phonetic form. Moraic theory provides a means of expressing this generalization. Following Hayes (1989) and Archangeli (1989), I assume that syllable onsets are not moraic, but that many languages assign the same number of moras to a long vowel as they do to a sequence consisting of a short vowel plus a coda. This seems to be the case in Mayo, since there is no vowel length in (3). In order to express this generalization, Vowel Lengthening can be replaced by Mora Insertion, a rule which adds a mora to a monomoraic word:

7. Mora Insertion (MI): $\mu \longrightarrow \mu\,\mu$ / [___ $]_W$

The advantage of Mora Insertion over Vowel Lengthening is that the former can treat vowel length and coda consonants in a uniform manner, whereas the latter makes no predictions about consonants. Thus, while both rules are descriptively adequate for the data presented thus far, Mora Insertion is more general and thus makes stronger predictions. I return to this point in section 1.2, where I show that Mora Insertion accounts not only for vowel lengthening but also for consonant gemination. In the meantime, I assume that Mora Insertion is the correct version of the rule, and that the inserted mora copies the melody of the vowel in order to produce the attested surface forms in the above examples.[4]

So far I have based the rule of Mora Insertion on examples involving only two words, both of which are function words. However, Mora Insertion is independently needed to account for the fact that monomoraic words are unattested at phonetic representation in Mayo. Every word, even if monosyllabic, has at least two moras in phonetic representation. Two more examples of function words are given in (8). In each of these forms, the vowel length disappears when a clitic is added, just as it does in the earlier examples involving length alternations.

8. a. kee *ke 'not yet'

   b. hee *he 'yes'

Mayo also has a number of verb roots that are underlyingly monomoraic. Mora Insertion applies to these words as expected in (9) through (13). In the (a) form of each of these examples, the underlyingly monomoraic root gains a second mora from the present tense (PRES) suffix. In each of the (b) forms, however, Mora Insertion applies because the root has no affixes.

9.	a. wé-ye go-PRES	'go (SG)'	b. wée béchi'ibo go for		'in order to go (SG)'
10.	a. yá-wa make-PRES	'make'	b. yáa béchi'ibo make for		'in order to make'
11.	a. é-ya think-PRES	'think'	b. ée béchi'ibo think for		'in order to think'
12.	a. bá-re intend-PRES	'intend'	b. báa béchi'ibo intend for		'in order to intend'
13.	a. ho-yé sit-PRES	'sit (PL)'	b. hoó béchi'ibo sit for		'in order to sit (PL)'

I have been assuming that all of the words examined thus far contain only underlyingly short vowels; the alternative is to assume that these vowels are underlyingly long. If the latter were the case, then some kind of vowel shortening rule would be needed in order to account for the forms in (2), (3) and (5) as well as the (a) forms in (9) through (13). However, there are many Mayo words which contain a long vowel that never alternates with a short vowel. I know of no monosyllabic words containing an underlyingly long vowel, but several polysyllabic examples are given in (14)(a), (15)(a) and (16)(a).[5] Each of these underived forms contrasts with the unrelated (b) form, which contains only short vowels.

14. a. yóoko 'jaguar' b. yóka 'paint'

15. a. naáte 'begin' b. nátemae 'ask'

16. a. boorók 'toad' b. porówim 'type of lizard'

If it were the case that a long vowel shortens non-word-finally, as might be concluded from the data presented earlier, then there would be no explanation for the vowel length in (14)(a), (15)(a) and (16)(a). If, however, each of the vowels in (1) through (5) and (8) through (13) is underlyingly short, then Mora Insertion accounts for all the length alternations, and the vowel length in (14) through (16) is underlying. Thus, the correct approach to analyzing the length alternations presented thus far is to assume that underlyingly short vowels undergo lengthening as the result of a rule such as Mora Insertion. To assume the opposite, i.e., that underlyingly long vowels shorten in non-final position, would fail to account for the length contrasts in (14) through (16).

More examples of vowel lengthening are presented in (17) through (21). However, each of these words is bisyllabic, and the first vowel lengthens when the word occurs in phrase final position but not elsewhere.[6]

	Phrase-finally	Elsewhere	Gloss
17.	a. sií-ka	b. si-ká	'depart-PAST (singular)'
18.	a. chaá-ye	b. cha-yé	'shout-PRES'
19.	a. taábu	b. tabú	'rabbit'
20.	a. kiíchul	b. kichúl	'cricket'
21.	a. tuúrus	b. turús	'spider'

Since Mora Insertion has already proven useful in accounting for vowel length alternations, I follow the same approach in analyzing (17) through (21). In order to account for the fact that the environment for lengthening in these examples is not word-final, another rule is needed to create a word-final environment. This rule, which I call *Phrase-Final Invisibility* (PFI), is stated in (22), where σ represents a syllable, $]_W$ represents the right edge of a word and ## represents a phrase boundary.

22. <u>Phrase-Final Invisibility (PFI)</u>: $\sigma\ ]_W \longrightarrow\ ]_W\ \sigma\ /\ \_\_\_\ \#\#$

Following Inkelas (1989), I assume that invisibility effects result from the exclusion of some member of a morphological string from the corresponding rule domain. Thus, PFI has the effect of rendering a phrase-final syllable invisible with respect to all subsequent rule applications within that stratum. The effect of PFI upon stress assignment and Mora Insertion is illustrated in (23), where F corresponds to the stress foot and W corresponds to the bimoraic word template.

```
23.      F          F
        / \         |
     s i k a  -->  s i<k a>
       |   |        |\
       μ   μ        μ μ
        \ /         |/
         W          W
```

Each of the words in (17) through (21) has stress on the second syllable except when that syllable occurs phrase-finally, in which case stress shifts to the preceding syllable. This in itself is evidence for PFI, but additional evidence comes from the lengthening of the first vowel in each of the phrase-final forms. This lengthening is accounted for by MI only if the environment for MI is created in each instance by a rule such as PFI, which renders the final syllable of a phrase invisible to the subsequent application of all rules.

The above discussion assumes that each of the (b) forms in (17) through (21), rather than the (a) forms, more closely represents the respective underlying form. But what if the opposite were the case? If each of the words in (17) through (21) were assumed to have (a) as its underlying form, then two problems would arise. First, the shortening of the vowel in each of the (b) forms cannot be accounted for without stipulating a new rule or constraint. Second, the shifting of stress from the next syllable must be accounted for. It would not be correct to say that stress simply falls on the second mora, because this does not hold in (24). The form of (24)(a) is identical to that of (24)(b), with one exception: stress occurs on the first syllable in (a) and on the second syllable in (b).

	Phrase-finally	Elsewhere	Gloss
24.	a. bwík-su	b. bwik-sú	'sing-COMPLETE'

If one assumes that stress falls on whichever syllable contains the word's second mora, then (24)(b) has stress in the wrong place. If, on the other hand, stress always goes on the second mora and then shifts when that mora is non-vocalic, the shift is leftward in (24)(a) and rightward in (24)(b). All such awkward complications are avoided by assuming that stress normally falls on the second syllable in these words, and that phrase-final stress is forced to shift leftward. This is strong evidence for PFI.

Why is no vowel lengthening observed in (24)(a)? After the phrase-final syllable has been rendered invisible by the application of PFI, the remainder of the word still has two moras, so the application of MI is blocked. If the morphology does not create a consonant cluster, however, then MI applies and vowel lengthening is observed, as in (25)(a).

	Phrase-finally	Elsewhere	Gloss
25.	a. bwiík-a	b. bwik-á	'sing-PRES'

The same effect can be seen by comparing (26) with (27):

26.	a. noók-a	b. nok-á	'speak-PRES'
27.	a. nók-la	b. nok-lá	'speak-PERFECTIVE'

The failure of MI to apply to (24)(a) *(bwíksu *bwiíksu)* and (27)(a) *(nókla *noókla)* supports the claim that codas are moraic.

I have given two arguments demonstrating that the vowel length in the phrase-final forms in (17) through (21) is derived rather than underlying.

First, the words in (17) through (21) exhibit vowel length only in a very restricted environment, whereas the words in (14)(a), (15)(a) and (16)(a) have vowel length in all environments. This indicates that the latter set of examples, but not the former set, have underlying vowel length. Second, the stress alternations in (17) through (21) are accounted for by PFI, which is also needed to account for the vowel quantity alternations in the same data. Taken together, then, the rules of MI and PFI are able to account for all of the cases presented thus far in which vowel length alternates with non-length.

Thus far, however, I have considered only words whose stress alternates between the first and second syllables. Next, I examine the effects of MI and PFI on words which have stress on the first syllable in all environments.

1.2. First syllable stress and consonant gemination.

Approximately half of the words listed in Collard and Collard (1962) exhibit the regular stress alternations that are illustrated in the data of section 1.1. Words from the other half of the lexicon exhibit first syllable stress in all environments. This contrast is illustrated in (28) through (32): In each pair of examples, the difference in meaning between the (a) and (b) forms is carried entirely (or almost entirely) by the placement of stress.

	First syllable stress:		Second syllable stress:	
28. a.	nók-nake	'know language-FUT'	b. nok-náke	'speak-FUT'
29. a.	pón-a-ka	'pull-PRES-PART'	b. pon-á-ka	'play-PRES-PART'
30. a.	súw-a-ka	'kill (PL)-PRES-PART'	b. suw-á-ka	'look at-PRES-PART'
31. a.	kóba-ta	'head-ACCUSATIVE'	b. kobá-nake	'defeat-FUT'
32. a.	wát-e-ka	'fall (PL)-PRES-PART'	b. waté	'others'

An interesting feature of all words in both stress categories is that, whenever a prefix is added, stress always shifts one syllable to the left. This is illustrated in (33) through (35).

	First syllable stress:		Second syllable stress:	
33. a.	nó-noknake	'will always know lang'	b. no-nóknake	'will always speak'
34. a.	pó-ponaka	'always pulling'	b. po-pónaka	'always playing'
35. a.	sú-suwaka	'always killing (PL)'	b. su-súwaka	'always looking at'

Hagberg (1989 a) argues for three points, based on the above facts about stress, which are relevant to the analysis of vowel length in Mayo. First, all words with consistent first syllable stress have lexical accent, while words with second syllable stress (eg., those in section 1.2) do not have lexical accent. Following Halle and Vergnaud (1987 a and b), stress is assigned via a set of rules which produce the attested patterns. The details of this analysis are not relevant to the present discussion. The arguments presented here depend only upon the claim that lexical accent accounts for the contrast between first and second syllable stress in Mayo; it does not matter, for this paper, whether lexical accent is a feature of words with first syllable stress or whether it is a feature of words with second syllable stress. Therefore I assume, following Hagberg (1989 a), that lexical accent is a property of words with invariant first syllable stress. Henceforth I refer to words with second syllable stress as *unaccented* and words with first syllable stress as *accented.*

The second claim argued for in Hagberg (1989 a) is that stress in Mayo is *cyclic.* This means that every time an affix is added to a word, the former stress information is lost and the rules of stress are applied to the entire new form which resulted from affixation. The basis for this claim is that stress always shifts leftward when a prefix is added, and no residual stress remains on the syllable that was stressed in the unprefixed form.

The third relevant claim of Hagberg (1989 a) is that lexical accent floats in underlying representation, and that lexical accent is never lost as a result of cyclic affixation. Rather, lexical accent associates cyclically to the leftmost syllable. If this were not true, then the contrast between lexically accented versus unaccented words would be lost whenever a prefix is added. (33) through (35) demonstrate that prefixation does *not* neutralize the stress contrast, so I conclude that lexical accent floats in underlying representation and associates cyclically to the leftmost syllable.

The above analysis of stress is crucial to the analysis of underlying vowel length in section two. But before moving on to section two, I first consider the application of MI and PFI to accented words.

The rules of MI and PFI apply to accented words, but consonant lengthening is generally observed instead of vowel lengthening. This is illustrated in (36) through (39), where the onset of a phrase-final syllable geminates.

Phrase-finally	Elsewhere	Gloss
36. a. míssi	b. mísi	'cat'
37. a. chókki	b. chóki	'star'
38. a. kóbba	b. kóba	'head'
39. a. tónno	b. tóno	'knee'

Phrase-final onset gemination is not attested, however, when the first syllable is already bimoraic. This can be seen by comparing (40) with (41) and (42) with (43).

Phrase-finally	Elsewhere	Gloss
40. a. chúpp-a	b. chúp-a	'harvest-PRES'
41. a. chúp-su-k	b. chúp-su-k	harvest-COMPLETE-PAST'
42. a. mákk-a	b. mák-a	'give-PRES'
43. a. mák-nake	b. mák-nake	'give-FUT'

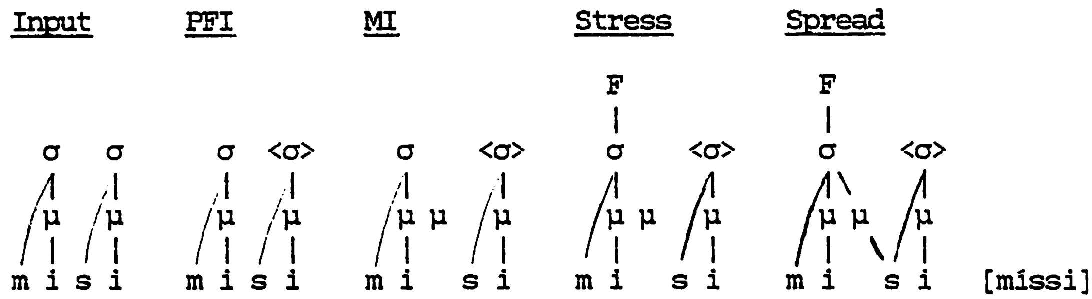
Thus, accented words exhibit quantity alternations in environments that correspond, in all relevant respects, to the environments that were attested for similar alternations in unaccented words in the previous subsection. However, the respective quantity alternations in these two classes of words differ in one crucial aspect: accented words exhibit onset gemination while unaccented words exhibit vowel lengthening.

How does one account for this asymmetry? The fact that accented words exhibit onset gemination in the same environment where unaccented words exhibit vowel lengthening strongly suggests that the two kinds of quantity alternations are, at some level, the result of a single process.[7] In what follows I propose an analysis that is consistent with this suggestion.

In section 1.1 it was assumed that an inserted mora receives its melody by copying a vowel. The data in (36) through (43), however, seem to indicate that an inserted mora receives its melody from a consonant. In order to reconcile these seemingly contradictory conclusions, I assume that an inserted mora always receives its melody through a rule of Leftward Spread, stated in (44), except when the application of this rule is blocked by some independent constraint. In the latter case, the empty mora is filled by rightward spread. This is in principle the same as Compensatory Lengthening (Hayes 1989), which is needed in many languages to account for derived vowel length which is triggered by the loss of a following coda. In the data under consideration here, no coda has been lost, but the environment for 'Compensatory Lengthening' has nonetheless been created by the application of MI. Since the term 'Compensatory Lengthening' might be misleading, henceforth I refer to this phenomenon in Mayo as *Vowel Copying*.

44. <u>Leftward Spread</u>: µ
```
µ
 \
  R
```

Leftward Spread simply states that an empty mora receives its melody by spreading from the closest root node on its right. The application of Leftward Spread to a phrase-final accented word is illustrated in (45), where F represents the stress foot and σ is a syllable node. Since PFI is a phrase-level rule, I assume that syllabification and stress apply on the word-level cycles that precede it. However, Hagberg (1989 c) shows that the cyclic rules of stress assignment apply even at the phrase level (after clitic movement, for example). Therefore, I assume that the input to PFI has no stress due to the erasure of metrical structure at the beginning of each cycle.

45. 

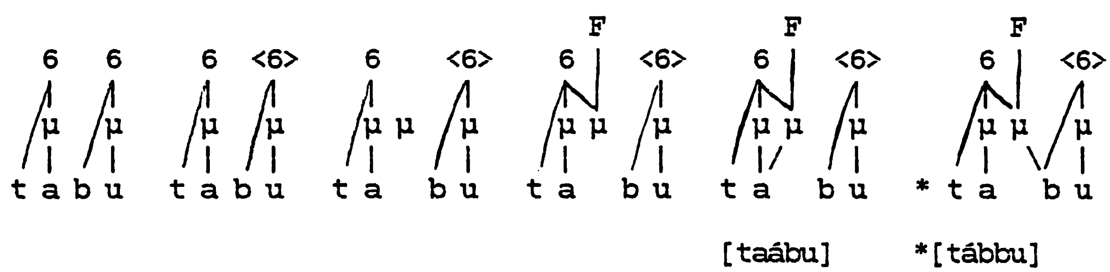
The final step of the derivation in (45) might be objected to on the grounds that the [s] which is spreading its root node is part of the invisible syllable. This objection is based on the claim of Inkelas (1989) (and others) that an invisible element is not available for participation in the application of a rule. However, Inkelas (1989) also claims that an invisible element becomes visible when passing into the next cycle of the derivation. If this is true, then the availability of [s] for spreading in (45) can be explained by assuming that Leftward Spread applies at a cycle that is later than the cycle in which PFI applies. For example, it may be that Leftward Spread applies during Phonetic Implementation.

Now consider the application of the same sequence of rules to an unaccented word in (47). In order to derive the correct result with respect to the direction of spreading, I assume (46), which is a paraphrase from Halle and Vergnaud (1987 a, page 138) regarding the application of the rules of stress assignment:

46. When a word is too short to construct a full metrical constituent, the constituent boundaries are constructed by default at the edges of the word.

Hagberg (1989 a) shows that Mayo stress feet are bounded, right-headed, and constructed from the left edge of the word non-iteratively. In unaccented words consisting of two or more syllables, this means that stress falls on the second syllable. (46) says that, if a word has only one syllable, stress should fall on the right edge of that syllable. If this syllable contains a short vowel, then the right and left edges are identical, for purposes of stress assignment, since consonants are universally prohibited from bearing stress.[8] If, on the other hand, a monosyllabic unaccented word contains a long vowel, then (46) predicts that stress should be assigned to the rightmost mora. I claim that this is in fact what happens in (47).

47. Input PFI MI Stress Spread



Since the second mora in (47) receives stress before Leftward Spread has a chance to apply, Leftward Spread is prevented from applying by the universal constraint against consonants bearing stress. The mora therefore gets its melody from the vowel.

I have thus provided an account of why accented words exhibit onset gemination while unaccented words exhibit vowel lengthening: An inserted mora receives its melody through Leftward Spread, except when the application of this rule is blocked by some independent constraint. In the latter case, the empty mora is filled by rightward spread.

Other explanations of these facts are also possible (see, for example, Hagberg 1989 b), but any analysis will have to distinguish between the creation of a moraic slot, which occurs in both unaccented and accented words, and the means by which that moraic slot gets filled. Even if the derivations given in (45) and (47) turn out to be incorrect, it still has to be the case that all the quantity alternations discussed thus far are derived rather than underlying.

The above analysis does make one prediction that can be tested. If MI is indeed responsible for all of the quantity alternations that have been discussed thus far, then a monomoraic accented word should exhibit vowel lengthening instead of consonant gemination if no consonant is available for copying. Such vowel lengthening in accented words is in fact attested in (48) and (49). Each of the (a) forms exhibits phrase-final consonant gemination while each of the (b) forms lacks it, as expected. In each of the (c) forms, however, the verb is underlyingly monomoraic, and vowel lengthening is attested.

48.	a. wé-yye go-PRESENT	b. wé-ye-'bwan go-PRESENT-'well'	c. wée béchi'ibo go for
49.	a. yá-wwa make-PRESENT	b. yá-wa-'bwan make-PRESENT-'well'	c. yáa béchi'ibo make for

This prediction is additionally borne out by (50) and (51). The occurence of stress on the first syllable in all environments indicates that *báre* and *táruk* have lexical accent, but they both exhibit phrase-final vowel lengthening instead of the expected onset gemination.

	Phrase-finally		Elsewhere	Gloss
50.	báare	*bárre	báre	'intend'
51.	táaruk	*tárruk	táruk	'roadrunner'

This unexpected vowel lengthening can be accounted for by positing the syllable structure constraint in (52), which states that [r] cannot occur in coda position.

52. $* \; r]_{\sigma}$

The syllable structure constraint in (52) is independently attested by the fact that [r] is never found in coda position in Mayo. Given this constraint, the vowel length alternations in (50) and (51) can be attributed to the blocking of the otherwise-regular onset gemination that is observed in (36) through (39). The constraint in (52) blocks this process from applying to [r] in (50) and (51), so the mora gets filled by Vowel Copying, just as predicted.

In summary, each of the quantity alternations presented thus far can be attributed to a sequence of two distinct processes. First, an empty mora is created by the application of MI, then that mora obtains its melody through the application of some other independent process. I have given two kinds of evidence to support this proposal. First, the environment for onset gemination is the same as the environment for vowel lengthening in all relevant respects. Second, Vowel Copying is observed when onset gemination is blocked from applying.

Since the goal of this paper is to explore the underlying representation of vowel length, I do not discuss any further the orthogonal question of why accented words exhibit onset gemination where unaccented words exhibit vowel lengthening. I simply assume, based on the above arguments, that a single set of prosodic rules and constraints is interacting with accentual information to produce the quantity alternations that are observed in both accented and unaccented words. In the next section I present some apparent violations of these rules and constraints and demonstrate that such violations can be explained only by an underlying contrast in vowel quantity.

2.0. Underlying vowel length.

In the previous section I showed that the twin phenomena of onset gemination and vowel lengthening can be derived from lexical accent plus a set of prosodic rules. These include the following: Mora Insertion (MI), Phrase-Final Invisibility (PFI), Leftward Spread and Vowel Copying. I also demonstrated that some vowels are underlyingly long, although I did not propose any formal representation of vowel length.

In section 2.1 I discuss the distribution of vowel length in Mayo and show that underlying vowel length occurs in both accented and unaccented words. I also demonstrate that the two moras of an underlyingly long vowel can be either tautosyllabic or heterosyllabic in phonetic representation. In section 2.2 I explore the implications of this surface contrast for the underlying representation of vowel length. I conclude that these two phonetic contrasts correspond to an underlying contrast between the MT and RT representations of vowel length.

2.1. The distribution of underlying vowel length.

Underlyingly long vowels occur in roughly 5% of Mayo words. As stated in the previous section, I consider a vowel to be underlyingly long if and only if its length cannot be accounted for by the application of MI. For example, in (53)(b) and (54)(b) there are enough syllables to block the application of MI, yet the first vowel surfaces as long.

	Phrase-finally	Elsewhere	Gloss
53.	a. yóoko	b. yóoko	'jaguar'
54.	a. béete	b. béete	'burn'

Underlying vowel length can occur on either the first or second vowel of a word. When vowel length occurs on the second vowel, only two stress patterns are attested. Stress can fall either on the first (short) vowel, as in (55), or else it can fall on the first mora of the second vowel, as in (56). Thus, it must be that (55) is accented and (56) is unaccented.

55. Accented: téwaa-tua 'name-CAUS'

56. Unaccented: tukáa-po *tukaápo 'at night'

The pattern in (56'), where stress occurs on the second mora of the second vowel, is never attested in Mayo.

56'. [(C)VCVV́]

The existence of (56) versus the absence of (56') from Mayo is significant. It implies that, whenever the first vowel of a word is underlyingly short, stress has only two possible locations where it can occur. In contrast, (57) through (59) show that, whenever the first vowel of a word is underlyingly long, stress has *three* possible locations where it can occur.

57. Accented: yóoko-ta 'jaguar-ACCUS'

58. Unaccented: booró'ok-im 'toad-PLURAL'

59. ? ? ? naáte-nake 'begin-FUTURE'

The data in (57) through (59) present a problem in that, when the initial vowel is long, there are *three* possible locations for stress rather than two. Since the presence of lexical accent produces first syllable stress, it is clear that *yóokota* must have lexical accent and that *booró'okim* must be unaccented, but what is the status of *naátenake*? The question is, how can these three stressed moraic positions be mapped into what otherwise appears to be a binary distribution of stress at the syllable level?

This question actually involves two distinct but related issues. I demonstrated in section one that Mayo stress feet are built on syllables rather than moras, and that only the first two syllables of a word are eligible to bear stress. Therefore, it must be the case that each of the three stress-bearing positions in (57) through (59) is contained within the first two syllables of the word. The first issue to be addressed, then, is this: What is the syllable pattern, at the phonetic level, of each of the words in (55) through (59)? Second, what is minimally required to be present in underlying representation in order for each of those syllable patterns to surface? In what follows I address the first issue; the second issue is discussed in section 2.2.

What, then, are the syllable patterns, at the phonetic level, of the words in (55) through (59)? In order to answer this question, consider the long vowel in *tukáapo*. If the two moras of its long vowel are tautosyllabic, then (60)(a) follows as a logical consequence. On the other hand, it is also possible that the two moras of the long vowel are heterosyllabic, as represented in (60)(b).

60. a. If moras are tautosyllabic: b. If moras are heterosyllabic:

```
   ´
   σ  -->   σ                            σ   σ
  / \      / \                           |   |
           ´                             |   |
 μ   μ    μ   μ              μ   μ  -->  μ   μ
  \ /      \ /                \ /         \ /
   a        a                  a           a
```

(60)(a) states that, if a stressed syllable contains a long vowel, then stress will be realized on the first mora of that vowel. (60)(b) simply represents the alternative possibility that the two moras of the long vowel in *tukáapo* are in different syllables. In either case, the absence of forms like **tukaápo* is accounted for: If a long vowel in the second syllable of a word is stressed, then (60)(a) states that stress has to be realized on the first mora of that long vowel. If, on the other hand, the two moras of that long vowel are distributed between the second and third syllables, then the stress rules of Mayo will assign stress to the first half of the 'long vowel' because only that portion of it is within the domain of stress.

The fact that forms like (56') are unattested is independent evidence for (60)(a).[9] If, contrary to (60)(a), it were possible for stress to occur on the second mora of a syllable, then forms like (56') would be expected to exist. However, the fact that such forms are not attested does *not* imply that the moras of *every* long vowel are tautosyllabic. Rather, it implies that *if* they are tautosyllabic in a given form, then stress can occur on the first but not the second mora.

Now consider *naátenake*. Up to now its accentual status has been an open question, but (60)(a) implies that the two moras of *naátenake* must be heterosyllabic. If they were *not* heterosyllabic, then two conclusions would follow. First, (60)(a) would be falsified, leaving no explanation for the absence of forms like *tukaápo*. Second, *naâtenake* would have to have lexical accent, since first syllable stress occurs only in accented words (apart from the effects of PFI on unaccented words). But this second conclusion is falsified by the following argument:

Recall from the preceding section that prefixation has the net effect of shifting stress leftward because stress assignment is cyclic and lexical accent refloats at the beginning of each cycle. This idiosyncracy of stress provides a diagnostic for determining the accentual status of *naátenake*. (61) unambiguously demonstrates that *naátenake* is unaccented. If it did have accent, the first syllable in (61)(b) would be stressed, but this is not the case.

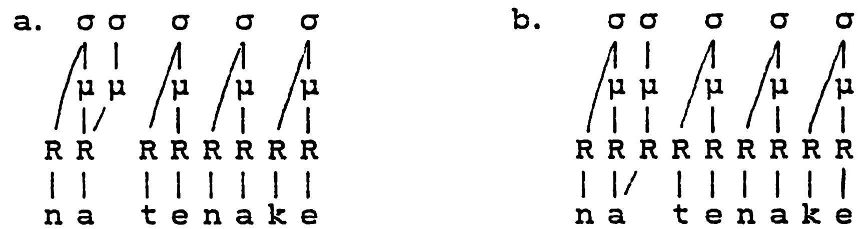
61. a. naáte-nake b. na-náate-nake *ná-naate-nake
begin-FUT HAB-begin-FUT

Since (61) does not have lexical accent, it follows that the two moras of the long vowel must in fact be heterosyllabic: stress falls on the second mora simply because it is in the second syllable.

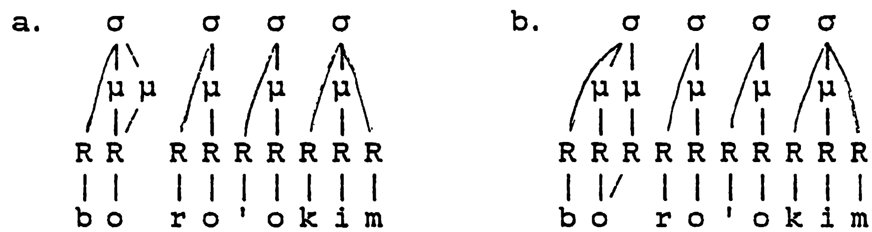
I thus claim that the two moras of the long vowel in *naátenake* are heterosyllabic, but this cannot be true of every instance of underlying vowel length. If the two moras of the long vowel in *booró'okim* were heterosyllabic, then this would be an instance of exceptional third syllable stress. While it would be possible to 'patch up' the analysis with some additional device such as lexical extrametricality, there is no independent motivation for such a device in Mayo. Therefore, I reject this approach and conclude that the moras of some long vowels are tautosyllabic, as in *booró'okim*, while others are heterosyllabic, as in *naátenake*.

We have now narrowed down the range of possible syllable patterns for *naátenake* and *booró'okim* at the phonetic level of representation. The set of logically possible PR's for *naátenake* is presented in (62), and the set of logically possible PR's for *booró'okim* is presented in (63). I demonstrated earlier that all 'long vowels' are bimoraic at PR, and I have just shown that the long vowel in *naátenake* has two syllable nodes at PR, whereas the long vowel in *booró'okim* has only one. This information is represented in both the (a) and (b) versions of (62) and (63). The only remaining question is with regard to root nodes. The MT approach claims that a long vowel has only one root node, as in the (a) versions of (62) and (63), and the RT approach claims that a long vowel has two root nodes, as in the (b) versions.

62.



63.



In the next section I consider these representations in the light of syllabification theory in order to decide between the (a) and (b) versions in each case, and to determine what underlying representations they might correspond to.

2.2. <u>The underlying representations of vowel length</u>.

In the preceding section I concluded that Mayo has two kinds of vowel length which differ at PR as to whether their respective moras are tautosyllabic or heterosyllabic. I showed that each kind of vowel length has two possible phonetic representations with respect to the root tier. These representations are given in (62) for *naátenake*, whose long vowel spans two syllables, and in (63) for *boorő'okim*, whose long vowel is contained within a single syllable. The question to be addressed in this section is this: Given the possible phonetic representations in (62) and (63), what can be inferred about their respective underlying representations? Two conclusions are possible: Either (a) Underlying vowel length has to be represented on the moraic tier for some words and on the root tier for other words; or (b) Syllable nodes must be represented underlyingly (at least for long vowels).

I reject (b) because it calls for additional underlying structure which has not conclusively been shown to be necessary in any language. It might appear to be necessary in Mayo, but in what follows I show that option (a), which posits only independently needed structure, is both descriptively adequate as well as falsifiable. In contrast, I know of no way in which (b) could ever be falsified, at least for Mayo.

Having rejected (b), what does (a) imply? There are only a few possible underlying representations; I claim that the representations in (64) are correct, whereas the alternatives in (65) are incorrect.

```
64. a.                         b.    μ
                                     |
      R R R R R R R R R           R R R R  R R
      | |/  | | | | | |           | | | |  | |
      n a   t e n a k e           b o r o' k m

65. a.  μ                      a'.  μ μ                   b.   μ μ
        |\                          | |                        |/
      R R R R R R R R R           R R R R R R R R R          R R R R  R R
      | |/  | | | | | |           | |/  | | | | | |          | | | |  | |
     *n a   t e n a k e          *n a   t e n a k e         *b o r o' k m
```

Notice in (64) that the underlying difference between the long vowel in *naátenake* and that of *booró'okim* is that the former has an extra root node but no mora, while the latter has a mora but no extra root node. In what follows I argue that the 'long vowel' in *naátenake* is best interpreted as a sequence of two identical root nodes, while that of *booró'okim* is a single vowel with underlying length.

In order to argue in favor of (64) and against (65), a theory of syllabification is needed. Following Selkirk (1982), Ito (1986, 1989) and Hayes (1989), I assume the following features of Mayo syllabification:

66. <u>Syllabification</u> (Selkirk 1982; Hayes 1989; Ito 1986, 1989):

a. Directional application of a template (right to left for Mayo).

b. Each [-high] ROOT gets its own σ node.

c. A maximum of one onset and one coda is adjoined to each σ node.

d. Vowel epenthesis resolves consonant clusters.

The features in (66) consist of the language-particular setting of universal parameters. For example, syllabification is claimed to be directional in all languages, but the direction varies on a language-particular basis.

An example of vowel epenthesis is given in (67); the derivations in (68) and (69) demonstrate that syllabification is from right to left.

67. /yebs/ 'sit (SG)' + /-la/ 'PERF' ⟶ yebísla *yebsíla 'has sat (SG)'

```
68. R to L:           σ           σ    σ       σ  σ    σ
                     /|          /|\  /|      /| /|\  /|
   yebsla ⟶ yebs l a ⟶ y e b i s l a ⟶ y e b i s l a   [yebísla]

69. L to R:       σ           σ    σ        σ    σ  σ
                 /|\         /|\  /|       /|\  /| /|
   yebsla ⟶ y e bsla ⟶ y e b s i la ⟶ y e b s i l a   *[yebsíla]
```

The derivations in (68) and (69) are incomplete; Hayes (1989) and Archangeli (1989) argue that at least some moras are assigned subsequent to syllabification via a separate rule of Weight-by-Position, stated in (70).

70. Weight-by-Position (Archangeli 1989):

```
         μ
         |
R  ⟶   R
```

(70) simply states that a root node gets assigned a mora. This may seem too powerful, but Archangeli (1989) points out that the two independent constraints stated in (71) are sufficient for Yawelmani to ensure that moras end up in the right place.

71. Independent Constraints (Archangeli 1989):

a. Onsets are not moraic

b. Maximal σ = [μ μ]

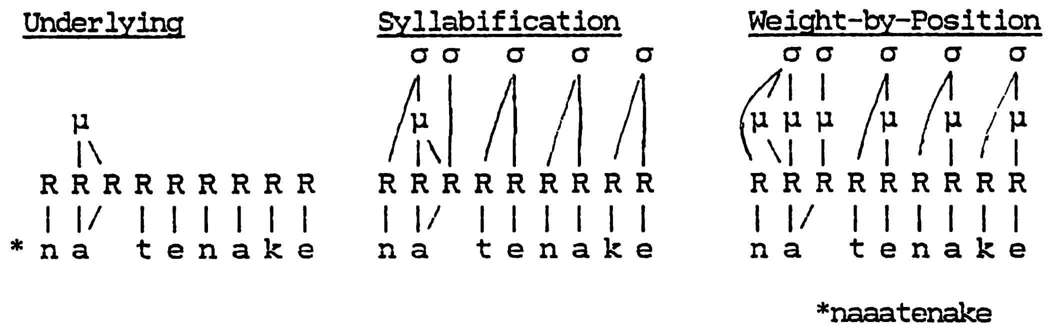
The constraints proposed for Yawelmani in (71) are exactly what is needed in Mayo as well. I already showed in section one that codas are moraic in Mayo; in what follows I demonstrate that Mayo syllables are maximally bimoraic.

Having laid the groundwork in terms of syllabification theory, I now return to the question of why (65) is ruled out in favor of (64). There are five reasons. First, the presence of most of the underlying moras in (65) is redundant. Only one mora is needed for *booró'okim*, as in (64), and no underlying moras are needed for *naátenake*. This is because all vowels, whether short or long, are moraic at phonetic representation. Since all vowels (except glides, which I discuss below) are distinguishable from consonants by their melodic features alone, it is not necessary to include moras in underlying representation unless, as in *booró'okim*, a single long vowel contrasts with short vowels.[10]

The second reason for selecting (64) in favor of (65) comes from Hayes (1989), which points out that some consonants have an underlying mora even though they cannot be syllabic. Thus, the presence of an underlying mora is not what qualifies a segment as syllabic.

The third reason is that the application of the syllabification algorithm (66) to each of the representations in (64) yields the correct results. In contrast, the application of syllabification to (65) yields extra-long vowels which do not violate any of the rules or constraints discussed thus far, yet which are unattested. This is illustrated in (72).

72. Application of syllabification and Weight-by-Position to (65)(a):



The fourth reason for selecting (64) in favor of (65) is distributional: Most of the logically possible sequences of non-identical vowels are attested in my data; a few examples are listed in (73).

73. iebu 'last year' siali 'green' temáe 'ask' hiókore 'forgive'

púa 'choose' éaka 'thinking' hóa 'house' haiti 'nauseated'

Since sequences of non-identical vowels can occur, it seems reasonable to expect sequences of identical vowels as well. *Naátenake* is one example of such a sequence.

The fifth and final reason for selecting (64) in favor of (65) is that the latter predicts the three-way contrast represented in (74). Such a contrast is unattested for non-high vowels.

74.
```
µ      µ µ
|      |/
R      R      R
```

The three-way contrast in (74) *is* attested (in phonetic representation) for [+high] vowels. This is because Mayo syllabification distinguishes between glides and [+high] vowels. This is illustrated in (75) through (77); in each case, a glide triggers epenthesis when it is unable to syllabify with an existing vowel.

75. Mexico-w   *Mexico-iw   'Mexico-toward'

76. Los Angeles-iw   *Los Angelesu   'Los Angeles-toward'

77. túuruy 'thick' + -mme 'they' —> túu.ru.yim.me   *túu.ru.im.me

In (75), [-w] 'toward' becomes a coda. In (76), the same morpheme is in a position where it would be expected to syllabify as a nucleus, but it does not. Instead, it triggers epenthesis and becomes a coda. The same is true of the final glide in *túuruy* 'thick'.

Mayo is not the only language that distinguishes between glides and high vowels. Guerssel (1986) describes a similar phenomenon in Berber. Hayes (1989) suggests that such a contrast makes it necessary to represent underlying vowel length in Berber with two moras; this allows for a short high vowel to be represented with a single underlying mora so as to be distinguishable from a glide, which has no underlying mora. However, I claim that the RT approach to vowel length allows the needed three-way distinction without requiring a single root node to have more than one underlying mora. The key is to represent a long high vowel as a sequence of two root nodes, each with its own mora. The only added stipulation is that high vowels do not undergo Weight-by-Position. This is in fact a falsifiable prediction, sketched out in (78).

78. <u>Prediction</u>:

```
                        σ                         * σ
                        |                           |\
           μ            μ                           μ μ
           |            |                           |/
           R    —>      R    —>     Weight-         R
           |            |         by-Position       |
       [+high]      [+high]          N/A         [+high]
```

The above prediction is testable in Mayo because, when an underlyingly long vowel is of the MT type (i.e., the two moras are tautosyllabic), the long vowel undergoes shortening in a closed syllable. This is illustrated in

(79); notice that the vowel length in (a) must be underived because the environment for MI is not met. The absence of vowel length in (b) proves that the two moras of the long vowel in (a) are tautosyllabic.

79. a. béeb-a-ka 'hit-PRES-PART' b. béb-la 'hit-PERFECT'

The shortening in (79)(b) can be accounted for in terms of a constraint upon the maximum size of a syllable in Mayo. Many languages have this same bimoraic syllable template. For example, Archangeli (1989) argues that Yawelmani has syllables of the form CVC and CVV, but none of the form CVVC. Whenever a long vowel occurs in a closed syllable, the vowel shortens just as (79) demonstrates for Mayo.

This shortening of a long vowel can be used as a diagnostic to test the prediction made in (78). The claim of (78) is that the MT type of vowel length does not exist for high vowels. Only the RT type of length is contrastive in high vowels because, according to (78), Weight-by-Position does not apply to them. Therefore, if the long vowel in (79) were high, shortening would be predicted not to occur because high vowels can be underlyingly long only if they have two root nodes. In other words, the two moras of a long high vowel have to be in separate syllables; therefore they are 'immune' to Shortening. I know of no counter-examples in Mayo, but the discovery of a counter-example would falsify my claim.

In the preceding section I gave examples of both types of 'vowel length' in unaccented words, but (79) is the only example presented thus far of an accented word which is unambiguously identifiable as to which kind of 'vowel length' it possesses. (79) is an example of the MT type of length in an accented word, but (80) and (81) are unequivocal examples of the RT type of length. If these two words had the MT type of length instead of the RT type, their 'long vowels' would shorten as in (79). Since they do not shorten, the 'long vowels' are actually sequences of identical short vowels which belong to separate syllables. In other words, these words possess the RT type of underlying 'vowel length'.

80. káakte'era 'raven'

81. sóok-te 'separate-CAUSE'

In summary, I have shown that Mayo distinguishes between a single underlyingly long vowel (MT approach) versus a sequence of two identical short vowels (RT approach). This underlying contrast can be represented without invoking any additional structure or rules beyond that which is needed for independent reasons. These two types of 'vowel length' are represented in (82).

```
82.   Moraic Tier:                          Root Tier:

          μ   μ                                R   R
           \ /                                  \ /
            R                                    P

  Hyman 1985, McCarthy & Prince 1986        Selkirk 1988
```

3. Conclusion.

It has been the goal of this investigation to explore the empirical claims, as applied to Mayo, of two different approaches to the representation of underlying vowel length. Through an examination of the prosodic rules of Mora Insertion, Phrase-Final Invisibility, stress assignment, Leftward Spread, Vowel Copying and specific syllabification processes, I have demonstrated that Mayo has two types of underlying vowel length which contrast in the manner in which they syllabify. One type of long vowel spans two syllables in phonetic representation, while the other type is entirely contained within a single syllable. This contrast is accounted for, without requiring any *ad hoc* devices, by assuming that Mayo represents underlying temporal contrasts sometimes on the root tier and sometimes on the moraic tier. A number of predictions stem from this assumption, and all of them seem to be instantiated in Mayo.

The theoretical implication is that Pike's (1947) approach to underlying vowel length is essentially correct: In some languages, an underlyingly long vowel is best analyzed as a sequence of two identical vowels, while in other languages it has to be treated as a single contrastively long vowel. A third logical possibility, given the first two, is that a single language might make use of both types of vowel length. Mayo is an example of such a language.

Endnotes

*Thanks to Diana Archangeli, Sandra Fulmer, Michael Hammond, Long Peng, and Wendy Wiswall for their many helpful comments and questions.

[1]Hayes (1989) suggests that underlying vowel length must be represented with two underlying moras in a language like Berber, where glides contrast with high vowels. Under that version of the Moraic Tier approach, short vowels have one underlying mora and long vowels have two; glides have none. However, I argue in section 2 that a single underlying mora is sufficient to represent vowel length even in languages like Berber.

[2]Actually, both the RT approach and the MT approach are represented in early literature. For example, Pike (1947, page 138) claims that underlying vowel length is best analyzed as a sequence of two short vowels in some languages and as a single contrastively long vowel in other languages. His criterion for distinguishing between the two is based on the overall segmental distribution of the language under consideration: If non-identical vowel sequences are found, then the long vowel is interpreted as a sequence

of two short vowels; otherwise, it is a single vowel which contrasts with short vowels.

[3]The sequence *ka-tím* comprises a single phonological word, as evidenced by the fact that it bears its own primary stress just like the verb *kó'okore* that follows it. Every Mayo word has one and only one stress, which is phonetically realized as high pitch. Secondary stress occurs only in compounds, where each of the two members is capable of occuring alone with primary stress. There are a few function words, such as *ka* 'no' and some of the pronoun sets, which cannot occur in isolation but which nevertheless serve as a domain for primary stress.

[4]An alternative would be to combine the two processes by simply saying that a vowel lengthens in a monomoraic word. However, Hagberg (1989 b) demonstrates that Mora Insertion accounts, in part, for onset gemination in certain environments as well as vowel lengthening in other environments. In order to capture this generalization, Mora Insertion needs to be stated independently of the two processes which it feeds.

[5]The absence of underlying vowel length in monosyllabic forms might be related historically to the synchronic rule of Mora Insertion. As Hagberg (1989 c) demonstrates, this rule applies prior to most affixation processes, so it is conceivable that underlying vowel length arose as a result of the lexicalization of certain morphological processes in certain words. For example *naáte* 'begin', which has an underlyingly long vowel in modern Mayo, may have resulted from the combination of a hypothetical stem of the form *na* plus *-te*, which is still productive as a causative suffix. If the Mora Insertion rule existed at that time and had the same domain as it has now, then it would have applied to *na* prior to affixation. If the stem and

suffix became inseparable through a diachronic process of lexicalization, then the presence of underlying vowel length in *naáte* would be explained. If this is correct, it would also explain the absence of underlying vowel length in monosyllabic stems.

[6]Except where otherwise noted, all data presented in this paper (including all the preceding examples) are in the 'elsewhere' (non phrase-final) form.

[7]Strictly speaking, an onset is not in the same environment as a syllable nucleus. In what follows, however, I show that MI in fact applies to accented words in exactly the same way as it applies to unaccented words. The difference in final outputs is due to the subsequent application of independent principles.

[8]This may not be true for all languages, in which case the constraint against consonants bearing stress would be reduced to a default parameter setting in Universal Grammar.

[9]Actually, I have demonstrated the validity of (60)(a) for unaccented words only. In the next section I show that (60)(a) holds for accented words as well.

[10]Thanks to Diana Archangeli for pointing this out to me.

References

Archangeli, D. (1989) 'Prosodic syllabification and Yawelmani "CVVC" syllables'. ms. University of Arizona.

Collard, H. and E. Collard (1962) *Vocabulario Mayo*. Serie de Vocabularios Indigenas Mariano y Silva Aceves, No. 6. Instituto Lingüistico de Verano: México, D.F.

Guerssel, M. (1986) 'Glides in Berber and syllabicity', *Linguistic Inquiry* 17:1-12.

Hagberg, L. (1988) 'Stress and length in Mayo', in W. Shipley, ed., *In honor of Mary Haas*. Mouton de Gruyter: Berlin, 361-75.

_____ (1989 a) 'Floating accent in Mayo', Arizona Phonology Conference II presentation to appear in *Proceedings*.

_____ (1989 b) 'Mayo reduplication', ms. University of Arizona.

_____ (1989 c) 'Stem, word and phrase as phonological domains in Mayo', ms. University of Arizona.

Halle, M. and J. Vergnaud (1987 a) *An essay on stress*. MIT: Cambridge.

_____ (1987 b) 'Stress and the cycle', *Linguistic Inquiry* 18:45-84.

Hayes, B. (1989) 'Compensatory lengthening in moraic phonology', *Linguistic Inquiry* 20:253-306.

Hyman, L. (1985) *A theory of phonological weight*. Foris: Dordrecht.

Inkelas, S. (1989) *Prosodic constituency in the lexicon*. Ph.D. dissertation, Stanford University, Palo Alto, CA.

Ito, J. (1986) *Syllable theory in prosodic phonology*. Ph.D. dissertation, University of Massachusetts, Amherst.

_____ (1989) 'A prosodic theory of epenthesis', *Natural Language and Linguistic Theory* 7.

McCarthy, J. and A. Prince (1986) 'Prosodic morphology', ms. University of Massachusetts, Amherst, and Brandeis University, Waltham, MA.

Pike, K. (1947) *Phonemics: a technique for reducing languages to writing*. The University of Michigan Press: Ann Arbor.

Selkirk, E. (1982) 'The syllable', in H. van der Hulst and N. Smith, eds., *The structure of phonological representations*. Foris: Dordrecht, 337-384.

_____ (1988) 'A two-root theory of length', NELS 19.