Mapping Phonological Structure to Phonetic Timing: Moras and Duration in Two Bantu Languages
Author(s): Kathleen A. Hubbard

Please see “How to cite” in the online sidebar for full citation information.

Please contact BLS regarding any further use of this work. BLS retains copyright for both print and screen forms of the publication. BLS may be contacted via http://linguistics.berkeley.edu/bls/.

The Annual Proceedings of the Berkeley Linguistics Society is published online via eLanguage, the Linguistic Society of America's digital publishing platform.
Mapping Phonological Structure to Phonetic Timing: Moras and Duration in Two Bantu Languages
Kathleen A. Hubbard
U.C. Berkeley

The mora has a number of phonological roles: it is sometimes used as a unit of weight, sometimes as a unit of quantity (length), and sometimes as a unit of timing. In this paper I ask whether the phonological unit ‘mora’ has any phonetic reality, and I conclude that at least for the languages under examination, the answer is yes. First I review the role of the mora as a unit of phonological representation, and then I look to see how such representations are reflected in phonetic output. For this purpose I present data from two Bantu languages (a family that figures prominently in the development of moraic theory, because of its rich prosodic structure involving tone, length distinctions, and quantity-sensitive phonological and morphological rules). Finally, I outline a theory of mapping between the underlying timing system of a mora-timed language and the surface output observed in speech.

Timing-tier theory arose because linear segmental representation of phonological structure was insufficient to account for prosodic phenomena such as light vs. heavy syllable distinctions, short vs. long segments, compensatory lengthening, stress and tone assignment, etc. (Ingría 1980, Hyman 1984 and 1985, etc.). In the timing model in (1), proposed by Clements and Keyser (1983), syllables do not directly dominate segments but instead are built on timing units (shown here as C’s and V’s).

(1)

\[ \sigma \\
C \quad V \quad C \\
| \quad | \\
p \quad i \quad t \\

A major insight of this model is that long segments involve a single segmental feature matrix linked to multiple timing units, and that phonological processes can affect either just the segmental or just the CV tier. Thus compensatory lengthening (as formulated in Clements 1986) involves reassociation of a segment from its own V unit to the following C, and spreading of the preceding vowel to fill the vacated timing slot:

(2)

\[ C \quad V \quad V \quad C \quad V \\
| \quad | \quad | \quad | \\
t \quad a \quad n \quad d \quad a \\
C \quad V \quad V \quad C \quad V \\
| \quad | \quad | \\
t \quad a \quad n \quad d \quad a \\
C \quad V \quad V \quad C \quad V \\
| \quad \check{\beta} \quad | \\
t \quad a \quad n \quad d \quad a \\

underlying representation

Prenasalization

Linking Convention
The notion of compensatory lengthening as maintenance of syllable timing is phonological, not phonetic. It is of course related to phonetic duration, and makes predictions about relative duration: all other things being equal, a segment linked to two timing units will have greater duration than a segment linked to just one. But it is emphasized that the mapping between the discrete phonological timing units used here and the nondiscrete relations of phonetic duration is as yet poorly understood, and thus the CV model should not be seen as one that directly encodes durations in milliseconds. This disclaimer about the phonetic predictions made by duration-based phonological representations persists into later versions of timing-tier theory as well.

Moraic theory evolved because CV theory did not sufficiently address the issue of phonological weight (Hyman 1985). In this theory, the mora is the unit that is relevant for syllable weight, vowel length, tone assignment, etc. Instead of mediating syllabic structure and segmental content through a tier that assigns one timing unit to each segment, moraic theory assumes that timing has more to do with the bearers of weight, stress, and tone. As shown in (3), moras are what dominate segments and are dominated by syllables (Hyman 1984 and 1985, McCarthy and Prince 1986, Hayes 1989). There is some disagreement about the details of this representation, but they are not crucial to this paper. What everyone agrees on is that a light syllable has one mora, a heavy syllable two. (In some languages both CVV and CVC are bimoraic (e.g. Latin, Hausa), while in others only CVV is bimoraic (i.e. CVC counts as light, as in Lardil and Huasteco).)

(3)

\[
\begin{align*}
\text{a.} & \quad \sigma \\
& \quad \mu \\
& \quad t \ a
\end{align*}
\]

\[
\begin{align*}
\text{b.} & \quad \sigma \\
& \quad \mu \ \mu \\
& \quad t \ a \ a
\end{align*}
\]

\[
\begin{align*}
\text{c.} & \quad \sigma \\
& \quad \mu \ \mu \\
& \quad t \ a \ t
\end{align*}
\]

\[
\begin{align*}
\text{d.} & \quad \sigma \\
& \quad \mu \\
& \quad t \ a \ t
\end{align*}
\]

This model represents more straightforwardly certain phenomena that had to be stipulated in CV theory: e.g., onsets never contribute weight while codas may, and segment deletion only triggers compensatory lengthening if it occurs in the syllable rhyme. The moraic model is also supported by a wide range of historical compensatory lengthening processes (Hock 1986).

But what is the relationship of such a representation to phonetic realization? The durational issue that the moraic model hints at but does not explicitly address is *isochrony*: the usual notion of isochrony (as discussed in Lehiste 1970) is that certain events occur at regular intervals in the speech stream (this is what we hear as the rhythmic aspect of speech). It has been claimed that this regular event is main word stress in English, the syllable boundary in French, and the mora in Japanese — in other words, there is supposed to be in English a regular durational interval between main stresses (regardless of how many syllables intervene), while in French each syllable takes up roughly the same durational span, and in Japanese the regular durational unit is the mora (constituted by CV, moraic N, or the first half of a geminate C). Thus these
languages are described as stress-timed, syllable-timed, and mora-timed, respectively.

But phoneticians have long since demonstrated that this notion is incorrect— at least insofar as the acoustic reflexes of stress and syllable/mora boundaries indicate. It is simply not the case that every main stress in English comes along with metronome-like regularity, and that the duration of intervening syllables is adjusted to compensate. There is some evidence that isochrony does exist perceptually, and some articulatory data that suggest why: the action of certain muscles is isochronous, but the execution of the articulation takes varying lengths of time to achieve depending on the nature of the segment (Fischer-Jørgensen 1964 [cited in Lehiste 1970], Lehiste 1984). So perhaps as listeners we correct for the differing mobility of the various articulators; that is, we extract from a non-isochronous acoustic signal a percept of isochronous articulatory triggers (though it is not clear how). In this paper I do not address the perceptual or articulatory basis for moraic timing, but I mention isochrony because it seems to be at the root of the notion “mora”: since we know that for phonological purposes CV = V = CVC in some languages, and in others CVV =VV = CVC, it is natural to ask whether this relationship holds in phonetic realization or whether it is strictly an abstract fact.

The present phonetic study examines the phonetic correlation between moraic structure and surface duration in Runyambo (a Bantu language of Tanzania) and Luganda (a closely related language of Uganda). Earlier studies of this relationship are mostly on Japanese, where the mora has a long history in traditional linguistics: it is said to be an isochronous timing unit, such that a CV syllable, a moraic nasal, and the first half of a geminate consonant are supposed to take up the same amount of time. The mora is a unit that all literate speakers are conscious of, since the kana writing system and the long poetic tradition both make use of it (and thus it is taught in school). Phonetic studies such as Beckman (1982), however, claim that the mora cannot be the kind of unit it is traditionally thought to be: for one thing, the inherent durations of different segments cause sequences that count as a mora phonologically to have very different durations. Further evidence in Beckman’s study against the mora as a constant unit of phonetic timing comes from the behavior of geminates and of devoiced (or deleted) high vowels.

But a later phonetic study of Japanese, Port, Dalby and O’Dell (1987), shows that looking just at long consonants or devoiced-vowel syllables or plain CV’s in isolation restricts one’s view to a stretch of speech that is too narrow for investigating moraic timing. Port et al. conclude that the mora is in fact a phonetically real timing unit in Japanese, but that segmental adjustment (compensation) to maintain mora count is done at the level of the word. That is, if you extend a sequence of segments by units that are phonologically one mora at a time, the duration of the word increases by roughly constant increments. Likewise, all phonologically-three-mora words fall into the same range of duration, regardless of syllable structure or segmental content, as do all four-mora words, five-mora words, etc.

This is precisely the conclusion I came to in an earlier study of Runyambo syllable timing (Hubbard 1992), rather by accident. In looking at the timing of prenasalized consonants and compensatorily lengthened vowels, I discovered that my corpus of words sorted neatly into groups by total word duration, and that those groups corresponded to phonological mora count. Since those tokens were
not balanced for the counting of moras and syllables, I performed a new experiment that was more carefully controlled to look for exactly this effect.

Bantu languages provide a natural way to extend a sound sequence by one mora or syllable at a time, namely verb affixes. I elicited a number of Runyanmb verb infinitives with different numbers of prefixes and suffixes, as well as roots of varying shapes. These included long vowels, vowel-nasal-consonant sequences, and plain CV sequences, such that mora count and syllable count matched in some cases and not in others. The corpus is given in (4):

(4)

roots + affixes
kunógoora  ‘to mold’ kuguruka  ‘to jump’
kunógoorera  ‘to mold for/at’ kugurucira  ‘to jump for/at’
kujinógoora  ‘to mold it (clay)’ kucigurucira  ‘to jump over it for/at’
kujinógoorera  ‘to mold it for/at’ kucitugurucira  ‘to jump over it for us’
kujeenda  ‘to go’ kukóma  ‘to tie’
kujeendera  ‘to go to’ kukómera  ‘to tie for/at’

roots of minimally different shape
kugoba  ‘reach, arrive’ kukuba  ‘fold’
kugooba  ‘bend’ kukuuba  ‘polish’
kugomba  ‘desire’ kukumba  ‘tilt, fall over’
kusiba  ‘to imprison’ kutana  ‘fester’
kusiiba  ‘pass time at’ kutáana  ‘go separate ways’
kusimba  ‘erect sthg’ kutánga  ‘forbid’
kusáaga  ‘be left over’ kusona  ‘sew’
kusanga  ‘come upon’ kusonda  ‘peck’
kubona  ‘to see’

The hypothesis was that total word duration would correlate closely with mora count, and that as phonological sequences were extended mora by mora, their duration would increase by nearly the same amount each time.

All target words were elicited in the frame “Gamba ______ rumo” (‘say ______ once’), and each word appeared three times on a randomized list. A native speaker of Runyanmb read the list onto tape; the tokens were then digitized at 10kHz and durations were measured from waveforms and wide-band spectrograms. The total length of each utterance was noted (from the release of /g/ in gamba to the onset of /o/ in rumo) so as to control for speech tempo; anomalously long or short tokens were discarded. The duration of each target word was logged and the results were analyzed statistically.

The mean values for each item are given in (5); mean values by number of moras are shown in (6). Analysis of variance to determine the correlation of word duration with mora count is compared with syllable count in (7); though I do not give here the results of all the post-hoc statistical tests, I have determined that they show (like the F-values here) that although syllable count does correlate with word duration, the correlation is much weaker than that with mora count.
(5) Mean Durations of Runyambo Tokens

<table>
<thead>
<tr>
<th>word</th>
<th>moras</th>
<th>sylls</th>
<th>dur</th>
</tr>
</thead>
<tbody>
<tr>
<td>kuguruka</td>
<td>4</td>
<td>4</td>
<td>572</td>
</tr>
<tr>
<td>kugurucira</td>
<td>5</td>
<td>5</td>
<td>677</td>
</tr>
<tr>
<td>kucigurucira</td>
<td>6</td>
<td>6</td>
<td>807</td>
</tr>
<tr>
<td>kucitugurucira</td>
<td>7</td>
<td>7</td>
<td>874</td>
</tr>
<tr>
<td>kunogoora</td>
<td>5</td>
<td>4</td>
<td>675</td>
</tr>
<tr>
<td>kunogoorra</td>
<td>6</td>
<td>5</td>
<td>772</td>
</tr>
<tr>
<td>kujinogoora</td>
<td>6</td>
<td>6</td>
<td>849</td>
</tr>
<tr>
<td>kujitunogoora</td>
<td>8</td>
<td>7</td>
<td>957</td>
</tr>
<tr>
<td>kusaaga</td>
<td>4</td>
<td>3</td>
<td>548</td>
</tr>
<tr>
<td>kusanga</td>
<td>4</td>
<td>3</td>
<td>562</td>
</tr>
<tr>
<td>kusona</td>
<td>3</td>
<td>3</td>
<td>476</td>
</tr>
<tr>
<td>kusonda</td>
<td>4</td>
<td>3</td>
<td>549</td>
</tr>
<tr>
<td>kukuba</td>
<td>3</td>
<td>3</td>
<td>488</td>
</tr>
<tr>
<td>kukuuba</td>
<td>4</td>
<td>3</td>
<td>562</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>word</th>
<th>moras</th>
<th>sylls</th>
<th>dur</th>
</tr>
</thead>
<tbody>
<tr>
<td>kukumba</td>
<td>4</td>
<td>3</td>
<td>575</td>
</tr>
<tr>
<td>kutana</td>
<td>3</td>
<td>3</td>
<td>468</td>
</tr>
<tr>
<td>kutaana</td>
<td>4</td>
<td>3</td>
<td>528</td>
</tr>
<tr>
<td>kutanga</td>
<td>4</td>
<td>3</td>
<td>561</td>
</tr>
<tr>
<td>kusiba</td>
<td>3</td>
<td>3</td>
<td>448</td>
</tr>
<tr>
<td>kusiiba</td>
<td>4</td>
<td>3</td>
<td>528</td>
</tr>
<tr>
<td>kusimba</td>
<td>4</td>
<td>3</td>
<td>546</td>
</tr>
<tr>
<td>kukoma</td>
<td>3</td>
<td>3</td>
<td>484</td>
</tr>
<tr>
<td>kukomera</td>
<td>4</td>
<td>4</td>
<td>581</td>
</tr>
<tr>
<td>kubona</td>
<td>3</td>
<td>3</td>
<td>469</td>
</tr>
<tr>
<td>kugoba</td>
<td>3</td>
<td>3</td>
<td>464</td>
</tr>
<tr>
<td>kugoba</td>
<td>4</td>
<td>3</td>
<td>529</td>
</tr>
<tr>
<td>kugomba</td>
<td>4</td>
<td>3</td>
<td>549</td>
</tr>
<tr>
<td>kujeenda</td>
<td>4</td>
<td>3</td>
<td>513</td>
</tr>
<tr>
<td>kujeendera</td>
<td>5</td>
<td>4</td>
<td>621</td>
</tr>
</tbody>
</table>

(6) Duration by Mora Count

<table>
<thead>
<tr>
<th>moras</th>
<th>duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>476</td>
</tr>
<tr>
<td>4</td>
<td>550</td>
</tr>
<tr>
<td>5</td>
<td>658</td>
</tr>
<tr>
<td>6</td>
<td>786</td>
</tr>
<tr>
<td>7</td>
<td>859</td>
</tr>
<tr>
<td>8</td>
<td>957</td>
</tr>
</tbody>
</table>

Note that the mean difference between categories in (6) is 96 ms; that is the apparent “standard” duration for a mora in Runyambo.

(7) ANOVA’s: Correlation with mora count vs. syllable count

<table>
<thead>
<tr>
<th></th>
<th>F-value</th>
<th>degrees of freedom</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>mora</td>
<td>452.814</td>
<td>5, 79</td>
<td>.0001</td>
</tr>
<tr>
<td>syllable</td>
<td>185.253</td>
<td>4, 80</td>
<td>.0001</td>
</tr>
</tbody>
</table>

The main result is that the total duration of a word in Runyambo is very closely correlated with mora count, much moreso than with syllable count. Though the syllable criterion also gives a statistically significant result, it is much smaller than that of the mora criterion. In addition, we can see that a four-syllable word containing a long vowel, such as kunogoora (675 ms), is much closer in duration to a five-syllable word with all short vowels such as kugurucira (677 ms) than to a four-syllable all-short word such as kuguruka (572 ms). Likewise kukumba (575 ms) is much closer in duration to words like kukomera (581 ms) than to words like kusona (476 ms). This relationship holds throughout the data set.
The same patterns were visible in the data from my pilot study in 1990, even though that experiment was not controlled specifically for this effect. As with Beckman’s (1982) results for Japanese, it is not the case that every phonologically mORAic sequence has the same duration:

(8) \( \text{one mora:} \)
\[
\begin{align*}
/ku/ &= 194 \, \text{ms} \\
/go/ &= 136 \, \text{ms}
\end{align*}
\]
\( \text{two moras:} \)
\[
\begin{align*}
/tana/ &= 351 \, \text{ms} \\
/gom(b)/ &= 300 \, \text{ms}
\end{align*}
\]

But it is the case that words with the same mora count have very comparable durations. From these patterns it is possible to conclude that although Runyambo does not have a rigidly isochronous mora-timing scheme, some kind of segmental compensation is taking place that is sufficient to override large inherent durational differences between segment types and roughly maintain a weight constant.

To extend this finding, I extracted some Luganda utterances from a tape that was made for another study to see if the same sort of mora effect appears in that language. The data set was originally elicited for a study of intonation, but contains a number of tokens much like those I used for the Runyambo study. It includes a number of single-word utterances that differ by the addition of one syllable or mora at a time, as well as a set of two- and three-word utterances. The tokens, produced by a native speaker of Luganda and recorded by Larry Hyman and Ian Maddieson, were elicited in four contexts: first in declarative intonation, then in two tonally different interrogative intonations, and finally in surprise intonation. Of these I used only declaratives. There was no carrier sentence, since the parameter to be measured was the intonation contour over utterances of a specific length. Each subset of four to eight items was read twice in the same order (which probably resulted in some rhythmic and order effects on duration). The tokens I selected are shown in (9):

(9) \( bάl\text{lima}^2 \)
\( bάm\text{ùlima} \)
\( bάl\text{limira} \)
\( bάm\text{ùlimira} \)
\( bάl\text{lima}l\text{ima} \)
\( bάm\text{ùlima}l\text{ima} \)
\( bάl\text{im}l\text{ira}l\text{ima} \)
\( bάm\text{ù}l\text{im}l\text{ira}l\text{ima} \)
\( bάg\text{ùl}\text{ir}\text{ira} \)
\( bάm\text{ùg}\text{ù}l\text{ir}\text{ira} \)
\( bάá\text{g}\text{ùl}\text{ir}\text{ira} \)
\( bάám\text{ù}g\text{ù}l\text{ir}\text{ira} \)
\( t\text{eb}g\text{ùl}\text{ir}\text{ir}\text{a} \)
\( o\text{mul}i\text{mi} \)
\( o\text{mup}ákās\text{i} \)
\( o\text{mu}s\text{èlf}kālē \)
\( o\text{mugōbā} \)
\( m\text{ùlam}u\text{zi}^3 \)
\( m\text{ùngēlēzā} \)

\( \text{‘they are cultivating’} \)
\( \text{‘they are cultivating it’} \)
\( \text{‘they are cultivating for/at’} \)
\( \text{‘they are cultivating it for him’} \)
\( \text{‘they are cultivating here and there’} \)
\( \text{‘they are cultivating it here and there’} \)
\( \text{‘they are cultivating for/at here and there’} \)
\( \text{‘they are cultivating for him here and there’} \)
\( \text{‘they are bribing’} \)
\( \text{‘they are bribing him’} \)
\( \text{‘they bribed’} \)
\( \text{‘they bribed him’} \)
\( \text{‘they are not bribing’} \)
\( \text{‘farmer’} \)
\( \text{‘porter’} \)
\( \text{‘soldier’} \)
\( \text{‘driver’} \)
\( \text{‘judge’} \)
\( \text{‘Englishman’} \)

\( \text{abali\text{mi}} \)
\( \text{abapākās\text{i}} \)
\( \text{abasēlfkālē} \)
\( \text{abagōbā} \)
\( \text{mūsf\text{ì}gile}^4 \)
\( \text{mūvūbūkā} \)

\( \text{‘farmers’} \)
\( \text{‘porters’} \)
\( \text{‘soldiers’} \)
\( \text{‘drivers’} \)
\( \text{‘deputy’} \)
\( \text{‘adolescent’} \)
These tokens were digitized and measured in the same way as before. There were some problems with measurement, including the difficulty of identifying the beginning of a vowel after pause and the end of a vowel before pause; this speaker often has periods of voiceless formant energy at pause boundaries, or trails off at the end of an utterance in voiced but very attenuated vowel offsets. Nonetheless, a number of fairly reliable measurements were obtained, and these durations were statistically analyzed as before.

Mean durations for some words (a representative sample) and for pooled data are shown in (10) and (11). ANOVA’s for correlation of word duration with mora count and syllable count follow in (12); in this case, all post-hoc tests were significant to 99% for mora count, but only some were significant for syllable count (and even those numbers were again much weaker).

(10) Mean Durations of Selected Luganda Verbs in Isolation

<table>
<thead>
<tr>
<th>word</th>
<th>moras</th>
<th>syllables</th>
<th>duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>balima</td>
<td>3</td>
<td>3</td>
<td>426</td>
</tr>
<tr>
<td>bamulimira</td>
<td>4</td>
<td>4</td>
<td>553</td>
</tr>
<tr>
<td>balimira</td>
<td>4</td>
<td>4</td>
<td>538</td>
</tr>
<tr>
<td>bamulimira</td>
<td>5</td>
<td>5</td>
<td>599</td>
</tr>
<tr>
<td>balimaalima</td>
<td>6</td>
<td>5</td>
<td>769</td>
</tr>
<tr>
<td>bamulimaalima</td>
<td>7</td>
<td>6</td>
<td>869</td>
</tr>
<tr>
<td>balimiralimira</td>
<td>7</td>
<td>7</td>
<td>817</td>
</tr>
<tr>
<td>bamulimiralimira</td>
<td>8</td>
<td>8</td>
<td>901</td>
</tr>
<tr>
<td>bagulirira</td>
<td>5</td>
<td>5</td>
<td>610</td>
</tr>
<tr>
<td>baagulirira</td>
<td>6</td>
<td>5</td>
<td>852</td>
</tr>
<tr>
<td>baamugulirira</td>
<td>7</td>
<td>6</td>
<td>914</td>
</tr>
</tbody>
</table>

(11) Duration by Mora Count

<table>
<thead>
<tr>
<th>moras</th>
<th>duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>426</td>
</tr>
<tr>
<td>4</td>
<td>545</td>
</tr>
<tr>
<td>5</td>
<td>604</td>
</tr>
<tr>
<td>6</td>
<td>810</td>
</tr>
<tr>
<td>7</td>
<td>867</td>
</tr>
<tr>
<td>8</td>
<td>901</td>
</tr>
</tbody>
</table>

Here, the mean difference between categories, or apparent moraic constant, is 95 ms -- but this is not as consistent as in Runyambo.

(12) ANOVA’s: Correlation with mora count vs. syllable count

<table>
<thead>
<tr>
<th></th>
<th>F-value</th>
<th>degrees of freedom</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>mora</td>
<td>45.646</td>
<td>5, 97</td>
<td>.0001</td>
</tr>
<tr>
<td>syllable</td>
<td>12.368</td>
<td>5, 105</td>
<td>.0001</td>
</tr>
</tbody>
</table>
As in Runyambo, it appears that in Luganda mora count is a more relevant timing factor than syllable count. The results are not as consistent — standard deviations are greater, and categories are not as distinct — but this is at least partly because the tokens were not elicited in the same way. For example, there is a large durational effect of position within an utterance: a given word is longer in isolation or utterance-finally than utterance-initially, and so on. In Runyambo, all target words occurred in the same position within the utterance, so this effect did not interfere with other timing factors. Luganda timing needs to be explored in greater depth with measurements of more comparable tokens; although it is presumably timed similarly to Runyambo (given similarities of vowel quantity distribution, compensatory lengthening processes, etc.), Hubbard (1992) shows that there are significant differences between the two languages in both phonological representation and surface duration (as seen in compensatory lengthening and tone assignment). So careful examination of both categorial and gradient differences between the timing systems of these languages is called for. Nonetheless, it appears that at least for the two speakers in this experiment, there is a significant effect of mora maintenance and segmental compensation on surface timing. Although syllable boundaries do not line up neatly across different words (as was implicitly claimed in Herbert 1975 for Luganda), there is some higher-order timing pattern such that (for instance) four-mora words containing long consonants like /k/ and /s/ are not radically longer than four-mora words containing short consonants like /r/ and /b/.

(13) Segment durations

<table>
<thead>
<tr>
<th>word</th>
<th>total</th>
<th>C2</th>
<th>V2</th>
<th>C3</th>
</tr>
</thead>
<tbody>
<tr>
<td>kusona</td>
<td>476</td>
<td>107</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>kubona</td>
<td>469</td>
<td>51</td>
<td>98</td>
<td>72</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>word</th>
<th>total</th>
<th>C2</th>
<th>V2</th>
<th>C3</th>
<th>C4</th>
</tr>
</thead>
<tbody>
<tr>
<td>kusanga</td>
<td>562</td>
<td>125</td>
<td>133</td>
<td>78</td>
<td>13</td>
</tr>
<tr>
<td>kutanga</td>
<td>561</td>
<td>88</td>
<td>167</td>
<td>76</td>
<td>15</td>
</tr>
</tbody>
</table>

Though I have not yet determined exactly how this compensation is achieved, it appears that vowels adjust around the inherent durations of consonants to maintain approximately constant duration for words of n moras. The natural next step in this research is to find out exactly where this compensation occurs (i.e. at what point within the syllable and word), and whether it indeed occurs only in vowels. Meanwhile, we can take this phonetic result as evidence that moraic representation does have a systematic relation to physical output, and thus we can use phonetic data as evidence for and against specific versions of phonological representations.

One obstacle in testing such representations, though, is the lack of a comprehensive model for how phonological timing gets mapped to surface realization. Models for other aspects of the mapping between phonological structure and phonetic output have been developed recently; see among others Beckman and Pierrehumbert (1986), Browman and Goldstein (1986), Keating (1988). These have dealt primarily with intonation, coarticulation, and the nature of phonetic targets and the trajectories between them. I assume that the same principles apply to timing -- specifically, that the path from abstract phonological structure to speech production looks something like (14) (the schematic is taken from Cohn (1990)): 
phonological rules
↓
language-specific phonetic rules
↓
universal phonetic implementation rules
↓
physical output

Given this model, we can assume that there is a point at the output of the phonological derivation where duration is assigned to segments and syllables. In a language like English, a chief priority at this point will be assigning greater duration to stressed syllables. In a language like Runyambo, the highest priority will be to assign duration to segments dominated by a mora: these are the ones that determine distinctive quantity, and thus are subject to the constraint of maintaining a distinctive durational contrast, while other segments may receive their duration specification solely from other sources (such as their place and manner features).

There is some evidence for a timing model like this, from research in speech synthesis. Campbell and Isard (1991), in an effort to find a workable algorithm for calculating duration in synthesized speech, examined natural speech corpora of British English and concluded that the syllable was the most relevant unit of programming for duration (not the segment or anything else). Thus they consider the following factors in their model: (1) number of segments in the syllable, (2) nature of the syllable nucleus (tense/lax vowel, sonorant consonant, etc.), (3) position of the syllable in the foot, (4) position of the syllable in the phrase and clause, (5) stress, and (6) the function vs. content role of the word. The first thing to be calculated is not segment duration but syllable duration, taking into account stress, phrasal position, etc. and leaving out phonetic detail. Then the appropriate durations for individual segments are computed within that syllable span, considering factors such as inherent durations, the nature of pre-boundary lengthening (whether it applies to all segments in a syllable or just the last ones), etc.

Based on my measurements of Runyambo and Luganda, I conclude that a similar process takes place in natural language in the mapping of phonological timing to phonetic duration. This means for a general theory of timing we would want to take into account at least the following:
Does the language make phonological reference to moras?
   Yes: assign minimum durations to maintain quantity distinction
   No: next step

Calculate appropriate duration for the syllable:
   how many syllables in phonological word?
   how many words in phrase?
   position in phrase?
   location of stress?

Calculate appropriate duration for each segment:
   feature specification?
   specifications of neighboring segments?

At this stage in the computation of timing, language-specific allophonic rules would apply, and their output would be subject to universal physiological constraints on relative timing. This means that a number of factors may influence the final realization of duration after the end of the phonological derivation. But this does not necessarily obscure underlying relationships; indeed in languages with distinctive quantity we know that it does not.

Clearly, more phonetic data must be gathered before we can fill in the details of such a timing model. Articulatory studies can tell us more about universal constraints as well as language-specific patterns; perceptual studies can help us determine the degree of "slack" there is in how accurately timing relations must be executed to maintain necessary contrasts. And, more fundamentally, further acoustic studies on more languages will provide the basic data on duration that will allow us to compare surface timing with phonological structure. Then it will be possible to determine the true role of the mora.

Notes

* The work on Luganda in this paper was supported in part by National Science Foundation Grant BNS91-09234.

1 The /b/ is not included in the measurement; it is shown to indicate that the VN sequence occurred before a stop, the environment in which compensatory lengthening takes place (cf. (3)). I have argued elsewhere (Hubbard 1992) that in Runyambo, this compensatory lengthening does not result in the nasal giving up all of its mora, as it does in Luganda. Rather, as Maddieson (1992) found for Sukuma, the nasal ends up sharing its mora with the preceding vowel. Thus it is not possible in Runyambo to get a "two-mora" measurement just from a compensatorily lengthened vowel; the nasal must also be included.

2 Acute accent marks High tone, grave accent Low tone, unmarked = toneless.

3 Phonetically [m(u)lamudzi]. These four tokens have no augment (initial vowel) because they were elicited after a negative verb, an environment in which nouns lose their augment.

4 Phonetically [mu:ɕire].

5 In Runyambo, for example, this step will assign greater duration to the penultimate syllable of the phonological word (see Hubbard 1992).
References


McCarthy, John J. and Alan S. Prince. 1986. Prosodic Morphology. MS, University of Massachusetts, Amherst and Brandeis University.


hubbard@garnet.berkeley.edu