4 Quantity

STUART DAVIS

1 Introduction

Quantity plays an important role in the analysis of a wide variety of phonological and morphological phenomena in many languages including the analysis of word stress, tone, compensatory lengthening, shortening processes, minimal word requirements, templatic restrictions and allomorphy selection. These phenomena frequently distinguish between syllables that are short (or light) from those that are long (or heavy). In the first section of this chapter we introduce the theory of moraic phonology of Hayes (1989a), a representational theory of quantity. In subsequent sections we overview a number of issues that emerge from moraic theory. In our discussion of these issues we will refer to the wide variety of processes in which quantity plays a role.

In modern studies of phonology, the term “quantity” refers to either segmental duration or syllable weight. With respect to segmental duration, quantity differences among segments are said to be phonemic in languages that contrast a long and short form of a vowel of the same quality and in languages that contrast a geminate versus non-geminate consonant. The Japanese examples in (1) illustrate both types of contrasts.

(1) Japanese quantity contrasts (Tsujimura 2007)
   a. [su] ‘vinegar’
   b. [su:] ‘inhale’
   c. [saka] ‘hill’
   d. [sak:a] ‘author’
Phonetically, though, a single phoneme may be pronounced longer or shorter depending on the nature of the environment in which the segment occurs. For example, cross-linguistically a vowel tends to have longer duration before a voiced consonant than before a voiceless one (Chen 1970) and tends to be longer if it precedes a fricative as opposed to a stop (Peterson and Lehiste 1960; House 1961). However, such phonetic environmental differences are not relevant for processes like stress placement that often distinguish syllables with long vowels. Thus, for instance, we do not find rules of stress assignment that place stress on a syllable containing a vowel preceding a voiced consonant or a fricative. On the other hand, phonemic contrasts in quantity as in (1) often play an important role in the phonology of languages that have such contrasts.

With the advent of generative phonology, the major issue concerning quantity has been the nature of its phonological representation. Chomsky and Halle (1968) use the feature [+long] to characterize segmental quantity. Under such a characterization, the difference between the Japanese words in (1a) and (1b) is that [u] would have the feature [+long] in the latter and [−long] in the former. However, it was subsequently noted (e.g. Leben 1980) that this representation was insufficient since long consonants can behave like a sequence of two segments for certain phenomena. Further, inalterability effects were noted by such researchers as Kenstowicz and Pyle (1973), Schein and Steriade (1986) and Hayes (1986) whereby long segments seemed to be immune to certain phonological processes that shorter segments of the same quality underwent. Such observations motivated an autosegmental representation of segmental quantity, in which long vowels and geminates are linked to two slots on a timing or prosodic tier while a short vowel or singleton consonant is linked to one slot. There has been much literature on the nature of this timing or prosodic tier (see Kenstowicz 1994b; Broseelow 1995, and Hermans 2006 for overviews). A common view of this tier was that it either consisted of CV-slots (e.g. McCarthy 1979a, 1981; Halle and Vergnaud 1980; Clements and Keyser 1982) or X-slots (Levin 1985). This is shown in (2) and (3) below where (2) gives both the CV-tier and X-tier representation of the Japanese word in (1b), and (3) gives the CV-tier and X-tier representation of the Japanese word in (1d).

(2)  a. CV-tier representation of [su:] b. X-tier representation of [su:]

\[
\begin{array}{c|c|c}
\text{C} & \text{V} & \text{u} \\
\hline
s & u
\end{array}
\quad
\begin{array}{c|c|c}
\text{X} & \text{X} & \text{X} \\
\hline
s & u
\end{array}
\]


\[
\begin{array}{c|c|c|c|c}
\text{C} & \text{V} & \text{C} & \text{C} & \text{V} \\
\hline
s & a & k & a
\end{array}
\quad
\begin{array}{c|c|c|c|c}
\text{X} & \text{X} & \text{X} & \text{X} & \text{X} \\
\hline
s & a & k & a
\end{array}
\]
Notice that in (2) and (3) the prosodic shape of the word is encoded segmentally as a sequence of CV-slots or X-slots. Because of this encoding, the CV or X-tier is often referred to as a prosodic tier.

In a highly influential paper, Hayes (1989a) rejected the segmental nature of the prosodic tier and argued instead for its characterization as moraic. Hayes’s main argument against a segmental CV or X-tier is that it fails to properly identify which types of segmental deletions lead to compensatory lengthening. To see what is at issue, consider Turkish compensatory lengthening discussed, by Sezer (1986) and more recently by Hermans (2006). In Turkish, the phoneme /v/ can optionally delete in certain postvocalic environments as shown in (4).

(4) Turkish optional /v/ deletion (Sezer 1986: 228)
   a. [davul] – [daul] ‘drum’
   b. [savmak] – [sa:mak] ‘to get rid of’

In a CV-tier or X-tier representation, there is no non-stipulative explanation for why the deletion of the consonantal phoneme /v/ should lead to compensatory lengthening in (4b) but not in (4a). Under Hayes’s (1989a) theory the difference between examples like those in (4a) and (4b) is that in (4b) the deleted /v/ is a moraic segment in that it is in the coda of the (first) syllable, but in (4a) it is not moraic since it is in the onset of the (second) syllable. Hayes (1989a) makes a strong case that compensatory lengthening involves the loss of a moraic segment without the deletion of the mora. While the notion of mora in current phonological theory goes back at least to Trubetzkoy’s (1939) discussion of syllable quantity, Hayes (1989a) develops a formal theory of moraic phonology in which the prosodic tier is characterized as moraic. Specifically, in Hayes’s theory, a short vowel is underlyingly monomoraic while a long vowel is bimoraic. With respect to geminate consonants, a geminate consonant differs from a short consonant in that the geminate is underlyingly moraic while a short consonant is non-moraic. Sample moraic representations are given in (5) where (5a) shows a short vowel, (5b) a long vowel, (5c) a short consonant, and (5d) a geminate.

(5) Underlying moraic representation (Hayes 1989a)
   a. \( \mu \)
   b. \( \mu \)
   c. \( \mu \)
   d. \( \mu \)

In Hayes’s theory a (non-geminate) coda consonant is not underlyingly moraic. Rather, in some languages a coda consonant acquires moraic status by the rule of Weight-by-Position shown in (6). (We indicate the syllable with the symbol \( \sigma \)).
In (7) we provide examples of syllabification with the surface moraic structure shown.

While the sequence of phonemes in (7c) and (7d) is identical, they differ in that the rule of Weight-by-Position in (6) has applied to (7c) but not to (7d). That is, the coda consonant in (7c) is moraic while the coda in (7d) is not. (7e) shows the mora structure and syllabification of a form with a geminate consonant, which on Hayes's theory is underlyingly moraic. Now, given this view of moraic structure, it becomes quite clear why Turkish /v/ deletion in (4b) triggers compensatory lengthening while that in (4a) does not. In (8) we show the syllabification and moraic structure of the two words in (4) before compensatory lengthening takes place.
In Hayes’s theory, the deletion of /v/ in (8b) as reflected in (4b) results in compensatory lengthening while the deletion of /v/ in (8a) as reflected in (4a) does not trigger it. Assuming that Weight-by-Position applies in Turkish, the deletion of a moraic segment triggers compensatory lengthening in order to preserve the mora while deletion of a non-moraic segment does not result in compensatory lengthening since no mora is involved. (9a) and (9b) show the two Turkish words after /v/ deletion has applied with compensatory lengthening in (9b).

(9) Compensatory lengthening as mora preservation

\[
\begin{array}{ccc}
\sigma & \sigma \\
\mu & \mu & \mu \\
d & a & u & l = [daul] \\
\end{array}
\]

As seen by examining (8) and (9), compensatory lengthening preserves mora structure despite the deletion of the segment. Consequently, Hayes’s theory makes a prediction that it is only the loss of a moraic segment that can lead to compensatory lengthening. To the extent that this is correct, it provides a strong argument for the moraic representation of segmental quantity and for the moraic nature of the prosodic tier.\footnote{As seen by examining (8) and (9), compensatory lengthening preserves mora structure despite the deletion of the segment. Consequently, Hayes’s theory makes a prediction that it is only the loss of a moraic segment that can lead to compensatory lengthening. To the extent that this is correct, it provides a strong argument for the moraic representation of segmental quantity and for the moraic nature of the prosodic tier.}

The moraic representation of segmental quantity in which certain segments are either underlyingly moraic or surface as moraic (through Weight-by-Position) plays an important role in the characterization of syllable quantity (i.e. syllable weight) in current phonological theory. This is because it provides a formal distinction between a light syllable and a heavy syllable. A light syllable is one that is monomoraic whereas a heavy syllable is bimoraic (or greater). Thus, the syllables in (7a) and (7d) are light or monomoraic whereas those in (7b) and (7c) are heavy or bimoraic. (The issue of geminate consonants as in (7e) is ignored here but will be discussed in detail in Section 5.) The difference between light and heavy syllables is of significance for a wide variety of phonological and morphological processes including stress assignment (e.g. languages in which stress picks out heavy syllables), tonal realization (e.g. languages in which contour tones can be realized only on heavy syllables), compensatory lengthening, closed syllable shortening, minimal word effects (e.g. languages that require a bimoraic word minimum), and morphological phenomena such as templatic morphology (McCarthy and Prince 1986; Crowhurst 1991; Broselow 1995), allomorphy selection, and mora augmentation (Davis and Ueda 2006, Fitzgerald forthcoming).

Given the moraic structure shown in (7a–d), we expect to find at least two types of systems with respect to syllable weight in languages in which syllable weight plays a role. These are shown in (10).\footnote{As seen by examining (8) and (9), compensatory lengthening preserves mora structure despite the deletion of the segment. Consequently, Hayes’s theory makes a prediction that it is only the loss of a moraic segment that can lead to compensatory lengthening. To the extent that this is correct, it provides a strong argument for the moraic representation of segmental quantity and for the moraic nature of the prosodic tier.}
Syllable weight systems:

(a) Heavy: CVV
   Light: CV  
   CVC

(b) Heavy: CVV
   Light: CV
   CVC

The system in (10a) reflects languages where Weight-by-Position applies to coda consonants so that both CVV and CVC syllables are treated as heavy. The system in (10b) reflects languages in which Weight-by-Position does not apply so that only CVV syllables are treated as heavy while CVC syllables pattern with light CV syllables. The system in (10a) is shown with its moraic structure in (11), and the system in (10b) is shown with its moraic structure in (12). (We use “c” and “v” below to indicate a consonant phoneme and a vowel phoneme, respectively.)

Moraic structure for Heavy CVV, CVC vs. Light CV

Heavy: CVV
   a. \[ \sigma \]
   \[ \mu \] \[ \mu \] \[ c \] \[ v \]

Light: CV
   b. \[ \sigma \]
   \[ \mu \] \[ \mu \] \[ c \] \[ v \]

Moraic structure for Heavy CVV vs. Light CVC, CV

Heavy: CVV
   a. \[ \sigma \]
   \[ \mu \] \[ \mu \] \[ c \] \[ v \]

Light: CVC
   b. \[ \sigma \]
   \[ \mu \] \[ \mu \] \[ c \] \[ v \]

The system of moraic structure as presented in (11) and (12) reflecting the two types of weight systems illustrated in (10) is a commonly accepted view of moraic representation and syllable weight. Much current work on syllable quantity takes (10)–(12) as a starting point even if to argue against moraic representations or specific aspects of moraic theory. In the remainder of this chapter, we will take (10)–(12) as a basis for an overview of various issues and elaborations that have been brought up in the literature on the moraic theory of quantity since Hayes’s seminal article on the topic, such as conditions on coda weight, the possibility of more than two degrees of syllable weight, the apparent inertness of syllable onsets to syllable weight, the underlying moraicity of geminate consonants, and weight inconsistency whereby, for example, within one language CVC syllables can act as both heavy and light. The conclusion will include a brief discussion of the phonetics of syllable weight.
2 Coda Weight

In the moraic theory that is presented in (10)–(12), no distinction is made between coda consonants of different types: either codas are moraic (11b) or not (12b). However, work by Zec (1988, 1995b) that harks back to Trubetzkoy (1939) pp. 170–171 (pagination as in Baltaxe 1969) posits that languages can impose a minimal sonority threshold on what can be a moraic segment. In support of this, Zec presents evidence from a variety of languages in which a CVC syllable closed by a sonorant consonant acts as heavy while a CVC syllable closed by an obstruent acts as light. In these languages there is a minimal sonority threshold for moraicity: only vowels and sonorant consonants can be moraic. Consequently, Zec argues that in addition to the weight systems given in (10) some languages have the weight system in (13).

(13) Syllable weight distinction based on sonority ($S =$ sonorant consonant and $O =$ obstruent)

<table>
<thead>
<tr>
<th></th>
<th>Heavy</th>
<th>Light</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVV</td>
<td>CV</td>
<td></td>
</tr>
<tr>
<td>CVS</td>
<td>CVO</td>
<td></td>
</tr>
</tbody>
</table>

Zec cites Lithuanian as an example of a language that has the syllable weight system in (13). She gives two pieces of evidence in support of this division. In (14), we show Zec’s (1995) examples of a morphological ablaut process in Lithuanian that has the effect of lengthening the vowel of the root in certain forms of the verb. (The double vowel representation of length in (14) follows Zec.)

(14) Verbal ablaut in Lithuanian (Zec 1995b: 100–102)

<table>
<thead>
<tr>
<th>Verb stem</th>
<th>Infinitive</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. tup</td>
<td>tuup-ti</td>
<td>‘perch’</td>
</tr>
<tr>
<td>b. dreb</td>
<td>dreeb-ti</td>
<td>‘splash’</td>
</tr>
<tr>
<td>c. vag</td>
<td>voog-ti</td>
<td>‘steal’</td>
</tr>
<tr>
<td>d. vir</td>
<td>vir-ti (*)</td>
<td>‘boil’</td>
</tr>
<tr>
<td>e. mir</td>
<td>mir-ti (*)</td>
<td>‘die’</td>
</tr>
<tr>
<td>f. kar</td>
<td>kar-ti (*)</td>
<td>‘hang’</td>
</tr>
<tr>
<td>g. kau</td>
<td>kau-ti</td>
<td>‘beat’</td>
</tr>
</tbody>
</table>

Ablaut applies to the infinitive forms in (14a–c) lengthening the root vowel. Vowel lengthening does not occur in (14d–g). The verb stems in (14) are similar to one another in that they all consist of a single closed syllable (assuming that the off-glide in (14g) also closes the syllable and functions as a coda). The key difference between (14a–c) and (14d–g) is that the stem syllable ends in a sonorant in the latter but in an obstruent in the former. If a sonorant consonant in the stem syllable is viewed as being moraic then the lack of vowel lengthening in (14d–g)
can be understood as the avoidance of trimoraic syllables. In (14a–c) the stem-final obstruent is not moraic so the vowel can lengthen without creating a trimoraic syllable. Independent support for this moraic difference between coda sonorants and obstruents comes from a restriction on the occurrence of the circumflex accent in Lithuanian. The circumflex accent is an accent type in Lithuanian that is associated with a rising tone. Zec (1988, 1995b) observes that this accent type only occurs on syllables with a long vowel or closed by a sonorant consonant. It does not occur on a syllable closed by an obstruent. This suggests that the circumflex accent can only be realized on a bimoraic syllable, and as Zec (1995b: 98) proposes, Lithuanian has a minimal sonority threshold limiting moraic elements to those that are [+sonorant]; a coda obstruent does not count as moraic.

Zec’s discussion of Lithuanian seems to support the weight distinction in (13) between sonorant and obstruent codas, but several issues remain unresolved. One important question that Zec (1995b) raises is whether other types of restrictions on coda consonants, referred to as edge constraints by Zec, are reducible to sonority threshold restrictions on moraicity. Zec argues convincingly that these two types of constraints affecting coda consonants are different. This is made clear by her discussion of the Australian language Lardil (Hale 1973). In Lardil, codas are restricted so that only (non-dental) coronal consonants can close a syllable, but the coda does not count as moraic since CVC words do not satisfy a bimoraic minimality requirement. Thus, according to Zec, Lardil exemplifies a language that has edge restrictions on coda consonants, but they are not reducible to restrictions on moraicity. Zec (1995b: 111–112) contrasts Lardil with the Tanoan language Kiowa which limits the coda to the consonants /p t m n l/. This set includes the sonorant consonants plus two obstruents. There are other obstruents in the Kiowa inventory (such as fricatives and dorsal stops) but only /p t/ can occur in coda position among the obstruents. Zec notes that in Kiowa, long vowels shorten in closed syllables regardless of whether the syllable is closed by a sonorant consonant or an obstruent. Assuming that closed syllable shortening occurs to prevent trimoraic syllables, Zec maintains that all codas in Kiowa are moraic, which includes both sonorant and obstruent consonants. Zec thus concludes that in Kiowa all codas can be moraic (i.e. there is no sonority threshold on moraic segments), but there are independent edge constraints that prohibit dorsal consonants and fricatives in the coda. If we bring together Zec’s discussion of Lithuanian with that of Lardil and Kiowa, we see three different examples of the relation between coda restrictions and coda moraicity. In Lithuanian, there are no major restrictions on what can appear as a single coda; however, only sonorant codas are moraic. In Lardil, edge constraints limit the nature of the coda to (non-dental) coronals, but only vowels can be moraic; that is, the sonority threshold prevents consonantal elements from being moraic. In Kiowa, edge constraints allow only /p t m n l/ as codas, but there are no sonority threshold constraints on moras; any consonant (sonorant or obstruent) can constitute a moraic coda. Finally, languages may witness a fourth possibility where there are few or no edge constraints restricting the nature of codas and where there is no sonority threshold on what can be moraic. Latin can be cited as an example since there are no (relevant) edge constraints on what can be a coda and any type of consonant in the coda is moraic.
There is another issue that emerges from Zec’s (1988, 1995b) proposal. Zec (1995b) defines sonority based on the major class features, basically [+consonantal] and [±sonorant]. This gives a division between obstruents, sonorants, and vowels. However, there seem to be languages where, in addition to sonorants, high sonority obstruents are also included in the class of moraic segments. One example is Tiv which is mentioned by Zec (1995b), though not discussed, as having the weight division in (13). However, an examination of Tiv (see Pulleyblank 1988, in particular) reveals that the consonants that can appear in coda position are sonorants and voiced fricatives. While it may be possible to interpret voiced fricatives as sonorant, one can just as well note that since voiced fricatives are the most sonorant of the obstruent types in Tiv, the sonority threshold on what can be moraic in Tiv would just be any phoneme with the sonority value of at least a voiced fricative. While such cases are not discussed by Zec, they are not incompatible with her theory, assuming a finer division of the sonority hierarchy than that allowed by the major class features [+consonantal] and [±sonorant].

What would not be expected under Zec’s theory are languages in which both obstruents and sonorants appear in coda position but where only syllables closed by obstruents make a syllable heavy.

A related issue that can be raised regarding Zec’s proposal is whether there are restrictions on coda moraicity that are not sonority related, a problem highlighted by comparing Japanese and the Australian language Ngalakgan. Japanese obeys the Coda Condition in the sense of Itô (1986), whereby coda consonants must be assimilated or place-linked to the following onset, but where coda consonants are clearly moraic (e.g. Kubozono 1999). It is possible to interpret this as a combination of an edge constraint (namely the Coda Condition) with a sonority threshold that allows any segment type (obstruents included) to be moraic. To put it another way, Japanese would have the weight division in (10a) in which any coda consonant can be moraic along with the independent Coda Condition that bans coda consonants from having their own place features.

In light of this, it is interesting to consider the Australian language Ngalakgan discussed by Baker (1997, 2008) and Davis (2003). Consider the pattern of primary stress reflected by the data in (15) taken from Baker.

(15) Ngalakan

a. ciwi ‘liver’
b. céraTa ‘women’s ceremony’
c. páRamùnu ‘sand goanna’
d. cálapir ‘red ant (species)’
e. kúpuy ‘sweat (n.)’
f. purúTci ‘water python’
g. kipiTkuluc ‘frogmouth (bird)’
h. miRárppu? ‘crab’
i. puTólko? ‘brolga (bird)’
j. macaptúrka ‘plant (sp.)’
k. LáRkurca ‘vine (sp.)’
l. cálpurkic ‘fish (sp.)’
The data in first column show that primary stress in Ngalakgan falls on the leftmost (non-final) heavy syllable; otherwise, it falls on the initial syllable. One way of analyzing these data is to posit that a coda consonant is moraic in making a syllable heavy be it an obstruent (15a–b) or a sonorant (as in 15c–g). The data in the second column (15m–t) show that the leftmost closed syllable (underlined) fails to attract primary stress. However, the coda in the leftmost closed syllable in (15m–t) is one of three types: in (15m–p) the coda is a nasal homorganic to the following onset; in (15q–s) it is the first part of a geminate consonant, and in (15t) the coda is a glottal stop. What these three coda consonant types have in common is that they do not have their own place features: the place features are either shared with the following onset, or, in the case of the glottal stop in (15t), there is a lack of place features altogether. If we assume that the closed syllables in (15a–g) attract stress because they are bimoraic, then it would seem that Ngalakgan has a restriction that requires moraic elements to have independent place features (i.e. not shared with a following onset). Thus, while a wide variety of segment types can appear in the coda in Ngalakgan, only those having independent place features surface as moraic (but see Baker (2008) for a different interpretation). The Ngalakgan case as analyzed here would be problematic for Zec’s theory since it entails a restriction on moraicity that is not sonority related. (See Ní Chiosáin (1990) for discussion of western Irish wherein codas with shared place features do not act as moraic.) This suggests that the relationship between coda restrictions and coda moraicity is still in need of further study.15

3 Multiple degrees of weight

In the discussion presented so far, there has been a tacit assumption that there are only two degrees of syllable quantity: light / monomoraic vs. heavy / bimoraic, though languages may differ along the lines of (10) and (13) as to what is considered light and heavy. We have also mentioned the lack of lengthening in Lithuanian CVS syllables (14d–g) and closed syllable shortening in Kiowa whereby potentially trimoraic syllables consisting of a long vowel and moraic coda are avoided. The avoidance of potentially trimoraic syllables is a rather common phenomenon cross-linguistically. Consider the example in (16) from two different dialects of Arabic where a variety of research has shown that (non-final) coda consonants pattern as moraic (e.g. Kiparsky 2003b).
Avoidance of trimoraic syllables in Arabic dialects (See Broselow (1992) for an overview.)


b. Meccan Arabic: /baab + na/ – [báa.ba.na] ‘our gate’

As seen in (16) both Cairene and Meccan Arabic avoid the potentially trimoraic parse of the first syllable of /baab + na/ as [báab.na]. The dialects, though, differ in how they avoid the trimoraic parse. While Cairene Arabic favors closed syllable shortening, Meccan Arabic preserves the underlying vowel length by having vowel epenthesis apply between the two consonants to create open syllables, thus avoiding any trimoraic syllable. Examples like Arabic, Lithuanian, and Kiowa suggest that languages tend to avoid having more than the two degrees of syllable weight. This issue was discussed at length by Trubetzkoy (1939) who maintained that almost all cases where a language has been claimed to have more than two degrees of distinctive quantity for the syllable involve effects that are phonetic rather than phonological.

Despite a tendency in many languages to avoid trimoraic syllables, there are some languages that seem to allow them. One example discussed by Hayes (1995a: 276) is the Hindi dialect described by Kelkar (1968). In this dialect, primary stress falls on the rightmost heaviest (non-final) syllable of the word. The syllable types are listed in (17).

(17) Syllable types in Hindi (Kelkar 1968; Hayes 1995a)

a. superheavy: CV:C, CVCC
b. heavy: CV:, CVC
c. light: CV

The syllable types in (17) can be interpreted as reflecting different quantities with superheavy syllables being trimoraic, heavy syllables bimoraic, and light syllables monomoraic. Stress falls on the rightmost (non-final) syllable with the most moras.

In a constraint-based framework, one can view the difference between a language like Hindi allowing trimoraic syllables and Cairene Arabic, which disallows such syllables as the applicability of a constraint against trimoraic syllables, as in Broselow (1992). This constraint would be inviolable in Cairene Arabic but not in the Hindi dialect described by Kelkar (1968). Thus, there do seem to be cases of more than two degrees of syllable weight.

In addition to cases like Hindi where there are trimoraic CV:C and CVCC syllables, one also comes across cases cited in the literature of what look to be multiple degrees of syllable weight based on vowel quality. Some of these systems have been discussed in detail in the stress literature in Optimality Theory by researchers such as Kenstowicz (1994b), Zec (2003), and De Lacy (2002). One example of such a system discussed by both Kenstowicz and Zec is the Finno-Ugric language Mordwin in which stress is sensitive to the height quality of a vowel. Essentially, following the interpretation of the Mordwin data in Zec (2003),
primary stress in Mordwin words falls on the leftmost syllable containing a non-high vowel; if the word has only high vowels, then stress falls on the leftmost (initial) syllable. Thus, stress is attracted to a syllable with a non-high vowel. While one may be tempted to analyze the difference between non-high and high vowels in terms of quantity by referencing moraic structure, for example, assigning two moras to non-high vowels and one mora to high vowels, this difference does not seem to be related to quantity. Thus, these researchers do not analyze the Mordwin system (and other similar systems where stress is sensitive to vowel quality) in terms of quantity distinctions. Rather, they are analyzed in terms of preferred quality distinctions in positions of prominence. Kenstowicz (1994b) makes reference in his analysis of Mordwin to constraints that prefer lower vowels as syllable peaks while Zec (2003) references a constraint on the sonority threshold for the head of a foot that in Mordwin must be [−consonantal], [−high].

One of the most complex stress patterns reported in the literature is that found in Nanti, a Kampa language of Peru discussed by Crowhurst and Michael (2005). In this language both vowel quality and quantity are essential in determining the location of primary stress. Nanti has a syllable weight hierarchy somewhat similar to that of Hindi in (17), but vowel quality can play a role in determining stress among syllables of equal weight. Consider first the Nanti stress data in (18) from Crowhurst and Michael (2005) which illustrate the effect of syllable quantity on stress. In examining the data in (18), keep in mind that when there are no overriding factors, stress in Nanti is iambic and iteratively assigned from the left edge of the word with final syllables normally being extrametrical as in (18a–b). (We do not discuss the issue as to which of the stressed syllables in a word is assigned primary stress. We refer the reader to Crowhurst and Michael for details regarding this and other aspects of the stress system not discussed here.)

(18) Nanti stress (Crowhurst and Michael 2005)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>o.kò.wo.gó.te.ro ‘she harvests it’</td>
</tr>
<tr>
<td>b.</td>
<td>i.rì.pi.rí.ni.te ‘he will sit’</td>
</tr>
<tr>
<td>c.</td>
<td>o.tá.sòŋ.ka.kse.ro ‘she blew on it’</td>
</tr>
<tr>
<td>d.</td>
<td>òŋ. ko.wo.gó.te.ro ‘she will harvest it’</td>
</tr>
<tr>
<td>e.</td>
<td>pìn.kse.ma.wáa.kse.ro ‘you will have listened attentively to it’</td>
</tr>
<tr>
<td>f.</td>
<td>jo.bií.kái.ga.kse ‘they MASC drank’</td>
</tr>
<tr>
<td>g.</td>
<td>núo.ga.kse.ma.pà.ro ‘I will have consumed it’</td>
</tr>
<tr>
<td>h.</td>
<td>i.kà.man.tái.ga.kse.na ‘they MASC told me’</td>
</tr>
<tr>
<td>i.</td>
<td>o.sà.ráan.tái.ga.kse ‘they MASC tore it with a purpose’</td>
</tr>
</tbody>
</table>

The data in (18a–b) reflect the normal pattern of iterative iambic footing assigned from the beginning of the word when there are no overriding factors. (18c–e) show that closed syllables (CVC) attract the stress away from a light CV syllable disrupting the iambic pattern. (The only coda consonant allowed in Nanti is a nasal homorganic to the following onset consonant.) This suggests that CVC syllables are heavier than CV syllables. The data in (18f–g) illustrate the effect of
a CVV syllable (i.e. a syllable with a long vowel or diphthong). Such syllables attract stress away from CV syllables. Further, the data item in (18h) is consistent with CVV syllables being heavier than CVC syllables in that stress falls on the fourth syllable (CVV) rather than on the third syllable (CVC). Finally, the comparison of the third and fourth syllables in (18i) demonstrates that CVVC syllables are heavier than CVV syllables. Taken together, the data in (18) suggest the following “weight” hierarchy in (19) for determining which syllable in the Nanti foot receives a stress.

(19) “Weight” (or strength) hierarchy for determining the stressed syllable in the Nanti foot

CVVC > CVV > CVC > CV

This is an interesting hierarchy because both CVV and CVC would be bimoraic under a conventional view of moraic phonology, but nonetheless, CVV patterns as heavier than CVC. Because of this, Crowhurst and Michael take the position that only syllables with long vowels are bimoraic in Nanti and that the coda is not moraic. To account for the apparent heaviness of CV syllables, Crowhurst and Michael (2005: 57) posit an independent coda strength scale that makes closed syllables (CVC) stronger than CV syllables. Stress is sensitive to this scale as well as to the moraic makeup of the syllable. Consequently, CVV still would be “stronger” than CVC in Nanti since it is bimoraic whereas CVC is monomoraic; a bimoraic syllable is stronger than a monomoraic one.16

In addition, vowel quality plays a role in stress assignment when the syllables in a foot have equal strength. Essentially, when syllables of equal strength are in the same foot, stress falls on the syllable having the lower vowel. This is shown in (20).

(20) Vowel quality effect on Nanti stress (Crowhurst and Michael 2005)

a. à.wo.te.hái.g̣i .ri ‘we approached him/them’
b. nò.g̣i .wo.tâ.kse .ro ‘I placed it (vessel) mouth down’
c. non.kàn.tái.g̣a.kse ‘we will have said’
d. iŋ.ḳên.tâ.kse .ro ‘he will have pierced it (with an arrow)’
e. noo.g̣i .g̣ai .ga .ro ‘we ate it’
f. i ŭ .bī .kái.g̣a .kse ‘they masc will have drunk’

In (20a–b), the first two syllables are light CV; nonetheless, a stress falls on the first syllable rather than the second because in each case the vowel of the first syllable is lower than the vowel of the second syllable. In (20c–d) the first two syllables are CVC; here, in each of these words, a stress is on the second syllable, not on the first, since the vowel of the second syllable is lower than that of the first. In (20e) the first two syllables are CVV. A stress falls on the second syllable in (20e) due to the fact that it has a lower vowel peak than that of the first syllable. The example in (20f) shows the same pattern except that it is the third
and fourth syllables that are of relevance. Crowhurst and Michael account for the effect of vowel quality on the determination of stress by proposing a vowel quality scale in addition to the coda scale. When all else is equal, the quality scale will determine stress going on to the syllable with a lower vowel peak. To sum up the Nanti discussion from Crowhurst and Michael, Nanti only has two degrees of syllable weight: bimoraic syllables with long vowels and monomoraic syllables that can either be CV or CVC. However, other factors such as vowel quality and the presence of a coda consonant give the appearance of a system with more than two degrees of syllable weight. While a separate vowel quality scale can be motivated for other languages like Mordwin discussed above where quality determines stress placement, it remains to be seen whether coda consonants in multiple weight systems such as that in (17) for Hindi can be understood in terms of a separate coda scale and not as moraic weight. The division in Hindi in (17) where CVV and CVC pattern together seems more consistent with an analysis of three degrees of syllable weight as opposed to a coda scale as in Nanti.

4 Syllable onsets and weight

One of the consistent findings in studies on syllable weight that is implicit in Trubetzkoy (1939) is that onset consonants do not play a role in the determination of syllable weight. (We delay the separate issue of initial geminate consonants until Section 5). The onset is irrelevant to most processes that are sensitive to syllable weight. Consider the fairly common process of closed syllable shortening, where a long vowel shortens in a syllable closed by a coda (or moraic) consonant. In this process, a potential trimoraic CVVC syllable becomes bimoraic CVC. Now, if “heavy” or complex onsets contributed a mora to the weight of the syllable, we might expect to find vowel shortening processes where a potential trimoraic CCVV syllable would shorten to bimoraic CCV. However, to my knowledge, such cases of vowel shortening have not been reported in the literature. Further, minimal word effects do not ever seem to treat the onset consonant as moraic (again deferring the issue of initial geminate consonants until Section 5). For example, if a language requires prosodic words to be bimoraic, it does not distinguish between V, CV, and CCV potential words; all would be disallowed. Similarly, we do not find a minimal word pattern requiring that monosyllabic words begin with an onset whereas longer words can begin with a vowel. The length or even the presence of the onset is irrelevant in meeting a minimal word requirement.

Similarly, in tone languages, tone is realized on moraic elements. This can include a coda consonant (in addition to vowels and syllabic consonants). However, tone in the phonemic or contrastive sense (e.g. lexical and morphological tonal melodies) never seems to have consistent realization on an onset, at least there are no clear instances of it (but see Topintzi (2006) for a possible example of a tonal onset in Kpelle). Furthermore, as noted in Section 1, processes of compensatory lengthening are typically not triggered by the loss of an onset consonant,
as in the Turkish example illustrated in (8a). This is consistent with Hayes’s theory that when compensatory lengthening does occur, it involves the loss of a segment that was moraic. Potential counter-examples to this, such as the three cases noted by Kavitskaya (2002), all involve the deletion of a syllable initial sonorant consonant triggering vowel lengthening. As Hayes (1989a: 282) notes, such cases may involve insertion of a vowel before the deletion of the consonant, or perhaps one could suggest that these cases involve a sonorant being incorporated into the nucleus before it deletes. In general, then, in languages that have processes sensitive to syllable weight, the onset never counts as moraic. For example, one can observe a wide range of mora sensitive processes described for Japanese including external phenomena such as speech errors and language games (Kuboizono 1999; Tsujimura 2007) without encountering any phenomena where onset consonants contribute to syllable weight.

Nonetheless, there is one area in the literature regarding the possible participation of onsets in a weight-sensitive process: patterns of stress assignment. Do onset consonants ever play a role in stress processes, and, if so, does that provide evidence that they can be moraic or contribute weight to the syllable? We turn to this question, and suggest that onset consonants are never moraic.

Davis (1985) was probably the first to survey languages reported to have onset-sensitive stress and mentions about a dozen languages. Gordon’s (2005) more recent survey includes thirteen languages reported to have onset-sensitive stress (though see Topintzi (2006) for critical comments on some of his examples as well as some additional cases). From these surveys as well as from other works such as Davis (1988) and Downing (1998), there seem to be only two types of onset-sensitive stress rules. The first type, which is more common, found amongst Australian languages (e.g. Arrernte) and Native American languages (e.g. Banawá), is where stress falls on the initial syllable if the word begins with a consonant and on the second syllable if the word begins with a vowel. The second type is where a phonological feature on an onset consonant is one of several factors in the determination of stress. (Topintzi (2006) considers a third type that combines the two types.) It is clear from these surveys that one never finds a language where stress falls on the syllables containing a complex (branching) onset, analogous to languages that place stress on heavy syllables. That is, there are no languages where CCV syllables are targeted for stress. One can maintain that the two types of onset-sensitive stress systems that do occur do not constitute cases of onset weight. First, consider the Arrernte type stress system where stress falls on the initial syllable if the word begins with a consonant and on the second syllable if the word begins with a vowel. While one may be tempted to suggest that in such languages the onset is moraic and stress falls on the leftmost bimoraic syllable, there are superior ways of analyzing such data without referencing moraicicy. Downing (1998), for one, observes that onsetless syllables are exceptional to a range of different prosodic processes and can be considered as ill-formed syllables. She posits an analysis where there is a misalignment between morphological and prosodic constituents. Thus, a word-initial vowel would not be part of the prosodic word and outside the domain of stress. Other possible analyses
include having a constraint that aligns the left edge of a foot with a consonant (Goedemans 1996, 1998) or having a constraint that aligns the left edge of a stress syllable with a consonant (Topintzi 2006; Hyde 2007a). Any of these processes captures the stress pattern without having to posit a moraic onset and can be justified based on typological data. Thus, the most common type of onset-sensitive stress system that places stress on the initial syllable if the word begins with a consonant and on the second syllable if it begins with a vowel does not suggest that onset consonants are moraic.

With respect to languages in which some phonological feature of an onset consonant plays a role in the determination of stress, Davis (1988) posits that such cases involve rules of stress shift that occur after the stress has been assigned. That is, rules of stress shift (and stress deletion) may reference a feature on an onset. However, this would not necessarily entail that the presence of certain features on an onset consonant can make it moraic. On the other hand, Topintzi (2006) does make such a proposal. She observes that in three of the languages that have onset-sensitive stress, the South American languages Pirahã, Arabela, and Karo, stress can be attracted to a syllable with a voiceless consonant in its onset. In the well-known Pirahã case (Everett and Everett 1984; Everett 1988) primary stress falls on one of the last three syllables of the word, whichever one has a long vowel; if there is more than one with a long vowel or if there is none with a long vowel, stress falls on the rightmost one containing a voiceless consonant in its onset. In Arabela, primary stress on a final syllable will move to the penultimate syllable if the final syllable starts with a voiced consonant and the penultimate with a voiceless one. Finally, in Karo, a word-final stress will move to the penultimate syllable if the final syllable begins with a voiced obstruent and if the penultimate syllable begins with a voiceless obstruent or sonorant, as long as the final syllable does not have high tone, a nasalized vowel, or final sonorant consonant. Topintzi proposes that in all three of these languages voiceless consonants are moraic. Thus, they can attract stress. In support of this view, she notes a parallel with sonorant consonants being moraic in coda position. Sonorant consonants are the preferred coda type, so they can be moraic in coda position without obstruents being moraic. Similarly, voiceless obstruents are the preferred onset consonant type, so they can be moraic in the onset without treating other consonants as moraic in the onset. While this is an interesting proposal, all three of these languages can just as easily be analyzed as involving stress shift or with an onset prominence scale (or with a prominence projection referencing onsets as in Hayes’s (1995a) analysis of Pirahã). Moreover, if voiceless consonants are moraic, we would expect to find languages where stress falls on the syllables in a word containing a voiceless onset, or we might predict a tendency for languages to avoid long vowels in syllables that begin with a voiceless obstruent. Given that such phenomena do not seem to occur, it may be preferable to view the apparent onset-sensitive nature of stress in Pirahã, Arabela, and Karo as involving stress shift or an onset prominence scale, and not reflecting moraic weight in the onset. Consequently, it is still possible to maintain that (non-geminate) consonants in a syllable onset never contribute weight to the syllable.
5 Geminate Weight

No issue in Hayes’s (1989a) proposal regarding underlying moraic structure as in (5), repeated below as (21), has arguably generated as much controversy as his proposal in (21d) that geminate consonants differ from single consonants (21c) in that they are underlyingly moraic.\(^{23}\)

\[(21)\] Underlying moraic representation (Hayes 1989a)

\[\begin{array}{cccc}
\text{a. } & \mu & \text{b. } & \mu \mu \\
\text{a = } & /a/ & \text{a = } & /a:/ \\
\text{c. } & \text{d. } & t & t/
\end{array}\]

Hayes actually does not discuss the implications of the representation in (21d), but it is made clear in subsequent work by Selkirk (1990b) and Tranel (1991), namely that there should be languages having the weight system shown in (22) in which a syllable closed by a geminate and a syllable with a long vowel act as heavy or bimoraic while a CV syllable and a syllable closed by a non-geminate consonant act as light or monomoraic ($G =$ geminate consonant, $C =$ non-geminate consonant).

\[(22)\] Syllable weight distinction based on geminates being underlyingly moraic:

\[\begin{array}{ll}
\text{Heavy} & \text{Light} \\
\text{CVV} & \text{CV} \\
\text{CVG} & \text{CVC}
\end{array}\]

The system in (22) is predicted to occur under Hayes’s theory in any language that allows long vowels and geminate consonants but in which Weight-by-Position does not apply. The moraic representation with syllable structure of (22) is given in (23).

\[(23)\] Surface syllabification of the division in (22)

\[\begin{array}{cccc}
\text{Heavy} & \text{Light} \\
\text{a. } & \sigma & \text{b. } & \sigma \sigma \\
\mu & \mu & \mu & \mu \\
t & a = [\text{ta}:] & t & a = [\text{ta} ta] \\
\text{c. } & \sigma & \text{d. } & \sigma \\
t & a = [\text{ta}] & t & a = [\text{tat}]
\end{array}\]

As seen in (23b), a syllable closed by a geminate is bimoraic while one closed by a singleton consonant (23d) is monomoraic. Although there are many aspects of the geminate controversy that could be considered here, we will first focus our discussion on evidence concerning the division in (22). Specifically, are there
processes that treat CVG syllables as heavy while treating other CVC syllables as light? We will then discuss other evidence for the moraification of geminates in (21d) briefly reviewing evidence from morphology and the behavior of word-initial geminates.

As noted by Selkirk (1990b) and Tranel (1991), if the syllable weight distinction in (22) exists as Hayes’s theory predicts, we would expect to find languages where a syllable closed by a geminate (i.e. the first part of a geminate) acts as heavy while that closed by a non-geminate does not. Sherer (1994) and Davis (1994, 1999a, 2003) show from a variety of processes that there do seem to be languages that make such a weight division. We will take up evidence from closed syllable shortening and stress.24 If we first consider closed syllable shortening, with the weight division in (22) one would expect to find a language where a long vowel shortens in a syllable closed by a geminate but not in one that is closed by a single consonant. Shortening would occur in the potential CVVG syllable in order to avoid a trimoraic syllable while shortening would not occur in CVVC since that would only be bimoraic. Kiparsky (2008c) mentions Swedish as a language where vowel shortening occurs before a geminate but not before a single coda consonant. Another language displaying this pattern of shortening is the Dravidian language Koya discussed by Tyler (1969) and Sherer (1994) as well as by Davis (1999a) which the following discussion is based on. Koya has long vowels, coda consonants, and geminate consonants. There are words in Koya like those in (24a–c) where a long vowel can occur before a coda consonant. Crucially, as Tyler (1969: 6) observes, there are no words in which a long vowel occurs before a geminate. They are always short as in (24d). (All Koya data are cited from Tyler (1969) with the page numbers provided; vowel length is indicated by a colon; the transcription of the vowel quality is phonemicized and does not reflect the precise allophonic variant.)

(24)  a. le:ŋɡa ‘calf’ (p. 11)  c. nɛ:rs ‘learn’ (p. 76)
     b. aːnɛda ‘female’ (p. 8)  d. ett ‘lift’ (p. 76)

Moreover, cases are found where a stem-final long vowel shortens before a suffix beginning with a geminate, as in (25).

(25)  a. keː + tt + oːndu [kettoːndu] ‘he told’ (p. 39)
     b. oː + tt + oːndu [ottoːndu] ‘he bought’ (p. 38)

This shortening can be viewed as a way of avoiding trimoraic syllables. Shortening does not occur before a non-geminate consonant as the examples in (26) illustrate.

(26)  a. naːl + ke [naːlke] ‘tongue’ (p. 47)
     b. tung + anaː + n + ki [tunganaːki] ‘for the doing’ (p. 90)

In (26) a long vowel surfaces before a syllable-final singleton coda consonant. Since vowel shortening occurs before a geminate in (25), the Koya data in (24)–(26)
are consistent with the weight system in (22) in which CVV and CVG syllables are bimoraic but CVC syllables are light.\(^{25}\)

While the above examples of Koya and Swedish are cases where vowel shortening occurs in syllables closed by a geminate, one can also find languages where vowel lengthening processes are prevented in CVG syllables but not in CVC syllables. This suggests that in such languages geminates are underlyingly moraic, though coda consonants in general are not; vowel lengthening then does not apply before a geminate since that would create a trimoraic syllable. A good example of this comes from Seto (Southeastern Estonian) discussed by Kiparsky (2008c). According to Kiparsky, this language has feet that are required to be trimoraic and this is normally implemented by foot-final vowel lengthening. Because of this process, a foot with the underlying sequence CV.CVC surfaces as CV.CVVC. However, given an input structure where the final consonant of the foot is part of a geminate, that is CV.CVG, no vowel lengthening occurs. This provides evidence that the geminate is underlyingly moraic; foot-final vowel lengthening need not occur in CV.CVG since the foot is already trimoraic.

A different case of a language that avoids the surfacing of CVVG syllables can be found in the West African language Fula as discussed by Paradis (1988) and Sherer (1994). Fula avoids CVVG syllables by degemination of the consonant. Importantly, as shown in (27), Fula allows CVVC syllables both morpheme-internally and over a morpheme boundary.

(27) CVVC syllables in Fula (Sherer 1994: 176)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>kaakt-ε</td>
<td>‘spittle’</td>
</tr>
<tr>
<td>b.</td>
<td>caak-ri</td>
<td>‘couscous’</td>
</tr>
</tbody>
</table>

Fula has a suffixation process that triggers the gemination of a root-final consonant. This is exhibited in the singular/plural alternations in (28). Because of a constraint in Fula requiring geminates to be [−continuant], a root-final continuant segment changes to a stop when it geminates. (I thank Abbie Hantgan for help on the Fula data.)

(28) Fula morphological gemination (Paradis 1988: 78)

<table>
<thead>
<tr>
<th>Stem (sg.)</th>
<th>Suffixed form (pl.)</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. lew</td>
<td>lebb-i</td>
<td>month</td>
</tr>
<tr>
<td>b. lef</td>
<td>lepp-i</td>
<td>ribbon</td>
</tr>
</tbody>
</table>

Of relevance here is that when a long vowel precedes the stem-final consonant, gemination fails to occur, but the stem-final consonant nonetheless is realized as a stop. Consider the singular/plural alternations in (29).

(29) Lack of gemination after a long vowel (Paradis 1988: 80)

<table>
<thead>
<tr>
<th>Stem (sg.)</th>
<th>Suffixed form (pl.)</th>
<th>Expected form</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. laaw</td>
<td>laab-i</td>
<td>*laabb-i</td>
<td>road</td>
</tr>
<tr>
<td>b. lees</td>
<td>leec-ε</td>
<td>*leécc-ε</td>
<td>bed</td>
</tr>
</tbody>
</table>
Given that gemination is part of this suffixing process, we note that the expected forms in (29), where the initial syllable would be CVVG, fail to surface as such. Rather, given the nature of the occurring suffixes forms in (29), it appears that degemination has occurred. This can be understood as the avoidance of a trimoraic CVVG syllable. Since CVVC syllables are allowed in Fula as seen in (27), Fula seems then to be a language that instantiates the weight system of (22) where CVG syllables are heavy but not other CVC syllables.

With respect to the stress evidence for the weight division in (22), probably the strongest case against the moraic analysis of geminate consonants is the observation from Tranel (1991) that there do not seem to be quantity-sensitive stress systems that support the weight division in (22) where stress would be attracted onto a syllable with a long vowel or closed by a geminate consonant but not on one closed by a non-geminate. Tranel (1991) points to the Uralic language Selkup as a language that has geminates but where CVG syllables are ignored by stress assignment even though syllables with long vowels attract stress. Consider the data below in (30). (The data in (30a–f) are given in Halle and Clements (1983: 189), while the data in (30g–h) are reported by Ringen and Vago (2011) for the Taz dialect of Selkup (citing the Selkup scholar, Eugene Helimski, p.c.), which has the same stress pattern as that shown by Halle and Clements.)

(30) Selkup (Halle and Clements 1983)

| a. | qumó:qi | ‘two human beings’ |
| b. | ú:cqo | ‘to work’ |
| c. | u:c:mit | ‘we work’ |
| d. | quminik | ‘human being (dat.)’ |
| e. | ámirna | ‘eats’ |
| f. | ú:cikkak | ‘I am working’ |
| g. | éssükka | ‘(it) happens (occasionally)’ |
| h. | ess Já:qo | ‘to happen (already)’ |

In Selkup, primary stress falls on the rightmost syllable with a long vowel (30a–c) or on the initial syllable if there are no long vowels (30d). A CVC syllable does not count as heavy for stress (30e), even if the CVC syllable is closed by a geminate as seen in (30f–g). As noted by Tranel (1991), if stress is targeting bimoraic syllables and geminates are underlyingly moraic, then the second syllable in (30f–g) would be the rightmost bimoraic syllable. Both the vowel and the geminate would contribute a mora to the second syllable. The fact that (30f–g) do not receive stress on the second syllable seems to provide evidence against geminates being moraic.

The stress pattern of Selkup does not appear to be unique in ignoring geminate consonants. Davis (1999a: 41) points to the Altaic language Chuvash (Krueger 1961) which has an almost identical stress pattern to that of Selkup: stress is attracted to the rightmost syllable with a full vowel but CVG syllables are ignored.
Thus, in both Chuvash and Selkup, CVG syllables do not seem to function like bimoraic CVV syllables but instead act like monomoraic CV and CVC syllables. Furthermore, languages where the stress pattern supports the syllable weight division in (22) seem rare. Davis (1994) discusses a Hindi dialect described by Gupta (1987) in which stress is attracted to the leftmost heaviest syllable in the word. The dialect treats both CVV and CVG syllables as bimoraic while CVC syllables behave as light. But as Curtis (2003) has pointed out, the pattern described by Gupta may be unusual among Hindi dialects in distinguishing CVG from CVC syllables.

One example of a language where stress assignment distinguishes CVG syllables from CVC syllables is the Uto-Aztecan language Cahuilla, and this is noted by Hayes (1995a). Consider the data in (31).

(31) Cahuilla stress (Hayes 1995a and references therein)

a. tākaličem ‘one-eyed ones’
b. čexiwèn ‘it is clear’
c. tāxmuʔat ‘song’
d. qānkčem ‘palo verde (pl.)’
e. héʔi kākawlāːqà ‘his legs are bow-shaped’
f. čexxiwen ‘it is very clear’

Following Hayes (1995a), the Cahuilla stress pattern in (31) reflects the assignment of moraic trochees starting from the left edge of the word. This is clearly shown in (31a–b). The comparison between (31c) and (31d) is interesting. In (31d), the first syllable is being treated as bimoraic given that there is a secondary stress on the second syllable on its own. There is a secondary stress on the second syllable since that would be the head of the second trochaic foot in the word. In (31c), even though the first syllable is closed, it is not treated as bimoraic. There is no secondary stress on the second syllable. Thus, the comparison between (31c) and (31d) shows that a syllable with a long vowel (i.e. the first syllable in (31d)) is regarded as bimoraic whereas a CVC syllable (i.e. the first syllable in (31c)) is regarded as monomoraic. This is also made clear by the form [kākawlāːqà] in (31e) where the CVC second syllable is skipped for stress, and the CVV third syllable forms a bimoraic foot on its own. The stress on the final syllable in (31e) indicates that Cahuilla allows a degenerate (i.e. monomoraic) foot in word-final position as discussed by Hayes (1995a). Given this, it is noteworthy that the initial CVG syllable in (31f) counts as bimoraic forming a trochaic foot on its own. Just as in the case of (31d), there is a secondary stress on the second syllable in (31f). While Cahuilla provides evidence for the weight distinction in (22), the general lack of stress evidence for (22) poses a challenge for the inherent underlying moraic analysis of geminates as in (21d). Perhaps, if more languages with the right set of properties are considered (i.e. stress languages with long vowels, coda consonants, and geminates), other cases supporting (22) will be found.
We now briefly turn to evidence for the underlying moraification of geminates in (21d) that is independent of the weight system in (22). Based on Davis (1999a) there are cases where morphological allomorphy seems to be sensitive to the underlying moraification of the stem (as opposed to its surface moraic structure) and this can provide evidence for the underlying moraic nature of geminates. As one example, consider the Hausa plural pattern in (33)–(36) referred to as Class 3 plurals by Kraft and Kraft (1973) and discussed in such works as Newman (1972, 1992) and Leben (1980) (though the analysis to be presented here is somewhat different from these works). These plurals involve the suffixation of the two different allomorphs in (32) (where the C-slot in (32a) is realized as a consonant identical to the last root consonant).

(32)  a. -aaCee  b. -ayee

The data in (33) show nouns whose roots end in a single consonant while the data in (34) show nouns whose roots end in a consonant cluster. These roots select the plural allomorph in (32a). (As seen in the data, most singular nouns in Hausa end in a final long vowel extension that is not part of the root. Tones are not indicated in the data, but the affixation is accompanied by a HLH tone pattern over the whole plural form.)

### (33)
<table>
<thead>
<tr>
<th>Singular</th>
<th>Plural</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. dam-oo</td>
<td>dam-aamee</td>
<td>‘monitor’</td>
</tr>
<tr>
<td>b. wur-ii</td>
<td>wur-aaree</td>
<td>‘place’</td>
</tr>
<tr>
<td>c. kaf-aa</td>
<td>kaf-aafee</td>
<td>‘small hole’</td>
</tr>
</tbody>
</table>

### (34)
<table>
<thead>
<tr>
<th>Singular</th>
<th>Plural</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. gulb-ii</td>
<td>gul-aabee</td>
<td>‘stream’</td>
</tr>
<tr>
<td>b. birn-ii</td>
<td>bir-aanee</td>
<td>‘city’</td>
</tr>
<tr>
<td>c. kask-oo</td>
<td>kas-aakee</td>
<td>‘bowl’</td>
</tr>
</tbody>
</table>

Now consider the data in (35) and (36) which select the plural allomorph in (32b). In this allomorph the suffix’s second syllable has an onset realized by the default consonant [y] which Newman (1972) notes is found elsewhere in Hausa as an epenthetic consonant. (35) contains examples where the root vowel is long and (36) shows examples where the root vowel is a diphthong.

### (35)
<table>
<thead>
<tr>
<th>Singular</th>
<th>Plural</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. zoom-oo</td>
<td>zoom-aayee</td>
<td>‘hare’</td>
</tr>
<tr>
<td>b. kiif-ii</td>
<td>kiif-aayee</td>
<td>‘fish’</td>
</tr>
<tr>
<td>c. suun-aa</td>
<td>suun-aayee</td>
<td>‘name’</td>
</tr>
</tbody>
</table>

### (36)
<table>
<thead>
<tr>
<th>Singular</th>
<th>Plural</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. bauun-aa</td>
<td>bauun-aayee</td>
<td>buffalo</td>
</tr>
<tr>
<td>b. mais-oo</td>
<td>mais-aayee</td>
<td>disused farm</td>
</tr>
<tr>
<td>c. gaul-aa</td>
<td>gaul-aayee</td>
<td>idiot</td>
</tr>
</tbody>
</table>
Based on the data in (33)–(36) one could generalize that if the root syllable contains a short vowel the allomorph -aaCee is selected, but if the root syllable contains a long vowel or diphthong then -aayee is selected. However, the data in (37) show that this generalization is not quite correct.

(37) S/\text{ingular} & \text{Plural} & \text{Gloss} \\
\text{a. tukk-uu} & \text{tukk-aayee} & \text{bird crest} \\
\text{b. tall-ee} & \text{tall-aayee} & \text{soup pot} \\
\text{c. gamm-oo} & \text{gamm-aayee} & \text{heat pad} \\

The nouns in (37) have short root vowels, yet they nonetheless pattern like the nouns containing long vowels and diphthongs in (35) and (36) by taking the plural allomorph shown in (32b). The question that emerges then is what unifies the roots in (35)–(37) that distinguishes them from those in (33)–(34). The answer seems to reside in their underlying weight. The root forms in (35) and (36) have the shape CVV, those in (37a–c) have the shape CVG. These then are forms where the roots are at least bimoraic underlyingly given Hayes’s moraification algorithm as reflected in (21). On the other hand, the roots in (33) and (34) would be underlyingly monomoraic. We see then that the allomorph in (32a) only attaches to a noun root that is underlyingly monomoraic. The allomorph in (32b) attaches to noun roots that are at least bimoraic underlyingly. To be clear, the plural allomorphy, which is a lexical process in Hausa, is sensitive to the underlying mora structure of roots and not to the surface mora structure. This is because there is much evidence in Hausa showing that surface CVC syllables are always bimoraic in Hausa. (Newman (1972) mentions vowel shortening in closed syllables among other processes.) Though the initial CVC syllables in the singular forms in (34) would surface as bimoraic because of the weight of the coda consonant, the plural allomorphy, which is sensitive to the underlying moraic structure treats (34) as monomoraic. This, then, can be seen as providing evidence for an underlying distinction between geminate consonants, which are moraic and singleton consonants which are not.

A final type of evidence for the moraic nature of geminates as in (21d), independent of the weight system in (22), comes from the behavior of word-initial geminates. Though such geminates are rare, they are attested in a number of languages. (In fact the dissertations of Muller (2001) and Topintzi (2006) are exclusively on initial geminates.) Muller (2001), whose study includes acoustic analyses of word-initial geminates, concludes that initial geminates are moraic in some languages but not in others. Topintzi (2006, 2008) focuses on languages where initial geminates pattern as moraic, and argues that such geminates constitute moraic onsets, thus providing a case where onsets carry weight. An example of a language where a word-initial geminate patterns as moraic is Trukese (also called Chuukese). Consider the data in (38) and (39) (cited from Davis 1999b and Davis and Torretta 1998, and see references cited therein) that reflect a minimal word constraint on Trukese nouns. (Note that Trukese forms are given in transcription rather than in Trukese orthography.)
### Trukese and Leti: Word-Initial Geminates

Trukese has a general process whereby a word-final long vowel shortens, as in (39). However, as (38) shows, shortening does not apply if the result would be monomoraic. This is because Trukese has a minimal word constraint that requires nouns to be bimoraic. The fact that the word-final vowel does shorten in (39b–d) strongly suggests that the initial geminate is moraic. That is, an output such as [tto] in (39b) is bimoraic with a mora being contributed by both the vowel and the geminate. As another example of a word-initial geminate acting as moraic, Topintzi (2006, 2008 and references cited therein) refers to stress evidence from Pattani Malay. In Pattani Malay there are no long vowels, and geminates only occur in word-initial position. Normally in Pattani Malay primary stress falls on the final syllable of a word except when the word begins with a geminate consonant, in which case the initial syllable is stressed. This, too, can be taken as evidence for the moraicification in (21d) where a geminate is underlingly moraic.

Nonetheless, there are cases where word-initial geminates do not pattern as moraic. A good example is that of the Austronesian language Leti as discussed by Hume et al. (1997). Leti like Trukese has initial geminates and a bimoraic minimal word requirement, but unlike Trukese, Leti does not have words of the pattern shown in the output forms in (39b–d) consisting of an initial geminate followed by a short vowel. The lack of such words can be taken as strong evidence that the initial geminate does not count as moraic in Leti. Perhaps one can understand the difference between Trukese and Leti in terms of the language-specific phonotactics. In Trukese, word-initial geminates are permitted but there are no word-initial consonant clusters. In Leti on the other hand, not only are word-initial geminates permitted but almost any possible word-initial sequence of two consonants can occur with no apparent sonority restrictions between them. Given this patterning, one could analyze the first consonant of a word-initial cluster in Leti as being extraprosodic. The initial consonant of such a cluster is unrestricted and can be identical to the following consonant. This means that the word-initial geminate of Leti consists of a sequence of identical consonants; the first consonant of the sequence would be extraprosodic just like the first consonant of any other word-initial cluster. Still, the difference between Leti and Trukese suggests that there is no consistent behavior in the weight properties of word-initial geminates.32

---

### Table: Trukese Geminates

<table>
<thead>
<tr>
<th>Underlying Representation</th>
<th>Output Form</th>
<th>Gloss</th>
<th>Unattested output</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. /maa/</td>
<td>[maa]</td>
<td>behavior</td>
<td>*ma</td>
</tr>
<tr>
<td>b. /taa/</td>
<td>[taa]</td>
<td>islet</td>
<td>*ta</td>
</tr>
<tr>
<td>c. /oo/</td>
<td>[oo]</td>
<td>omen</td>
<td>*o</td>
</tr>
</tbody>
</table>

### Table: Leti Geminates

<table>
<thead>
<tr>
<th>Underlying Representation</th>
<th>Output Form</th>
<th>Gloss</th>
<th>Suffixed form</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. /etiruu/</td>
<td>[etiru]</td>
<td>coconut mat</td>
<td>[etiruu-n]</td>
</tr>
<tr>
<td>b. /tt00/</td>
<td>[tt0]</td>
<td>clam sp.</td>
<td>[tt00-n]</td>
</tr>
<tr>
<td>c. /ɛčča/</td>
<td>[čča]</td>
<td>blood</td>
<td>[čča-n]</td>
</tr>
<tr>
<td>d. /ssɔɔ/</td>
<td>[ssɔ]</td>
<td>thwart of a canoe</td>
<td>[ssɔɔ-n]</td>
</tr>
</tbody>
</table>

### Notes:
- Trukese has a general process whereby a word-final long vowel shortens, as in (39).
- Shortening does not apply if the result would be monomoraic.
- In Pattani Malay, primary stress falls on the final syllable of a word except when the word begins with a geminate consonant.
- In Leti, word-initial geminates are permitted, but there are no word-initial consonant clusters.
- Leti geminates are analyzed as extraprosodic, allowing for sequences of identical consonants.
- The difference between Trukese and Leti suggests no consistent behavior in the weight properties of word-initial geminates.
Given this inconsistent behavior of geminate consonants, we suspect that the topic of the phonology of geminate consonants will continue to be a controversial one.

6 Weight Inconsistencies

So far in this chapter we have assumed that languages witness a weight consistency, such that, for example, if CVC syllables are heavy in a language (or if CVS syllables are heavy in a language), they consistently pattern as bimoraic throughout the language. Nonetheless, researchers have observed that weight inconsistencies not only occur but are quite common. These typically involve CVC syllables sometimes patterning as light and sometimes as heavy within the same language.

In this section we will briefly examine two different situations where weight inconsistencies are found. In Section 6.2 we will consider what Hyman (1992) terms “moraic mismatches” or prosodic inconsistencies where within the same language certain syllable types act as heavy for one process but light for another. First, though, in Section 6.1, we will consider the somewhat different phenomenon of context dependent weight (Hayes 1994; Rosenthal and van der Hulst 1999) where the weight of a CVC syllable is based on its context within a word.

6.1 Context Dependent Weight

Context dependent weight is a phenomenon noted by such researchers as Kager (1989), Hayes (1994, 1995a), Alber (1997), Rosenthal and van der Hulst (1999), and Morén (1999, 2000) whereby CVC syllables surface as heavy (bimoraic) only in certain contexts. Specifically, one must consider the makeup of the other syllables within a word to know whether a particular CVC syllable will pattern as bimoraic. Context dependent weight of CVC syllables occurs in certain stress systems whereby a CVV syllable in a word receives stress, but if a word has no CVV syllables, a CVC syllable receives stress. This exemplifies context dependent weight since a CVC syllable is heavy (bimoraic) only in words without long vowels. As an example of context dependent weight, both Rosenthal and van der Hulst (1999) and Morén (2000) discuss Kashmiri stress. Consider the Kashmiri data in (40).

(40) Kashmiri stress pattern (Rosenthal and van der Hulst 1999; Morén 2000)

- a. vah.ráa.vun ‘to spread’
- b. báa.laa.dar ‘balcony’
- c. yu.nu.vár.si.ti ‘university’
- d. jóm.bir.zal ‘narcissus’
- e. á.ní.ga.ti ‘darkness’

The generalization illustrated by the stress patterns in (40) is that stress falls on the leftmost heavy syllable. While a CVV syllable is always bimoraic, a CVC
sylable can be heavy only in a word without long vowels (i.e. other bimoraic syllables). Concerning the data in (40), the items in (40a–b) show that in words containing long vowels, primary stress goes on the syllable containing the leftmost long vowel; (40c–d) indicate that if the word has no long vowels then primary stress goes on the leftmost closed syllable; otherwise, stress goes on the leftmost (initial) syllable as in (40e). (Note that the references cited for Kashmiri indicate that final syllables are never stressed and would be extrametrical.) The word-initial CVC syllable in (40a) does not pattern as bimoraic while the word-initial CVC syllable in (40d) and the CVC syllable in (40c) do since they receive stress. Thus, we see from the Kashmiri data that a CVC syllable can only surface as bimoraic in words lacking any CVV syllables.

In terms of a formal optimality-theoretic analysis (similar but not identical ones are offered by Rosenthal and van der Hulst (1999) and Morén (2000)), the Weight-to-Stress constraint forces a coda to surface as moraic in words lacking inherently bimoraic CVV syllables. In examples like (40c), then, the CVC syllable surfaces as bimoraic with stress since there are no other potentially bimoraic syllables to stress (given that the language does not lengthen underlying short vowels to make a syllable bimoraic). It is interesting that this analysis of context dependent weight can account for the occurrence of certain syllable weight hierarchies that were mentioned in Section 3. Given the Kashmiri stress data in (40), one could suggest that Kashmiri has a syllable weight hierarchy in which CVV syllables are heavier than CVC syllables which in turn are heavier than CV syllables. Stress would then fall on the leftmost heaviest (non-final) syllable in the word. However, given the contextual weight of CVC syllables, some of the cases of syllable weight hierarchies reported in the literature can be viewed as reflecting a system of context dependent weight. For example, the syllable strength hierarchy for Nanti in (19) is probably better analyzed as involving context dependent weight rather than as reflecting an independent coda strength hierarchy.

It is important to mention another phenomenon that is considered by Rosenthal and van der Hulst (1999) to be a type of context dependent weight. Without doubt, the most common case of variable coda weight is final consonant extrametricality where in languages in which CVC syllables are normally bimoraic, a word-final CVC syllable functions as light. This is a pervasive phenomenon that occurs in many unrelated languages. (See Hayes (1995a: 58–60), in particular, for arguments supporting the notion of final extrametricality and Hyde (2007b) for an elaborated theory of non-finality.) In languages like Arabic, a CVC syllable normally attracts stress if it constitutes the penultimate syllable of the word but not if it constitutes the final syllable. In such languages, a penultimate CVC syllable would be bimoraic but a final CVC syllable would be monomoraic. In an optimality-theoretic analysis such as that of Rosenthal and van der Hulst (1999) the constraint ranking would normally select a coda that is moraic in the output, but a high-ranking non-finality constraint prevents a word-final coda from being moraic; thus a word-final CVC syllable patterns as light.

A very intriguing observation regarding final extrametricality of CVC syllables put forward by Ham (1998) is that final CVC syllables are always extrametrical
in languages that have word-final geminates. This is because a word-final geminate is moraic and would need to be distinguished in final position from a potential moraic coda. Given the underlying representation of geminates as in (21d), final extrametricality of CVC syllables is able to preserve the contrast between an underlying final geminate and the corresponding final singleton consonant. The geminate of a final CVG syllable would surface as moraic while the singleton coda of the final CVC would be non-moraic. This difference is found in Arabic dialects where a final CVG syllable attracts stress making it distinct (i.e. bimoraic) from a final CVC syllable which is light (monomoraic) and does not attract the stress. In a variety of other languages having word-final geminates examined by Ham (1998) the same distinction is made between final CVG and CVC syllables. If Ham’s (1998) observation holds up to further scrutiny, it constitutes an interesting argument for the underlying moraification of geminate consonants. (See also Topintzi (2008: 175) for discussion on this point.)

6.2 Moraic Mismatches

Another type of weight inconsistency found in languages involves CVC syllables being treated differently depending on the specific process. One would expect that if CVC syllables act as bimoraic in a language they would act as bimoraic for all relevant weight-sensitive processes in that language. Nonetheless, researchers such as Crowhurst (1991), Steriade (1991), Hyman (1992), Broselow (1995), Hayes (1995a) have noted what Hyman (1992) refers to as moraic mismatches and others such as Fitzgerald (forthcoming) call prosodic inconsistencies. This is the case where within a single language CVC syllables sometimes act as heavy and sometimes as light depending on the process at issue. Hyman (1992) gives a variety of examples from Bantu languages. Consider the data Hyman (1992: 258–259) provides from the Runyambo-Haya dialect cluster of Tanzania in (41). (High tone is represented by an acute accent.)

(41) Runyambo-Haya dialect cluster (Tanzania)

a. Assign a high tone to the second mora of a verb stem
   1. ni-tu-rim-á ‘we are cultivating’
   2. ni-tu-siig-a ‘we are smearing’
   3. ni-tu-jend-a ‘we are going’

b. Reduplication (the final a vowel reduplicates as long if the stem syllable is monomoraic, but it reduplicates as short if the stem syllable is bimoraic)
   1. (ku-) lim-a → (ku-) limaa-lima ‘to cultivate’
   2. (ku-) siig-a → (ku-) siiga-siiga ‘to smear’
   3. (ku-) genda → (ku-) genda-genda ‘to go’ *(ku-) gendaa-genda

In (41a) we see a process whereby a high tone is assigned to the second mora of a verb stem (underlined in 41a). As shown in the third example in (41a), a preconsonantal nasal consonant does not add a mora to the verb stem. In (41b) we
see a verbal reduplication process whereby the final vowel /a/ reduplicates as long if the verb stem is monomoraic but as short if the stem syllable is bimoraic. Here, the third example in (41b) shows that the preconsonantal nasal of the verb stem does add a mora. Thus, according to Hyman (1992a), Runyambo-Haya exemplifies a moraic mismatch. Mora count is different for tone (41a) versus reduplication (41b). Tone assignment treats the preconsonantal nasal as non-moraic while the same nasal adds a mora to the verb stem with respect to reduplication.

Along similar lines, Steriade (1991) notes that in Khalkha Mongolian the stress rule treats CVC syllables as light or monomoraic since stress is attracted to a CVV (bimoraic) syllable skipping over CVC syllables, but a bimoraic constraint on minimal verb stems regards both CVV and CVC as heavy (bimoraic), given that no CV stems are found but CVV and CVC stems occur. Steriade (1991) suggests that it is only processes like stress and tone that are involved in moraic mismatches treating otherwise bimoraic CVC syllables as light. Specifically, according to Steriade (1991), languages may restrict tone and stress bearing elements to those above a certain sonority threshold where pitch realization would be clearest, and this does not necessarily mean that CVC syllables are monomoraic.

The notion of moraic mismatches, that certain weight-sensitive processes in a language may consider CVC syllables as light while others treat them as heavy, is most fully developed in a series of important works by Gordon (1999, 2001, 2002, 2004b). Gordon argues for a radical departure viewing weight as process-driven rather than language-driven. In Gordon’s view, moraic mismatches should be the expected case. It just depends on the process whether certain CVC syllables will be treated as heavy, not on the language. In order to show this, Gordon (2004) reports on a survey of six weight-sensitive phenomena in approximately 400 languages. Phenomena that he surveyed included stress, tone, poetic metrics, compensatory lengthening, minimal word requirements, and templatic restrictions. As an example of his findings supporting his view, he found that most languages treated syllables closed by an obstruent (CVO) as heavy with respect to minimal word constraints but as light with respect to tonal phenomena. Another finding is that syllables closed by a sonorant, (CVS), frequently pattern as heavy with CVV syllables for tonal phenomena, but this patterning is rare for stress. An example of this, also discussed by Blevins (2004), is Lhasa Tibetan which has both tone and stress. Lhasa contour tones, which are only realized on heavy syllables, treat CVV and CVS as heavy, but CVO as light. On the other hand, Lhasa stress, which treats CVV syllables as heavy, considers all CVC syllables as light, including CVS syllables. Lhasa Tibetan represents a typical finding in Gordon’s survey. Generally, with respect to stress patterns, Gordon found that languages either treated just CVV syllables as heavy (as in Lhasa Tibetan) or both CVV and CVC syllables as heavy; it was rare for a language to treat CVS as heavy with respect to stress without also treating CVO as heavy. But, as mentioned, CVS syllables often patterned with CVV syllables as heavy with respect to tonal phenomena.

Gordon maintains that the reason why tone and stress pick out different syllable types as heavy is because they have a different phonetic basis. Gordon (2004:
285–286) explains the patterning of CVV and CVS as heavy for tone along the following lines:

The physical correlate of tone is fundamental frequency, which is only present in voiced segments . . . Crucially, the fundamental frequency profile of a segment or syllable (and hence its tonal profile) is cued not only by the fundamental itself but also by the higher harmonics . . . The presence of harmonics greatly enhances the presence of fundamental frequency . . . [T]he more crucial harmonics for the perception of the fundamental, the low frequency harmonics (House 1990), are typically present in sonorants . . . In contrast to sonorants, obstruents provide either minimal or no cues to fundamental frequency.

Thus, tone is sensitive to the presence of certain harmonics found in sonorants but not in obstruents. Consequently tonal phenomena can treat both CVV and CVS syllables as heavy. On the other hand, according to Gordon, stress is sensitive to the overall auditory energy in the syllable rhyme. Sonorancy and voicing are two of the best features for predicting higher energy values in the rhyme. Gordon (2004) makes the interesting claim that a major factor in determining whether or not CVC syllables act as heavy for stress is the nature of the language’s coda inventory. For example, if we compare Khalkha Mongolian which treats CVC syllables as light for stress with Finnish which treats them as heavy, Gordon observes that in Khalkha Mongolian there are more voiceless consonants that can occur in coda position than voiced ones (sonorants included) while in Finnish there are more voiced consonants in coda position than voiceless ones. Consequently, in general, Finnish codas have more auditory energy than Khalkha codas and so Finnish CVC syllables behave as heavy with respect to stress while Khalkha Mongolian CVC syllables behave as light. Thus, not only does Gordon argue for the process-specific nature of syllable weight, he argues further that the processes do not share a single phonetic basis for what makes a syllable heavy.

While Gordon’s research offers a new and different perspective on syllable weight, a perspective that focuses on process types rather than on issues that emerge from a moraic theory of quantity as has been the focus of this chapter, critical assessment of Gordon’s work can be found in Curtis (2003), Topintzi (2006), and de Jong (2000). Curtis (2003: 290) notes that all of Gordon’s mismatches involve CVC syllables with varying codas and do not involve CVV syllables. She concludes that his findings support phenomenon-specific weight-by-position rules that can reference features such as sonorant, but they do not challenge the inherent weight of vowels or even of geminates, nor the structural representation of syllable weight. Topintzi (2006) focuses her criticism on stress being related to overall auditory energy noting, among a variety of potential problems, the lack of sonorant onsets affecting stress in languages where onsets seem to matter for stress. As was mentioned in Section 4, Topintzi observes that when a feature of an onset consonant influences stress it is usually a voiceless consonant that attracts the stress. This is the opposite of what Gordon’s theory of auditory energy would predict. An important criticism of Gordon’s theory connecting the patterning of
CVC syllables with respect to stress with overall auditory energy in the syllable rhyme comes from Ahn’s (2000) typological survey of stress systems, which is discussed in detail by de Jong (2000). Ahn makes a strong claim based on stress descriptions of 136 languages that quantity-sensitive unbounded stress systems (such as that of Khalkha Mongolian) always treat CVC syllables as light while bounded systems such as Finnish can treat them as heavy. This difference between how CVC syllables are treated in bounded vs. unbounded systems does not fall out from Gordon’s theory. De Jong (2000) suggests that the underpinnings of stress are different in the two systems: bounded systems can reflect delimitative intonation marking over syllables while unbounded systems reflect loudness or prominence. (See de Jong (2000) for details as to why this can result into a different treatment of CVC syllables.)

7 Conclusion

In this chapter we have surveyed a variety of issues that emerge from the moraic representation of quantity. Our focus has been almost exclusively on phonological issues concerning the patterning of possible moraic elements and their role in determining syllable weight. We have spent less time examining alternative representational proposals regarding syllable weight (though some of these are mentioned in Section 1 and in the notes), not only because of space limitations but also because recent critical works (e.g. Curtis 2003; Kraehenmann 2001; Ringen and Vago 2011) assume familiarity with moraic theory.

We have dealt with the phonetics of syllable weight only briefly here, in the discussion regarding Gordon’s work in Section 6.2. One might expect that if CVV and CVC syllables stand out for a variety of processes it is because they are longer than CV syllables. But a direct correlation with duration was already observed to be problematic by Trubetzkoy (1939), who notes phonetic contextual effects on duration that are not moraic, such as contextual effects on vowel duration depending on the nature of surrounding elements. A good example showing that phonetic duration does not necessarily translate to syllable weight is the observation by Chen (1970) that vowels are longer before voiced consonants. This seems to have no effect on stress placement. We do not find stress systems where primary stress falls on a syllable containing a vowel before a voiced consonant. Perhaps for such reasons, Hayes (1995a: 271) emphasizes phonological duration rather than phonetic duration in his discussion of weight: “weight can be thought of as a property of the time dimension: a syllable is heavy because it is long. This is the viewpoint of moraic theory: the moras form an abstract characterization of a syllable’s phonological duration.” But even a view of phonological duration was called into question by Newman (1972: 320) who maintained that, “... the distinction between heavy and light syllables cannot be assumed on a priori grounds to be phonologically analyzable in terms of units of duration nor to be phonetically correlated with actual-time differences.” Gordon (2004) as well takes a position against a strictly durational characterization of syllable weight noting that it is
problematic in languages that have both CVV and CVC syllables but where only CVV syllables pattern as heavy, since he finds in his durational study of Khalkha Mongolian (which treats only CVV as heavy with respect to stress) that the duration of rhymes in both CVV and CVC syllables are distinct from rhyme duration of CV syllables. Thus, he concludes that duration alone is not a good fit for determining the basis of syllable weight in such languages. What emerges from phonetic explorations into syllable weight like that of Gordon as well as de Jong (2000) is that there is no one phonetic correlate of syllable weight. Thus, one could maintain that syllable weight is an abstraction and that moras form its abstract characterization.

ACKNOWLEDGMENTS

This chapter could not have been written without the work done by the participants in the syllable weight seminar at Indiana University in Spring 2007. In particular, I thank Majdi Bamakhramah, Nick Henriksen, Chisato Kojima, Dongmyung Lee, Paco Montaño, Jung-yueh Tu, Junghyoe Yoon, and Kenji Yoshida. I also thank Steve Chin, Chris Green, Iskra Iskrova, Brian José, Jungsun Kim, Traci Nagle, Nathan Sanders, and Natsuko Tsujimura for giving me detailed comments on a preliminary version of this chapter. Further, I thank Brett Baker, Ken de Jong, Daniel Dinnsen, Tracy Alan Hall, Abbie Hantgan, Michael Marlo, Paul Newman, Anne Pycha, Robert Vago, and Islam Youssef for discussion on specific aspects of this paper. Finally, I thank the editors, especially John Goldsmith, for their valuable comments in improving the quality, if not the quantity, of this chapter.

NOTES

1 In this chapter, length is transcribed either by a colon or by a sequence of identical letters. For example, geminate-p will be transcribed either as [p:] or [pp].

2 Note, though, that such contextual phonetic differences may eventually become relevant phonologically as in the development of compensatory lengthening in Friulan, as discussed by Kavistskaya (2002).

3 This is not to say that issues of representation were unimportant in the discussion of quantity in the pre-generative period. Relevant discussion can be found in Trubetzkoy (1939: 173–174) of the English translation and in Hockett (1955: 76–77). Somewhat relatedly, with respect to English there is a long standing debate on whether the tense-lax vowel distinction is really one of quantity or quality and how the difference should be represented. DuPonceau (1818) is critical of those who do not recognize a quantity distinction in the description of English vowel sounds (pp. 239–240) and further suggests that the English length distinction should be represented by a diacritic on a short vowel. Trager and Bloch (1941) took vowel quantity in English as h since length and h can be interpreted as being in complementary distribution. On the other hand, Chomsky and Halle (1968) viewed the English vowel contrast as one of quality (tense vs. lax) rather than quantity.
4 All references to Trubetzkoy 1939 in this chapter are to the 1969 English translation of it. See Anderson (1985: 100–106) for discussion regarding Trubetzkoy’s view of quantity.

5 According to Hock (1986b: 90) a notion like the mora can be traced back to the Sanskrit Grammarians of the fifth century B.C.E. The earliest linguistic reference to mora mentioned by the OED is from volume XI (p. 591) of the first edition of the Encyclopedia Americana (1832) where the term is equated with a short syllable. Bloomfield (1933: 110) refers to mora as “an arbitrary unit of relative duration.” Important works within the framework of generative phonology that incorporate or develop a theory of mora include Newman (1972), Hyman (1985), Hock (1986a), and McCarthy and Prince (1986). These all slightly differ from one another and from Hayes (1989a). It is Hayes’s theory that is most influential in current work within moraic phonology.

6 We use the terms “onset,” “nucleus,” and “coda” to make reference to the different positions within the syllable, with onset being the syllable initial consonant or consonants, the nucleus being the vowel or peak of the syllable, and the coda being the syllable-final consonant or consonants. Here, we make no assumption regarding the formal status of these as constituents of the syllable, but see Davis (2006) for an overview.

7 In this chapter it is assumed that a non-moraic consonant, whether an onset or a coda, attaches directly to the syllable node. While this is a fairly standard assumption for onset consonants and was assumed by Hayes (1989a), Katada (1990) argues that an onset shares a mora with the following vowel as in (i). With respect to the coda consonant, Broselow et al. (1997) give phonetic arguments for a non-moraic coda sharing a mora with a preceding vowel as in (ii):

\[
\begin{align*}
(i) & \quad \sigma \\
& \quad \mu \\
& \quad C \ V
\end{align*}
\]

\[
\begin{align*}
(ii) & \quad \sigma \\
& \quad \mu \\
& \quad V \ C
\end{align*}
\]

See the works cited for discussion of these technical issues.

8 In this chapter we make the common assumption that an intervocalic geminate is heterosyllabic as shown in (7e) and that an intervocalic consonant cluster syllabifies as heterosyllabic as well (ignoring cases where such clusters are also possible complex onsets). Any intervocalic consonant is normally assumed to syllabify as an onset, so that /VCV/ would syllabify as [VCV], but see Blevins (2004) for further discussion.

9 Kavitskaya (2002) notes three cases where the deletion of an onset consonant leads to compensatory lengthening. This is unexpected in Hayes’s theory since onsets are not typically moraic. It may be of relevance that each of these cases involves the deletion of a sonorant in the onset. See the discussion in Note 20 regarding compensatory lengthening in Samothraki Greek, one of the cases discussed by Kavitskaya.

10 Throughout this paper, unless otherwise noted, when syllable types such as CV, CVV, and CVC are listed or discussed, it should be understood that the initial C is to be interpreted as standing for C₀ (meaning zero or more initial consonants). This reflects the general weightlessness of syllable onsets, an issue that will be discussed in Section 4.

11 The moraic theory elaborated on here predicts that there are no languages in which CVC patterns as heavy while CVV patterns as light. This is because long vowels are
underlyingly bimoraic whereas a coda consonant is not underlyingly moraic; so a syllable with a long vowel would always be treated as bimoraic. As noted by Blevins (1995: 237) this prediction seems correct. A separate issue is whether there are languages that lack CVV syllables but where CVC syllables are heavy. Hayes (1989a: 290) mentions Ilokano and Spanish as possible examples and Hayes (1995a: 205) cites the Cariban language Hixkaryana. Seneca (Prince 1983) may be another example, given that its accent system is sensitive to closed syllables but lacks long vowels phonemically. However, Trubetzkoy’s (1939) discussion of syllable weight and Zec’s (1988, 1995b) proposal on sonority thresholds for moraic segments (to be discussed in Section 2) predicts that the presence of bimoraic CVC syllables in a language implies the presence of CVV syllables in that language. We will not be considering this issue further here.

While the system in (13) has [+sonorant] as a minimal sonority threshold on what can be moraic, the system in (10a) would have no minimal sonority threshold on what can be moraic (so even obstruents can be moraic); the system in (10b) would have a strict sonority threshold in which only [+sonorant], [-consonantal] elements (i.e. vocalic phonemes) can be moraic.

There are a variety of syllable structure issues concerning diphthongs that go beyond the scope of this chapter. These include whether the glide element of a diphthong is part of a syllable nucleus or syllable margin and whether rising diphthongs (i.e. a sequence of an on-glide and a vocalic peak) pattern as monomoraic or bimoraic. See Davis and Hammond (1995), Baertsch (2002), Smith (2003) and Levi (2004) for relevant discussion.

While both Tiv and Lithuanian allow only higher sonority consonants to be moraic, they differ with respect to the surfacing of lower sonority consonants such as voiceless stops in coda position. In Lithuanian they surface as non-moriaic in coda position, but in Tiv they never surface in codas. This suggests that Weight-by-Position (i.e. the constraint or rule that requires codas to be moraic) still plays a role in languages where there is a minimal sonority threshold on what can be moraic. Weight-by-Position can be violated in Lithuanian but not in Tiv.

There are languages that witness an interesting interaction of sonority constraints on moraicity and the Coda Condition. In Ponapean (Goodman 1997) only sonorant codas are permitted but they must obey the Coda Condition (ignoring some loanwords). Thus, in Ponapean, the only codas permitted are sonorant consonants that share place features with the following onset and these behave as moraic. We can analyze Ponapean as having minimal sonority threshold on moraics elements requiring them to be sonorant along with the Coda Condition as an edge constraint. In Campidanian Sardinian (Davis and Baertsch 2005, and references cited therein), at least in initial syllables which have less restricted phonotactics than other syllables, a coda consonant must be either a high sonority rhotic (e.g. [ar.ba] ‘white’) or a consonant that obeys the Coda Condition, such as a nasal homorganic to a following onset or the first part of a geminate. The effect of this is that a consonant of lower sonority in the coda must obey the Coda Condition whereas a high sonority coda (i.e. rhotic) need not obey it. See Davis and Baertsch (2005) for a partial optimality-theoretic analysis and discussion. Further Basbøll (2005) offers an analysis of Danish in which both obstruents and sonorants appear in the coda position, but only sonorants are moraic; this frequently cooccurs with the stød.

An alternative way of understanding the hierarchy in (19) is to consider it a case of context dependent weight along the lines of Rosenthal and van der Hulst (1999). On such a view, CVC syllables are normally monomoraic but can be bimoraic in a
context where there would otherwise be only monomoraic syllables in the word. This is discussed in Section 6.1.

The various scales that Crowhurst and Michael (2005) employ (such as the vowel quality scale) to help determine stress location can be seen as a type of prominence projection along the lines of Hayes (1995a). Hayes employs a prominence projection for atypical situations where factors other than syllable weight, such as tonal quality or certain features on phonemes, play a role in stress.

One issue not discussed here is whether a segment can ever display a three-way contrast in quantity. Trubetzkoy (1939: 180–181) doubted such cases existed phonologically even in languages like Estonian and Lapp (Saami), suggesting that the appearance of multiple degrees of quantity were a phonetic effect resulting from other factors in these two languages; however, see Bye, Toivonen, and Sagulin (2008) for evidence of a three-way phonemic consonantal length contrast in Inari Saami.

Odden (2006) discusses a minimality pattern in the Bantu language Zinza where bisyllabic words must either begin with a consonant or a long vowel, but never a short vowel. Superficially, this may look like a case where a CV (initial) syllable patterns together with a syllable beginning with a long vowel, perhaps suggesting the bimoraicity of CV syllables. Odden, though, accounts for the Zinza pattern by referencing a high ranked constraint militating against prosodic words beginning with short vowels. The potential moraicity of onsets is not at issue.

The clearest case in the literature where the deletion of an onset consonant triggers compensatory lengthening involves /r/ deletion in Samothraki Greek where a word-initial /r/ and /r/ as a second member of an onset cluster deletes triggering compensatory lengthening of a following vowel; /r/ does not delete in coda position, but does delete intervocally without triggering compensatory lengthening. Hayes (1989a: 283) suggested an analysis involving epenthesis, but as Topintzi (2006) shows, such an analysis is not motivated synchronically. Topintzi presents a thorough discussion of the Samothraki data and analyzes it as an instance where the deletion of a non-moraic segment results in the addition of a mora by compensatory lengthening, thus constituting a clear counter-example to Hayes’s (1989a) theory. A different way of looking at the Samothraki Greek data is to posit, abstractly, that /r/ deletes when it is forced into the nucleus of the syllable; thus it is moraic when it deletes. This is suggested by Kiparsky (2011). Key to such an analysis is that /r/ can never surface as an onset. Kavitskaya (2002) provides a similar view suggesting that the /r/ is vocalic enough to be interpreted as additional vowel length.

One can speculate that it is the higher pitch that occurs on vowels after voiceless consonants that affects stress placement in Arabela and Karo. This could be the case in Karo where stress is retained on a final syllable that has a high tone even if it begins with a voiced consonant and the penultimate begins with a voiceless one. Hyman (2006b), though, notes that pitch does not seem to be a factor in the determination of stress in Pirahã and that the phonetic length of a voiceless consonant may indeed be making a contribution in the determination of stress.

Perhaps a perceptual basis for the weightlessness of syllable onsets can be found in Goedeman’s (1998) experimental study where using CVC synthetic stimuli he observed that listeners were less sensitive to duration fluctuations in the onset consonant than in either the vowel or coda consonant. However, as Goedemans notes, this may be an effect of onset weightlessness; that is, the weightlessness of onsets influences listeners’ ability to perceive durational contrast in the onset. See Gordon (2005) for discussion on Goedemans’ work.
This controversy is reflected in certain important response articles that argued against this underlying moraic representation of geminates such as Selkirk (1990b) which argued for a two root node theory of geminates and Tranel’s (1991) article which posited a principle of equal weight for codas whereby geminates were moraic or non-moraic depending on the patterning of other codas in the language. More recently, Ringen and Vago (2011) have argued for a universal segmental length representation of geminates as in (i) with no inherent weight properties:

(i) \[ C \ C \ \alpha \]

Ringen and Vago maintain that all evidence for a single root analysis of geminates as illustrated in (21d) is reanalyzable with structure like that in (i) and that there are phenomena that are only compatible with that in (i). A recent defense of the moraic theory of geminates can be found in Topintzi (2008). Other researchers have proposed composite representations for geminates such as Schmidt (1992) and Hume et al. (1997) who incorporate both an X-tier and a moraic tier in their analysis of geminates. Specifically, geminates are represented as a single phoneme linked to two X-slots and could be moraic in coda position if other codas in the language are moraic. Moreover, the controversy over geminates has fostered a number of dissertations which have a focus on the phonology of geminates. Some of the more important ones include Curtis (2003) who proposes a model of moraic representation that combines the two root node theory of geminates with moraic theory, Ham (1998), Keer (1999), Kraehenmann (2001), Morén (1999), Muller (2001), Sherer (1994), and Topintzi (2006). It is impossible in this chapter to discuss the wide variety of interesting issues and proposals that are raised in these dissertations.

See Davis (1994, 1999) for the range of processes considered to bear on the weight division in (22) and see Curtis (2003) and Ringen and Vago (2011) for critiques. Curtis (2003: 169–170) suggests that the lack of word-internal CVVG syllables in Koya may be due to a shortening effect that geminate consonants have on preceding vowels since the perceptual cues for vowel length can be blurred in CVVG syllables; thus, Curtis maintains that vowel shortening before geminates is independent of the issue of the moraic status of geminates.

Proponents of the moraic theory of geminates have suggested various strategies for such languages whereby geminate consonants do not seem to pattern as moraic. Topintzi (2008), who for the most part maintains the underlying moraic view of geminates, suggests that weightless geminates in a language like Selkup are represented by double consonants with two root nodes rather than as a single root node linked to a mora like (21d). Davis (2003) suggests that the stress pattern of languages like Selkup do not necessarily argue against the moraic theory of geminates, rather, from an optimality-theoretic perspective, the pattern can be a consequence of certain high-ranking stress constraints that have the effect of ignoring the bimoraicity of any CVC syllable. See Davis (2003: 95) for elaboration.

Hayes (1995a) notes that Cahuilla CVC syllables closed by glottal stops are also treated as bimoraic, though not other CVC syllables. In this light, it is worth recalling the Ngalakgan data in (15) which constitutes an almost opposite pattern to Cahuilla in closed syllables whereby syllables closed by geminates and those closed by a glottal stop are skipped over for stress and thus act as monomoraic while other CVC syllables...
(at least those where the coda has its own place features) are treated as bimoraic. In order to maintain the moraic theory of geminates, given the Ngalakgan stress data, one would need to maintain that underlingly Ngalakgan geminates are indeed moraic; that is they have the underlying moraic structure in (21d); they just do not surface as moraic because of a constraint that requires moraic elements to have their own place features. In this way, Cahuilla geminates would reflect their underlying moraic structure while Ngalakgan geminates do not. With respect to the issue in Cahuilla of a syllable closed by a glottal stop being treated as heavy, it may be possible to view this as reflecting a sonority threshold condition on what can be moraic. It is well-known that glottal type consonants can either pattern as highly sonorous elements or as obstruents though this has not been much discussed in the literature, but see Churma and Shi (1995) and Parker (2002). This variable behavior can be understood as to whether a language treats the laryngeal articulator as a place of articulation in the vocal tract or as just reflecting a particular state of the glottis. In the case of the latter, the laryngeal consonant would be a sonorant since there would be a free flow of air in the vocal tract. If this is the case for Cahuilla, then it would provide an interesting case where moraic codas are either highly sonorous consonants (glottal stops) or a geminate. Such a combination of moraic codas can be understood through Morén’s (1999, 2000) distinction between coerced weight and distinctive weight. See Note 33 for discussion on this distinction.

28 A very interesting case of a language in which CVV and CVG syllables pattern together with respect to stress is San’ani (Yemen) Arabic as described by Watson (2002: 81–82) who specifically notes their patterning together as opposed to CVC syllables. In this language, stress falls on the rightmost non-final CVV or CVG syllable in the word. If there are no such syllables (and ignoring a possible final superheavy syllable), then stress falls on the rightmost non-final CVC syllable up to the antepenultimate syllable; otherwise, it falls on the leftmost CV syllable. Thus, in words where there is a (non-final) CVC syllable and a (non-final) CVV or CVG syllable, it is either the CVV or CVG syllable that receives stress. The priority of CVV and CVG syllables in this language implies that CVC only acts as heavy in words in which there are no underlingly bimoraic syllables (CVV or CVG). That is, Weight-by-Position could only apply in a word that would otherwise have no bimoraic syllables. The patterning of CVG with CVV is best understood here if the geminate is underlingly moraic. The underlying moraicity of geminates in San’ani Arabic is further supported by Watson’s (2002: 82) observation that CVC syllables are never stressed in pre-antepenultimate position whereas both CVV and CVG syllables can be. See Watson (2002) for an analysis that incorporates the view that geminates are underlingly moraic. We suspect that syllables closed by geminates are special (and distinct from CVC syllables) with respect to weight properties in other dialects of Arabic as well, but this topic has yet to be fully explored.

29 The Hausa data in this section have been discussed with Paul Newman, the leading authority on the Hausa language (e.g. Newman (2000), though he disagrees with the analysis suggested here that the allomorph selection between (32a) and (32b) reflects an underlying weight distinction.

30 The Hausa plural pattern is made somewhat more complicated by the data in (i) that involve roots ending in a nasal followed by a homorganic stop.

(i) | Singular | Plural | Gloss |
---|---|---|---|
| a. kund-ii | kund-aaye | ‘notebook’ |
| b. gunt-uu | gunt-aaye | ‘stub’ |
Davis (1999) argues that these “partial geminates” can be considered underlyingly moraic and so can pattern with the roots in (37) with a geminate. (Davis (1995b) surveyed the weight behavior of partial geminates showing that depending on the language, they could pattern like geminates.) Newman (1992) argues for an analysis of (i) where the nasal is incorporated into the nucleus and so they would pattern with other roots having complex nuclei like in (35) and (36). Newman (electronic communication, September 3, 2007) further suggests that the different behavior of CVCCV singulars (e.g. 34 and 37) with respect to the plural allomorphy is due to the segmental characteristics of the abutting consonants rather than to underlying moraicitiy or other metrical factors.

31 Following a suggestion in Hayes (1995a), Davis (1999) proposes that word-initial geminates are moraic but that the mora is not part of the syllable onset. His representation is in (i) while Topintzi’s moraic onset representation is given in (ii) (where the vowel of the syllable is also shown).

\[
\begin{array}{c}
(i) \quad \sigma \\
\mu \\
\mu \\
c \quad v
\end{array}
\begin{array}{c}
(ii) \quad \sigma \\
\mu \\
\mu \\
c \quad v
\end{array}
\]

One difference between (i) and (ii) is that (ii) predicts that onset geminates could occur word-externally not just at the beginning of the word. In support of (ii) Topintzi (2008) provides interesting evidence from Marshallese that word-internal geminates are syllabified as onsets and are not heterosyllabic (7e) as commonly assumed in moraic theory.

32 José and Auger (2005) show that within a single language not all initial geminates pattern the same. According to them, in Vimeu Picard (phrase-)initial geminates differ in their representation as to whether they have a single set of features or two sets of identical features linked to two root nodes. Based on phonological patterning they argue that initial [ll] has the former representation while initial [nn] has the latter.

33 Of relevance to this topic is an important distinction that Morén (1999, 2000) makes between distinctive weight and coerced weight. Distinctive weight refers to underlying moraic structure that is reflected on the surface; geminate consonants, for example, would have distinctive weight. Coerced weight occurs when a non-moraic input segment surfaces as moraic. Context dependent weight can be seen as a type of coerced weight. However, coerced weight is more extensive than context dependent weight. For example, consider a language where Weight-by-Position always applies so that codas consistently surface as moraic. Such a language would have coerced weight since a non-moraic input consonant surfaces as moraic (in coda position), but it would not be an instance of context dependent weight since coda weight is not dependent on the makeup of the other syllables in the word. Morén makes an important observation that distinctive weight is not subject to sonority restrictions while coerced weight is. Thus, while Weight-by-Position may be restricted to sonorant consonants as discussed in Section 2, there are no universal sonority restrictions on distinctive weight; one does not find implications such that the presence of an obstruent geminate implies the presence of a sonorant geminate. This seems to be a correct observation about geminates as noted by Morén (1999), Blevins (2004) and implicit in Trubetzkoy’s
(1939) discussion of geminates, though see Kawahara (2007) concerning certain tendencies regarding which consonant types are more likely to be geminate.  

Klamath is another language cited in the literature (e.g. Hayes 1995a; Blevins 2006b) that has a weight hierarchy in which CVV syllables are heavier than CVC syllables which in turn are heavier than CV syllables. Stress falls on the rightmost heavy syllable in Klamath (with final CV and CVC syllables being extrametrical), but CVC syllables only count as heavy in words without long vowels. This can be analyzed in terms of contextual weight as in Rosenthal and van der Hulst (1999). See also Yu’s (2005) analysis of Washo stress and reduplication that makes crucial use of contextual weight of CVC syllables.  

The specific nature of final extrametricality varies among languages. For example, in Cairene Arabic it is just the word-final consonant that is extrametrical while in Latin it is the entire final syllable. Also languages vary as to whether and how they incorporate final extrametrical elements into higher prosodic structure. For pertinent discussion regarding the analysis of English and German word-final syllables see Hall (2002).  

More detailed discussion and analysis of the behavior and representation of preconsonantal nasals in Bantu languages can be found in Hyman and Ngunga (1997) and Downing (2005).  

A somewhat similar mismatch to that in Khalkha Mongolian can be found in the Uto-Aztecan language Tohono O’odham. Fitzgerald (forthcoming) documents that in Tohono O’odham CVC syllables behave as consistently bimoraic for the prosodic morphology of the language that includes processes of reduplication and gemination, but such syllables do not inherently attract stress. The language has a quantity-insensitive trochaic stress system that assigns alternating stress from the left edge of the word ignoring the apparent bimoraic nature of CVC syllables. CVV syllables in Tohono O’odham are restricted to word-initial position (at least in the native vocabulary), which is the location of primary stress.  

The occurrence of moraic mismatches brings up certain representational issues regarding moraic structure that are discussed by Broselow (1995) which I will not discuss here, other than to mention Hayes’s (1995a: 300) proposal of a moraic grid whereby sonorous (i.e. vocalic) moras would have two levels of grid marks and less sonorous moras would have one level. Certain processes like Haya tone or Khalkha Mongolian stress would make reference to the second level of grid marks while other processes such as minimal stem or word constraints would make reference to the lower level of grid marks. There are other ways of analyzing these mismatches in a constraint-based approach to phonology, which we do not discuss here.  

This suggests that lengthening due to phonetic factors is different from length that is phonological or distinctive. An interesting study that shows such a difference is Pycha (2007, 2009). She observes that Hungarian has two lengthening processes: a phrase-final process that lengthens a segment immediately adjacent to the phrase boundary and a morpho-phonological lengthening process whereby certain suffixes trigger gemination of a stem-final consonant. Pycha shows that when an affricate is targeted to be lengthened, the two processes implement the lengthening in different ways even though the overall duration is essentially the same. Specifically, with phrase-final lengthening which Pycha considers to be phonetic, the lengthening of the targeted affricate occurs mainly in the fricative part of the affricate, but in the morpho-phonological lengthening triggered by the suffix the lengthening mainly occurs in the closure part of the affricate.