3   Geminates

Stuart Davis

1   Introduction

The term "geminate" in phonology normally refers to a long or "doubled" consonant that contrasts phonemically with its shorter or "singleton" counterpart. Such contrasts are found in languages like Japanese and Italian, as exemplified by the minimal pairs in (1) and (2), respectively.[1] Languages such as English and Spanish do not have geminates.

(1)   Japanese geminate contrast (Tsujimura 2007)[2]

    a.   [saka]   'hill'
    b.   [sakka] 'author'

(2)   Italian geminate contrast

    a.   [fato]   'fate'
    b.   [fatto] 'fact'

The issue of the phonological representation of geminates has engendered much controversy over the past thirty years. The main issue revolves around how to distinguish formally a geminate consonant from its singleton counterpart in a way that captures the cross-linguistic phonological patterning of geminate consonants. The featural representation of geminate consonants posited in Chomsky and Halle

---

[1]   Languages with geminates vary considerably with respect to the durational difference between the geminate and its singleton counterpart. Idemaru and Guion (2008) report a 3:1 ratio in the duration of geminates to singletons in Japanese but only a 1.8:1 ratio for Italian. They further note that there may be other phonetic cues to geminates besides consonantal duration. These include pitch and intensity differences that may provide secondary acoustic cues to a geminate. However, this chapter will not focus on the phonetic properties of geminates, nor on the issue of which types of consonants are more likely to be geminated (but see Pycha 2007, 2009 and Kawahara 2007 for discussion on these issues). Instead, this article will focus on the phonological behavior of geminates and the matter of their representation in phonology.

[2]   In this chapter, geminate consonants are transcribed by a sequence of two identical letters; long vowels are represented either as a sequence of two identical vowel symbols or with the IPA length mark.

(1968) as being a single consonant possessing the distinctive feature [+long] has long been considered insufficient, since, as noted by researchers such as Leben (1980), long consonants can behave like a sequence of two consonants for certain phenomena. Leben posited an autosegmental representation of geminates in which a single phoneme is linked to two slots on a skeletal tier that encodes the prosody of the word. This skeletal tier is also referred to as a CV-tier, an X-tier, or a length tier depending on the specific conception of the researcher. Important earlier works that incorporate a CV-tier include McCarthy (1979, 1981), Halle and Vergnaud (1980), Clements and Keyser (1983) and Hayes (1986), while Levin (1985) posited that the tier consisted of X-slots (see CHAPTER 55: THE SKELETON). Geminate representation on this view is exemplified by the geminate [kk] of the Japanese word in (1b), as is illustrated in (3).

(3)   a.   CV-tier representation   b.   X-tier representation

C   V   C   C   V   X   X   X   X   X
| | \ / | | | \ / |
s   a   k   a   s   a   k   a
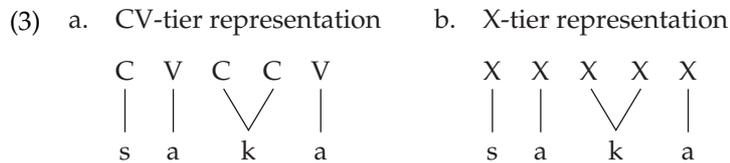
As seen in (3), a geminate consonant has one set of features indicated by the single consonant "k" on the phoneme (or melody) tier, whereas it is linked to two slots on a prosodic tier. In (4), we make clear the distinction between a geminate and a singleton using an X-tier that encodes prosody.

(4)   Prosodic length analysis of geminates

a.   X   X   b.   X
        \ / |
         k   k
     (geminate)   (singleton)
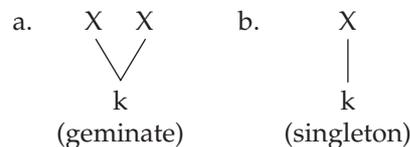
While the proposals for the representation of geminates in (4) go back thirty years, this representation is specifically argued for by Ringen and Vago (forthcoming), who refer to (4) as the segmental length analysis of geminates.

A different representation of geminates from that in (4) is the two-root node analysis of geminates posited by Selkirk (1990) shown in (5). The root node in a feature-geometric framework indicates the major class features of a sound (McCarthy 1988) and it dominates the rest of the specified features. Every phoneme has a root node, but a geminate under this view has two root nodes (RN = root node, c = consonant).

(5)   Prosodic length analysis of geminates

a.   RN RN   b.   RN
        \ / |
         c   c
   (geminate)   (singleton)
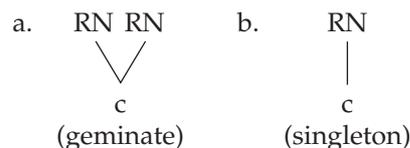
There are at least two main differences between the two-root node analysis of geminates in (5) and the segmental length analysis in (4). First, unlike the X-slots

in (4), a root node is not considered to be a prosodic unit. Second, the two-root node analysis can more readily capture certain phenomena whereby a single geminate splits into two phonemes, as in the case of Icelandic preaspiration: for instance, underlying /kappi/ 'hero' is realized as [kahpi]. (See Selkirk 1990 for a detailed discussion on how the two-root node theory captures this process, but also Keer 1998 for an optimality-theoretic analysis of Icelandic preaspiration that argues against the two-root node analysis.)

Probably the standard view of geminate representation in current phonological work is the moraic representation of geminates posited by Hayes (1989) and argued for in Davis (1994, 1999a, 2003) as well as by Topintzi (2008). On this view, geminates are represented to be underlyingly moraic or heavy, as shown in (6) (where UR = underlying representation, and μ = mora): a geminate does not have double linking, be it to two slots on the prosodic tier, as in (4a), or to two root nodes, as in (5a).

(6)   Moraic (weight) representation of geminates (Hayes 1989)

    a.   *Geminate in UR*   b.   *Single consonant in UR*

$$\begin{array}{cc} \mu & \\ | & \\ c & c \\ \text{(geminate)} & \text{(singleton)} \end{array}$$

This inherent weight approach to geminates is couched within the theory of moraic phonology as developed in Hayes (1989), which characterizes the prosodic tier as being moraic rather than segmental, as in (3). Specifically, in Hayes's theory of moraic phonology, a short vowel is underlyingly monomoraic while a long vowel is bimoraic; a geminate consonant differs from a short consonant in that the former is underlyingly moraic while the latter is non-moraic. Sample moraic representations are given in (7), where (7a) shows a short vowel, (7b) a long vowel, (7c) a singleton consonant and (7d) a geminate. We refer to the representation in (7d) as the weight analysis of geminates.

(7)   Underlying moraic weight representation (Hayes 1989)

    a. μ b. $\mu \quad \mu$ c.   d. μ

        a = /a/   a = /aː/   t = /t/   t = /tt/
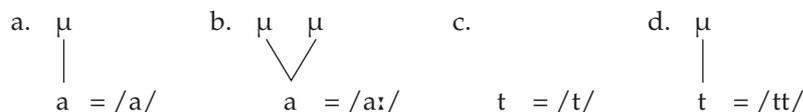
The moraic weight representation of geminates in (7d), where a single phoneme is linked underlyingly to a mora on the prosodic tier, is quite different from the length representation shown in (3), in which a single phoneme is linked to two C-slots (or X-slots) on the prosodic tier. These two different views of geminate representation make different predictions with regard to the patterning of geminates in phonology. For example, as noted by Ringen and Vago (forthcoming), if epenthesis is triggered by a word-final consonant cluster (i.e. a word ending in two C-slots), epenthesis would be predicted to occur in a word that ends in a final geminate since the word would end in two C-slots under the geminate representation in (3a). Ringen and Vago discuss Hungarian as a language with this

epenthesis pattern. On the other hand, if a geminate is represented as moraic, as in (6a) and (7d), epenthesis might not be predicted to occur with a word ending in a geminate, since the consonantal length of a geminate is not segmentally encoded. That is, there would not be two C-slots or two consonantal elements at the end of the word to trigger the epenthesis. Ringen and Vago point out that the Hungarian epenthesis pattern poses a problem for the moraic view. Further, given the weight analysis of geminates in (7d), geminate consonants are predicted to play a role in processes that are sensitive to syllable weight even when singleton (coda) consonants do not. Much of the recent research on geminates has focused on whether geminates display weight properties that are independent of other consonants. This will be discussed shortly.

Over the past twenty years a wide variety of phonological evidence has been brought to bear on the correct representation of geminates. The issue is still controversial.[3] All three views of geminate representation presented in this section, namely the prosodic length view in (4), the two-root node view in (5) and the moraic weight view in (6), have been argued for on the basis of the phonological patterning of geminates. Some composite views have even been proposed that combine aspects of the above representations, such as those of Schmidt (1992), Hume *et al.* (1997) and Curtis (2003). In §2 we will present specific evidence from a variety of phenomena to argue for the inherent weight representation of geminates. In §3, we will examine the behavior of geminates with respect to stress processes, cross-linguistically. In these sections, we will try to maintain a consistent view for the weight analysis in (6a) even when the data presented seem problematic for such a view. In §4 we will reconsider the representational issue and suggest that a composite view of the representation of geminates under a constraint-based approach can account for the patterning of geminates in the world's languages.

2   The weight analysis of geminates

The underlying weight analysis of geminate consonants, as proposed in Hayes (1989), views a geminate consonant as being underlyingly moraic as shown in (6a), whereas a non-geminate consonant is underlyingly non-moraic as in (6b). The weight representation of geminates in (6a) has a number of implications, which will be discussed in this section. One such implication is that if geminates are inherently moraic, they should count as moraic in considering minimal word effects: that is, the cross-linguistically common requirement that content words be at least bimoraic. In §2.1 we show that this is the case for Trukese. A specific structural aspect of the weight representation in (6a) is that geminates do not entail a double linking to two C-slots as in the length representation. This implies that there should be cases in which geminates do not pattern with a sequence of consonants. §2.2 discusses cases of the asymmetrical patterning of geminates and consonant

---

[3]   The controversy over geminates has fostered a number of dissertations with a focus on the phonology of geminates. Some of the important ones include Sherer (1994), Ham (1998), Keer (1999), Morén (1999), Kraehenmann (2001), Muller (2001) and Topintzi (2006). Although space does not allow me to discuss the wide variety of interesting issues and proposals that are raised in these dissertations and the different positions that are taken, some issues raised in these dissertations will be brought up in the course of this chapter.

sequences. A third implication that emerges from the weight representation in (6a) is the prediction that there should be languages that treat syllables closed by a geminate (CVG) as heavy but do not otherwise treat syllables closed by a (coda) consonant (CVC) as heavy. In §2.3, we will provide evidence for this prediction by discussing languages that avoid long vowels in syllables closed by a geminate (CVVG), but do not generally avoid long vowels in closed syllables (CVVC). We hold off until §3 the discussion of geminate behavior in weight-sensitive stress systems.

2.1 Trukese initial geminates

One type of evidence for the underlying moraic nature of geminates as in (7d) comes from the bimoraic minimal word requirement in Trukese (also called Chuukese) and the behavior of word-initial geminates with respect to it. Although word-initial geminates are rare, they are attested in a number of languages. (Indeed, the dissertations of Muller 2001 and Topintzi 2006 are exclusively on initial geminates; see also CHAPTER 48: INITIAL GEMINATES.)[4] Muller (2001), whose study incorporates acoustic analyses of word-initial geminates in a variety of languages, including Trukese, concludes that initial geminates are moraic in some languages but not in others, while Topintzi (2006, 2008), focusing on languages where initial geminates pattern as moraic, argues that such geminates constitute moraic onsets, thus providing examples in which onsets carry weight.[5] Trukese provides a clear example of a language where a word-initial geminate patterns as moraic. Consider the data in (8) and (9), which reflect a minimal word constraint on Trukese nouns. The data here are cited from Davis (1999b) and Davis and Torretta (1998), and are mainly taken from Dyen (1965) and Goodenough and Sugita (1980). The relevance of Trukese geminates for moraic phonology has previously been observed by Hart (1991) and Churchyard (1991).

---

[4] It is clear from typological surveys of geminate consonants such as Thurgood (1993) and from the discussion in Pajak (2009) that geminates are most commonly found in intervocalic position and least commonly found when not adjacent to any vowel (e.g. between two consonants). Languages that allow for geminates that are only adjacent to one vowel (e.g. word-initial or word-final geminates), although not common, are not as rare as languages that allow for geminates to occur not adjacent to any vowel. As noted by Pajak (2009), the typological facts correspond to perceptual saliency in that, the contrast between a singleton and a geminate consonant is most perceptually salient in intervocalic position and least salient in a position not adjacent to any vowel.

[5] Following a suggestion in Hayes (1989), Davis (1999b) 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).

(i)   σ   (ii)   σ

μ   μ   μ   μ

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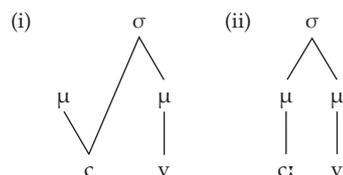
One difference between (i) and (ii) is that the latter predicts that onset geminates could occur word-internally, not just at the beginning of the word. In support of (ii), Topintzi (2008) provides interesting evidence from Marshallese that word-internal geminates are syllabified as onsets and are not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

(8)		*Underlying representation*	*Output form*		*Unattested output*
	a.	/maa/	[maa]	'behavior'	*[ma]
	b.	/təə/	[təə]	'islet'	*[tə]
	c.	/oo/	[oo]	'omen'	*[o]

(9) | | *Underlying representation* | *Output form* | | *Suffixed form* [n] = RELATIONAL |
|---|---|---|---|---|---|
	a.	/etiruu/	[etiru]	'coconut mat'	[etiruu-n]
	b.	/ttoo/	[tto]	'clam sp.'	[ttoo-n]
	c.	/ʧʧaa/	[ʧʧa]	'blood'	[ʧʧaa-n]
	d.	/ssɔɔ/	[ssɔ]	'thwart of a canoe'	[ssɔɔ-n]
	e.	/ffəne/	[ffən]	'advice'	[ffəne-n]

Trukese has a general process whereby a word-final long vowel shortens, as seen in (9a)–(9d). This is part of a more general process of final mora deletion, as evidenced in (9e). However, as (8) shows, final mora deletion does not apply if the result would be monomoraic, because Trukese has a minimal word constraint that requires nouns to be at least bimoraic. The fact that the word-final vowel does shorten in (9b)–(9d) strongly suggests that the initial geminate is moraic. That is, an output such as [tto] in (9b) is bimoraic, with a mora being contributed by both the vowel and the geminate. This is supportive of the underlying weight analysis of geminates.[6]

---

[6] While discussion on Trukese in works such as Muller (2001), Curtis (2003), Topintzi (2006) and Ringen and Vago (forthcoming) recognizes the inherent weight of Trukese geminates even though they do not all incorporate the underlying weight analysis of geminates (7d), these researchers have often contrasted the moraic behavior of initial geminates in Trukese with the clearly non-moraic behavior of initial geminates in Leti, as originally discussed in Hume *et al.* (1997). For example, although Leti has initial geminates. it lacks words consisting of an initial geminate followed by a short vowel such as [ppe]. Hume *et al.* maintain that the lack of such words argues against the moraicity of geminates, given the presence of a bimoraic minimal word condition. Following Davis (1999b, 2003), I maintain in this chapter that Leti is different from Trukese, because the initial geminates of Leti (but not Trukese) are extraprosodic, and that this is supported by the phonotactics of Leti. To see this, it is insightful to compare Leti geminates and word-initial clusters with those of Trukese. In Leti, underlying geminates occur only in word-initial position (Jennifer Muller, personal communication). In Trukese, in contrast, they occur in both word-initial and word-internal positions. Moreover, in Trukese, word-initial clusters other than geminates do not occur (with the exception of a few loanwords). On the other hand, word-initial clusters are pervasive in Leti, allowing for almost any possible sequence of two consonants at the beginning of the lexical word. There are no sonority restrictions on what these two consonants can be. The two consonants in a word-initial sequence can be an obstruent + sonorant, such as [pn pl pr tm tl tr vn vl vr], a sonorant + obstruent such as [mb ms mv ns rs rv], a sonorant + sonorant such as [mr nr rm rn rl], or two obstruents such as [pt tp pk kp tk kt]. Given this patterning, one could realistically 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. Such an analysis would explain the absence of Leti words like [ppe] or any other word of the shape CCV. With initial extraprosodicity, these forms would not comply with the bimoraic minimum. Given that underlying geminates only occur word-initially in Leti, and given the general phonotactics of word-initial clusters in Leti discussed above, I conclude that Leti presents a very different type of situation from Trukese where the geminate phonology is tightly integrated with the rest of the phonology (see Davis & Torretta 1998). I suggest that the Leti case has no bearing on the issue of the underlying representation of geminates. Leti allows initial extraprosodic consonants, and the apparent geminate is just a coincidental case where the extraprosodic consonant has the same quality as the following prevocalic consonant.

2.2   *Asymmetrical cases of the patterning of geminates and consonant clusters*

In the weight representation of geminates in (6a), repeated below in (10a), a consonant is underlyingly linked to a mora. This is contrasted with the length representation of geminates in which a consonant is linked to two X-slots or C-slots, as in (10b).

(10)   Contrasting representations for geminates

   a.   *Geminate in UR:*   b.   *Geminate in UR:*
        weight representation   *length representation*

$$\begin{array}{cc} \mu & \quad C \quad C \\ | & \quad \backslash \! / \\ c & \quad c \end{array}$$

        (geminate)   (singleton)

An important difference between the two representations is that the length representation in (10b) tacitly assumes that geminates should pattern similarly to a sequence of two consonants for rules or constraints that reference the CV-tier. Crucially, the weight representation in (10a) does not make such an assumption. There can be cases where geminates will pattern differently than a sequence of two consonants. An example where there is a parallel in patterning between a geminate and a sequence of two consonants is the case of Hungarian epenthesis discussed by Ringen and Vago (forthcoming) and noted earlier in the chapter. In Hungarian, in some verb stems that end in two consonants, an epenthetic vowel occurs after the two consonants when a consonantal suffix is added (e.g. /önt-s/ → [öntɛs] 'pour-2sɢ'). No epenthesis occurs if the verb stem ends in a single consonant when the suffix is added (e.g. /kɑp-s/ → [kaps] 'receive-2sɢ'). This suggests a constraint for these forms that disallows a word-final sequence of three C-slots. The two representations of geminates in (10) seem to make different predictions for verb stems that end in a geminate when the consonantal suffix is added. On the length view (10b), epenthesis would be predicted since there would be three consecutive C-slots, but such a prediction is not clear, given the weight representation in (10a). As Ringen and Vago observe, epenthesis does occur in a verb stem ending in a geminate (e.g. /fygg-s/ → [fyggɛs] 'depend-2sɢ'), which clearly shows that a geminate patterns like a sequence of two consonants with respect to the CV-tier. While this is not conclusive evidence for a length representation, since one could express the epenthesis rule as being sensitive to mora structure, given certain other assumptions about Hungarian, the analysis is more straightforward under the length representation of geminates.

   Nonetheless, there exist striking cases where geminates do not pattern like a sequence of C-slots. One such example comes from Trukese geminates, shown in (9) above. As seen before, Trukese has word-initial geminates, but does not have word-initial consonant clusters. One would think that if a language allows for a word-initial geminate, it should also allow for a word-initial sequence of two consonants under the length representation of geminates in (10b). Moreover, Trukese has word-internal geminates that are intervocalic (e.g. [tikka] 'coconut oil'), but Trukese does not generally allow for intervocalic consonant clusters. These

observations would be hard to account for under the length representation of geminates. One would expect that if two C-slots could occur at the beginning of the word or intervocalically, those two C-slots should not be restricted to just geminates. Furthermore, Trukese geminates do not pattern exactly like single consonants either. This is clearly seen in the observation that words in this language can end in a single consonant (as in the suffixed forms in (9)), but cannot end in a geminate. This contrastive behavior of geminates and consonant clusters on the one hand and geminates with singleton consonants on the other is consistent with and reflective of the weight analysis of geminates in (10a), especially in light of the fact that a word like [tto] in (9), comprising an initial geminate followed by a short vowel, meets the bimoraic word-minimality requirement. The presence of word-final singleton consonants and the absence of word-final geminates can be seen as a reflection of a high-ranked constraint that disallows words to end in a moraic consonant. A final singleton consonant would not be considered moraic in Trukese. In this regard, it is important to note that Trukese lacks CVC words. This is expected, given bimoraic word minimality if a word-final consonant is not moraic. Thus the patterning of Trukese geminates provides evidence for the weight representation in (10a) and against the length representation in (10b). (For further details and arguments for the weight representation of geminates in Trukese, see Davis and Torretta 1998.)

Another example in which geminates pattern differently from consonant clusters concerns final geminates in Arabic. In some Arabic dialects, such as the Hadhrami dialect as spoken in the town of Ghayl Bawazir near the south coast of Yemen (Bamakhramah 2009, personal communication), consonant clusters are avoided in word-final position (e.g. ['girid] 'a monkey' from underlying /gird/, ['binit] 'a girl' from underlying /bint/); yet word-final geminates are allowed (e.g. ['rabb] 'Lord', [ʔaˈxaff] 'lighter'). Moreover, word-final geminates in Hadhrami Arabic are different from singleton consonants in that a word-final geminate attracts stress onto the last syllable of the word, but a word that ends in a singleton does not have such impact on stress (e.g. [ʔaxáff] 'lighter' *vs.* [ˈʔakbar] 'greater') (see also CHAPTER 135: WORD STRESS IN ARABIC). Under the length representation of geminates as in (10b) it would be difficult to explain why word-final geminates are allowed when word-final consonantal sequences are avoided. Moreover, the attraction of stress onto the final syllable of a word ending in a geminate is consistent with the weight representation, given that, as observed by Bamakhramah (2009), primary stress typically falls on the rightmost bimoraic syllable. Consequently, the patterning of geminates in languages like Trukese and Hadhrami Arabic calls the length representation into question.[7]

---

[7] One matter that is historically important to note, but not focused on in the current chapter, concerns properties of geminates that were discussed in the literature in the 1980s in works such as Hayes (1986) and Schein and Steriade (1986) regarding geminate integrity and geminate inalterability. Geminate integrity is the observation that rules of epenthesis tend not to split up a geminate consonant (at least not a "true" geminate consonant, i.e. one that is monomorphemic and non-derived); geminate inalterability refers to the tendency that geminates are resistant to certain rules of segmental phonology (e.g. spirantization in Tiberian Hebrew) that a priori should apply to them. Hayes (1986), in particular, argued that the assumption of a CV-tier can account for integrity and inalterability effects. Integrity effects were accounted for by a length representation of geminates as in (10b), repeated below in (i.), because the insertion of a specific vowel into a geminate as in (i.) would violate the prohibition on crossing association lines.

2.3   Avoidance of CVVG syllables

Hayes's theory of underlying moraic representation was presented in (7): geminate consonants (7d) differ from singleton consonants (7c) in being underlyingly moraic. Furthermore, Hayes distinguishes long vowels from short vowels by representing the former as underlyingly bimoraic and the latter as monomoraic. Moraic theory has an implication for the patterning of geminates with respect to weight-sensitive processes – an implication not discussed by Hayes (1989), but taken up by other researchers such as Selkirk (1990), Tranel (1991) and Davis (1999a, 2003). That is, there should be languages in which syllables with a long vowel (CVV) and those closed by a geminate consonant (CVG) count as heavy since they would be bimoraic, while CV syllables and CVC syllables (i.e. syllable closed by a non-geminate consonant) would be considered light or monomoraic. This weight distinction is shown in (11) (G = geminate consonant, C = non-geminate consonant).

(11)   Syllable weight distinction based on geminates being underlyingly moraic

heavy	*light*
CVV	CV
CVG	CVC

The system in (11) is predicted to occur under Hayes's theory in any language with long vowels and geminate consonants that do not regard coda consonants as moraic. The moraic representation for syllables with the structure of (11) is given in (12).

(i)   *Geminate in UR: Length representation*

(ii)   *Epenthesis into a geminate*
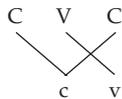
Geminate inalterability effects were handled by a condition on interpretation in segmental rules that association lines in structural description of rules had to be interpreted as exhaustive. A rule like spirantization in Tiberian Hebrew, which only applied to singleton consonants and not to geminates, would include in its rule environment a single C-slot linked to the phoneme. Since the rule environment did not explicitly show double linking as in (i.), the rule would fail to apply to geminates. Kenstowicz (1994: 410–416) summarizes important criticisms of the CV account of both geminate integrity and inalterability. He points out that geminate integrity could be called into question if epenthesis is viewed as a two-stage process of inserting a V-slot followed by a late default spell-out rule. However, it is worth noting that geminate integrity effects follow automatically from the weight representation of geminates as in (10a). With respect to geminate inalterability, Kenstowicz specifically calls attention to work by Selkirk (1991), who noted that rules of inalterability tended to always involve spirantization processes, thus suggesting a more general explanation that does not involve the length representation of geminates. Along these lines, Kirchner (2000) approaches the issue of geminate inalterability from a general theory of lenition within a functionally based optimality-theoretic framework.

(12)   Surface syllabification of the division in (11)

Heavy (bimoraic)

a.   σ   b.   σ   σ

                  μ   μ   μ   μ μ

        t   a = [taː]   t   a   t   a = [tat.ta]
                                                   (the first syllable is bimoraic)

Light (monomoraic)

c.   σ   d.   σ

              μ   μ

        t   a   = [ta]   t   a   t   = [tat]
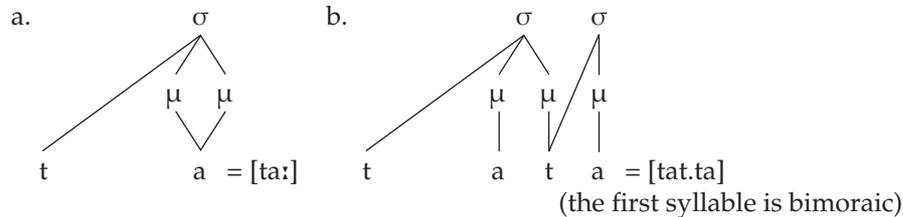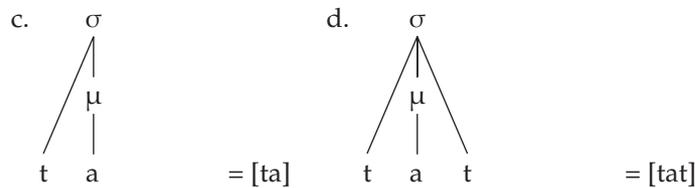
As seen in (12b), if a geminate is underlyingly moraic, the syllable closed by
a geminate (i.e. the first syllable of 12b) will be bimoraic, just like a syllable
containing a long vowel (12a). This should be contrasted with a syllable closed
by a non-geminate, as in (12d). The syllable in (12d) reflects the syllabification
in a language where the rule (or constraint) Weight-by-Position does not apply.
Weight-by-Position is the language-specific rule posited by Hayes (1989) that
makes a surface coda consonant moraic (see CHAPTER 57: QUANTITY-SENSITIVITY).
If the rule applies in a language, any closed syllable (CVC) in that language will
behave as bimoraic. If it does not apply, closed syllables should pattern as light
or monomoraic. However, because geminates are underlyingly moraic within
Hayes's theory in (7), then a syllable closed by a geminate (CVG) will always be
bimoraic, as in (12b), even if Weight-by-Position generally does not apply to codas.

   Under the assumption that the syllable weight distinction in (11) exists, as Hayes's
theory predicts, a specific prediction that emerges is that we would expect to find
languages where a long vowel in a syllable closed by a geminate (CVVG, where
G = the first part of a geminate) is avoided while a long vowel in a syllable
closed by a non-geminate is not (i.e. CVVC is allowed). Such avoidance of CVVG
syllables can be seen as a particular instance of a cross-linguistically common
phenomenon of avoiding trimoraic syllables (Prince 1990). When we have a lan-
guage that treats coda consonants as moraic or heavy, trimoraic syllables are often
avoided, as is the case with the dialectal Arabic examples in (13). (See Broselow
1992 and Kiparsky 2003 for overviews on Arabic syllables.)

(13)   Avoidance of trimoraic syllables in Arabic dialects

        a.   Cairene Arabic   /baab + na/ → ['bab.na]   'our gate'
        b.   Meccan Arabic   /baab + na/ → ['baa.ba.na]   'our gate'

Both Cairene and Meccan Arabic avoid the potentially trimoraic parse of the first
syllable of /baab + na/ as ['baab.na]. The dialects, however, differ in how they
avoid the trimoraic syllable. While Cairene Arabic favors closed-syllable shorten-
ing, Meccan Arabic preserves the underlying vowel length by having vowel

epenthesis apply between the two consonants to create open syllables, thereby avoiding any trimoraic syllable.

In the Arabic dialects, the rule (or constraint) of Weight-by-Position applies, resulting in a moraic coda. However, if we consider the weight division in (11) in which Weight-by-Position does not apply, we would expect to find a language where CVVG syllables but not CVVC syllables are avoided. A potential CVVG syllable would not surface, since it is trimoraic, while a CVVC could still occur, as it would only be bimoraic. Kiparsky (2008a) discusses Swedish dialects where vowel shortening occurs before a geminate but not before a single coda consonant. For example in West Swedish (Kiparsky 2008a: 191) /ruu-dde/ 'rowed' surfaces as [rudde], with its underlying long vowel shortened, but no shortening occurs when a long vowel is before a singleton coda consonant. The shortening before a geminate changes the potential trimoraic CVVG syllable to bimoraic CVG. There is no need for shortening in a CVVC syllable since it would be bimoraic, given that Weight-by-Position does not apply. Kiparsky specifically uses the Swedish data to argue for the moraic representation of geminate consonants.

Another language displaying this pattern of shortening is the Dravidian language Koya, brought up by Sherer (1994), based on Tyler (1969) and discussed in Davis (1999a, 2010). Koya has long vowels, coda consonants and geminate consonants. Sherer notes that there are words in Koya like those in (14a)–(14c), with a long vowel before a coda consonant. Crucially, as Tyler (1969: 6) observes, there are no words that contain a long vowel before a geminate. They are always short as in (14d). All Koya data are cited from Tyler (1969), with the page numbers provided. (The transcription of the vowel quality is phonemicized and does not reflect the precise allophonic variant.)

(14)   a.   leːŋga   'calf' (p. 11)   b.   aːɳɖa   'female' (p. 8)
       c.   neːrs   'learn' (p. 76)   d.   ett   'lift' (p. 76)

Sherer additionally notes cases where a stem-final long vowel shortens before a geminate-initial suffix, as the examples in (15) show.

(15)   a.   keː + tt + oːɳɖu   [kettoːɳɖu]   'he told' (p. 39)
       b.   oː + tt + oːɳɖu   [ottoːɳɖu]   '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 (16) illustrate.

(16)   a.   naːl + ke   [naːlke]   'tongue' (p. 47)
       b.   tuŋ + anaː + n + ki   [tuŋganaːŋki]   'for the doing' (p. 90)

In (16), a long vowel surfaces before a syllable-final singleton coda consonant. Since vowel shortening occurs before a geminate in (15), the Koya data in (14)–(16) are consistent with the weight system in (11), in which CVV and CVG syllables are bimoraic whereas CVC syllables are light.[8]

---

[8]   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. However, this does not explain cases like Fula in (19) where avoidance of CVVG is achieved by degemination rather than vowel shortening.

   While the above examples of Koya and Swedish are cases where vowel shortening occurs in syllables closed by a geminate, there are other 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, although coda consonants in general are not. Vowel lengthening then does not apply before a geminate since that would create a trimoraic syllable. This is illustrated by Seto (Southeastern Estonian), discussed by Kiparsky (2008b). According to Kiparsky, Seto has feet that are required to be trimoraic and such restriction is normally implemented by foot-final vowel lengthening. As a result, 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, i.e. CV.CVG, no vowel lengthening occurs. This provides evidence that the geminate is underlyingly moraic: that is, foot-final vowel lengthening need not occur in CV.CVG since the foot is already trimoraic.

   A different case 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 but, importantly, it allows for CVVC syllables as seen in (17).

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

   a.   kaakt-ɛ 'spittle'   b.   caak-ri 'couscous'

This language has a suffixation process that triggers the gemination of a root-final consonant. Consider the singular/plural alternations in (18). Because of an active constraint that requires geminates to be [−continuant] in Fula, a root-final continuant segment changes to a stop when it geminates. (I thank Abbie Hantgan for help with the Fula data.)

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

	stem (SG)	suffixed form (PL)	
a.	lɛw	lebb-i	'month'
b.	lɛf	lepp-i	'ribbon'

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. This is illustrated by the singular/plural alternations in (19).

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

	stem (SG)	suffixed form (PL)	expected_form	
a.	laaw	laab-i	*laabb-i	'road'
b.	lɛɛs	lɛɛc-ɛ	*lɛɛcc-ɛ	'bed'

Given that gemination is part of this suffixing process, we note that the expected forms in (19), where the initial syllable would be CVVG, fail to surface as such. Rather, the nature of the occurring suffixed forms in (19) makes it appear that degemination has occurred. This can be understood as the avoidance of a trimoraic CVVG syllable. Since CVVC syllables are allowed as in (17), Fula seems to

instantiate a language with the weight system of (11), where CVG syllables are heavy but not other CVC syllables. Thus, we see that a variety of languages have strategies to avoid CVVG syllables but not CVVC syllables. These languages can be understood as providing support for the inherent weight analysis of geminates.

3   The patterning of geminates in stress systems

One of the criticisms of the weight analysis of geminates proposed by Hayes (1989), discussed by Tranel (1991), comes from the observation that there do not seem to be quantity-sensitive stress systems that support the weight division in (11), repeated below as (20), where stress would be attracted onto a syllable with a long vowel or closed by a geminate consonant, but not on a syllable closed by a non-geminate.

(20)   Syllable weight distinction based on geminates being underlyingly moraic

heavy	*light*
CVV	CV
CVG	CVC

The system in (20) 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 generally apply to coda consonants. According to the division in (20), CVV and CVG syllables would syllabify as bimoraic, while CV and CVC syllables would syllabify as monomoraic, as was shown in (12). Since quantity-sensitive stress systems single out bimoraic syllables, it would be expected that at least some quantity-sensitive stress systems would reflect the weight division in (20) if the moraic representation of geminates were correct. Tranel suggests that weight systems like (20) do not exist and instead proposed a principle of equal weight for codas. Specifically, in languages in which codas pattern as moraic, geminates would be moraic; but in languages in which codas were not moraic, geminates would not be moraic. While our observation in §2.3 above – that in a variety of languages CVVG syllables are avoided but CVVC are not – can be taken as evidence for the weight division in (20), Tranel's observation is of importance. In this section, we will overview the behavior of geminates in quantity-sensitive stress systems. In §3.1, we will provide stress data from various languages which indeed support the division in (20) whereby CVV and CVG syllables pattern together, thus supporting the moraic weight analysis of geminates. These are languages whose stress patterns Tranel predicted not to occur. In §3.2, I will review the type of case mentioned by Tranel in which quantity-sensitive stress treats all closed syllables in the same manner whether they be CVG or CVC. These are the languages that motivated Tranel's principle of equal weight for codas and can be considered somewhat problematic for the weight analysis of geminates. In §3.3, I will present the case of the Australian language Ngalakgan (variably spelled Ngalagkan and Ngalakan; Baker 1997, 2008), in which CVC syllables can attract stress but apparently not CVG syllables. This seems to suggest that somehow geminates are resistant to carrying a mora. Thus, this section will identify three different types of geminate behavior with respect to quantity-sensitive

stress systems. In §4, we will try to reconcile this variability of behavior with respect to the representation of geminates.

3.1   Languages that uniquely treat CVG closed syllables as heavy with respect to stress

In this subsection I discuss two languages cited in the literature where syllables closed by a geminate (CVG) function as heavy, like a syllable with a long vowel in attracting stress, even when other closed syllables (CVC) may not act as heavy. Such languages provide stress evidence for the syllable weight division in (20). The two languages to be discussed are the Uto-Aztecan language Cahuilla and the San'ani (Yemen) dialect of Arabic.

Hayes (1995) noted that stress assignment in the Uto-Aztecan language Cahuilla distinguishes CVG syllables from CVC syllables: CVG syllables behave like CVV syllables in attracting stress. Consider the data in (21), taken from Hayes (1995, and see sources cited therein).

(21)   Cahuilla stress

a.	'ta.ka.ˌli.ʧem	'one-eyed ones'
b.	'ʧe.xi.ˌwen	'it is clear'
c.	'tax.mu.ˌʔat	'song'
d.	'he.ʔi.'ka.kaw.ˌlaː.ˌqa	'his legs are bow-shaped'
e.	'qaːn.ˌki.ʧem	'palo verde (PL)'
f.	'ʧex.ˌxi.wen	'it is very clear'

The data items in (21a)–(21d) help to establish Cahuilla as having iterative left-to-right stress on alternating syllables (i.e. trochaic foot structure), although CVC syllables are not distinguished from CV syllables.[9] (See also CHAPTER 42: THE FOOT and CHAPTER 46: THE IAMBIC-TROCHAIC LAW for more general discussion of trochaic foot structure.) The quantity-insensitive nature of CVC syllables is probably most clearly seen in the third and fourth syllables of the sequence in (21d), the trochaic foot ('ka.kaw). The second syllable of this foot is treated as light despite the presence of a coda consonant, which indicates the monomoraic nature of CVC syllables. On the other hand, the quantity-sensitive nature of the stress system can be seen not only in the stress-attracting nature of syllables with long vowels, but also in the observation that the syllable immediately after a long vowel receives stress. This is witnessed by the last two syllables in (21d), both of which have stress and comprise a foot, ('laː)('qa), and can be understood if Cahuilla foot structure is maximally bimoraic. Consequently, the last two syllables in (21d) cannot form a single foot, since such a foot would be trimoraic. The syllable with the long vowel ('laː) comprises a bimoraic foot on its own; the last syllable ('qa) is forced to comprise a (monomoraic) foot on its own, due to a constraint in the

---

[9]   Hayes (1995) notes that Cahuilla CVC syllables closed by glottal stops, but not other CVC syllables, are also treated as bimoraic. I shall not discuss this, other than to note that in the Ngalakgan stress pattern presented in (25) CVC syllables closed by a glottal stop exceptionally act as monomoraic. The variable behavior of coda glottal consonants with respect to syllable weight is known but infrequently discussed (but see Churma and Shih 1996).

language that requires exhaustive footing. The data item in (21e) is similar to the last two syllables in (21d) in that the syllable with a long vowel forms a foot on its own and the syllable immediately after it is at the beginning of a new trochaic foot, thereby receiving stress. We see then that bimoraic CVV syllables in Cahuilla are distinguished from monomoraic syllables, not only in bearing stress, but also by the presence of stress on the syllable immediately after them. CV and CVC syllables lack these two characteristics and function as monomoraic. It is interesting in this light to observe the patterning of syllables closed by a geminate, as in (21f), where the first syllable is CVG. This CVG syllable functions as bimoraic. It has stress, as would be expected of any initial syllable, but crucially it patterns exactly like a CVV syllable in that the syllable immediately after it also carries stress. This provides evidence that the CVG syllable comprises a bimoraic (trochaic) foot on its own. This constrasts with the initial CVC syllable in (21c) that forms a trochaic foot with the following syllable, suggesting the monomoraic nature of the CVC syllable. Cahuilla thus serves as a clear illustration for the weight distinction in (20) in which stress treats syllables with long vowels and those closed by geminates as bimoraic but not other types of syllables, be they CV or CVC.

San'ani Arabic (Yemen) presents a very interesting case in which CVV and CVG syllables pattern together with respect to stress. Watson (2002: 81–82) specifically notes that they pattern together as opposed to CVC syllables (see also CHAPTER 135: WORD STRESS IN ARABIC). Consider the data in (22), which illustrate the stress pattern in words without geminates.

(22)   San'ani Arabic (Watson 2002: 81–82)

     a.   mak.'tuːb   'office'
     b.   da.'rast   'I/you (MASC SG) learnt'
     c.   'saː.fa.rat   'she travelled'
     d.   miʁ.'saː.lih   'launderette'
     e.   mi.'gam.bar   'sitting'
     f.   'mad.ra.sih   'school'
     g.   mak.'ta.ba.tiː 'my library'
     h.   'li.bi.sat   'she wore/put on'
     i.   'ka.tab   'he wrote'
     j.   'ra.ga.ba.tih   'his neck'

Stress normally falls on one of the last three syllables of the word: it falls on a final superheavy syllable (CVVC or CVCC) if there is one, as in (22a)–(22b); it falls on the rightmost non-final heavy syllable (CVC or CVV) up to the antepenultimate, as in (22c)–(22f); otherwise, stress falls on the leftmost CV syllable, as in (22g)–(22j). The data in general show that the word-final segment does not play a role in the computation of weight so that the final syllable can only be stressed if it is superheavy. The word in (22g) illustrates two important aspects of the stress system. It shows that a word-final syllable ending in a long vowel does not attract the stress; it also indicates that a CVC syllable in pre-antepenultimate position fails to attract stress. The latter point is significant, since it suggests that Weight-by-Position, which assigns a mora to a coda consonant, is restricted to one of the last three syllables of the word. Now let us consider the data in (23) with words possessing geminate consonants.

(23)   San'ani Arabic (Watson 2002: 81–82): stress on words with geminate consonants

   a.   ji.'ħib.bu   'they (MASC) love/like'
   b.   mit.'ʔax.xi.raːt   'late (FEM PL)'
   c.   mu. 'sadʒ.dʒi.la.ti   'my recorder'
   d.   'haː.ka.ða.haː   'like this'
   e.   'daw.wart   'I/you (MASC SG) looked for'
   f.   'saː.fart   'I/you (MASC SG) travelled'

The comparison between (22) and (23) indicates the priority of CVG and CVV syllables for stress assignment in that CVG and (non-final) CVV syllables always attract stress even when in pre-antepenultimate position as in (23c) and (22d). The word in (22g), in contrast, shows that a CVC syllable does not receive stress in pre-antepenultimate position. The difference between CVG and CVC syllables can be readily understood on the inherent weight analysis of geminates. If a geminate is underlyingly moraic, it contributes weight to the syllable regardless of its location in the word. Recall that Weight-by-Position does not apply here, because it is restricted to one of the last three syllables in San'ani Arabic. In pre-antepenultimate position, only CVV and CVG act as bimoraic. Moreover, (23e) shows that CVG syllables have a priority of stress over a final superheavy syllable and should be compared with (22a) where a regular CVC syllable is devoid of such priority. It could be argued that Weight-by-Position in San'ani Arabic only applies to words that would not otherwise have bimoraic syllables (CVV or CVG). That is, there is no necessity for Weight-by-Position to apply in (23e) or (23f). While we do not pursue a full analysis here (but see Watson 2002), the priority given to both CVV and CVG syllables in stress assignment, especially as seen by the comparison of (23c) and (23d) with (22g), provides an interesting argument for the underlying moraic weight analysis of geminates, and, in turn, Tranel's (1991) claim of equal weight for codas.

   We have detailed above two cases where CVV and CVG syllables pattern together with respect to stress systems as predicted by the inherent weight analysis of geminates. Further support for the weight analysis of geminates is found in other languages. For example, Gupta (1987) discusses a Hindi dialect 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, although, as noted by Curtis (2003), such a pattern appears unusual among Hindi dialects. Additional support may come from the stress system of Pattani Malay, discussed by Topintzi (2006, 2008) and references cited therein. Pattani Malay has geminates that are restricted to word-initial position and the language lacks long vowels. Although primary stress typically falls on the final syllable of a word, stress occurs on the initial syllable in words that begin with a geminate consonant. This can be taken as evidence for the moraification in (10a) where a geminate is underlyingly moraic. That is, stress is attracted onto a syllable that is bimoraic.

   Despite the range of examples presented in this section, it remains rare to see languages that display the weight system in (20), grouping CVV and CVG syllables together as heavy. It is possible, on the other hand, that the rarity is due to the infrequent occurrence of the specific set of properties that is required for CVV

and CVG to pattern together in stress assignment; namely, the language would have to have quantity-sensitive stress, long vowels, coda consonants, and geminates. Perhaps when such languages, San'ani Arabic for one, are examined closely, more instances of the special properties of CVG syllables will emerge. In this connection, it is worth noting that in most Arabic dialects CVG syllables are special: they always attract stress when in word-final position. This property separates them from other CVC syllables, which do not attract stress in word-final position. The difference thus finds a logical explanation in the underlying weight analysis of geminate consonants (especially in those dialects, such as Hadhrami Arabic, discussed earlier, which disallow final consonant clusters).[10]

3.2   *Languages in which stress treats all codas equally*

There are two types of languages in which stress assignment treats all codas equally. In the first type, stress is quantity-sensitive and is attracted to a heavy syllable, be it CVV, CVC or CVG. Latin belongs to this group: any coda consonant makes a syllable heavy, so both CVC and CVG syllables are bimoraic. In the second type, which is more relevant for the representation of geminate consonants, both CVC and CVG syllables behave as light in a quantity-sensitive stress system. In such a language, stress is attracted to a CVV syllable, but both CVC and CVG syllables seem to pattern as monomoraic, treating CVG as light, just like other CVC syllables. As an illustration, consider the stress data from the Uralic language Selkup in (24). The data in (24a)–(24f) come from Halle and Clements (1983). The data items in (24g)–(24h) are reported in Ringen and Vago (forthcoming) from the Selkup language scholar Eugene Helimski, and reflect the Taz Selkup dialect, which seems to have the same stress pattern as that in Halle and Clements (1983).

(24)   Selkup stress

a.	qu'moːqi	'two human beings'
b.	'uːcɨqo	'to work'
c.	uː'cɔːmɨt	'we work'
d.	'qumɨnɨk	'human being (DAT)'
e.	'amɨrna	'eats'
f.	'uːcɨkkak	'I am working'
g.	'ɛsykka	'(it) happens (occasionally)'
h.	ɛs'sɪqo	'to happen (already)'

[10]   In many Arabic dialects, word-final CVC syllables behave as extrametrical. Ham (1998) puts forward the very intriguing observation that final CVC syllables are always extrametrical in languages that possess 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. With the underlying moraic weight representation of geminates as in (10a), 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 from a final CVC syllable (i.e. bimoraic), 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 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.)

In Selkup, primary stress falls on the rightmost syllable with a long vowel (24a)–(24c) or on the initial syllable if there are no long vowels (24d). A CVC syllable does not count as heavy (24e), even if the CVC syllable is closed by a geminate, as seen in (24f) and (24g). As noted by Tranel (1991), if stress targets bimoraic syllables and geminates are underlyingly moraic, the second syllable in (24f) and (24g) would be the rightmost bimoraic syllable. Both the vowel and the geminate would contribute a mora to the second syllable. The fact that (24f) and (24g) do not receive stress on the second syllable, however, seems to provide evidence against geminates being underlying moraic, favoring a representation of geminates that is different from that in (10a).

The stress pattern of Selkup does not appear to be unique in ignoring geminate consonants. Davis (1999a: 41) notes the Altaic language Chuvash (Krueger 1961), which exhibits an almost identical stress pattern to that of Selkup: stress is attracted to the rightmost syllable with a full vowel (interpreted as being bimoraic), but CVG syllables are ignored. Thus, in both Chuvash and Selkup, CVG syllables do not function as bimoraic CVV syllables but instead act like monomoraic CV and CVC syllables.

Data from languages like Selkup have been used by Tranel (1991) and Ringen and Vago (forthcoming) to argue against the underlying moraic weight representation of geminates. Ringen and Vago note that such languages are consistent with the length analysis of geminates as in (10b). In these languages, stress is sensitive to the presence of a long vowel, and ignores a coda consonant, whether the coda is part of a geminate or not. However, it is not that proponents of the weight representation are unaware of languages like Selkup. Topintzi (2008), who for the most part maintains the underlying moraic weight view of geminates, suggests that weightless geminates are represented by double consonants with two root nodes rather than as a single root node linked to a mora like (10a). But such a comment implies that there is language-specific variation in the representation of geminate consonants. On the other hand, Davis (2003) suggests that the stress pattern of languages like Selkup does not necessarily argue against the underlying moraic representation of geminates; viewed 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. As suggested by Steriade (1990: 275), there may be reasons in some languages to restrict the set of stress-bearing segments to those that are also tone-bearing, "for reasons that are clearly related to the fact that pitch is one of the main realizations of metrical prominence." Steriade's suggestion can be incorporated into an optimality-theoretic approach as a constraint that restricts pitch realization to vocalic elements: the constraint prefers to place stress on any CVV syllable over any syllable closed by a consonant, even if that consonant is part of a geminate. Thus, the lack of second syllable stress in (24f) and (24g) of the Selkup data need not reflect on the underlying moraicity of geminate consonants.

3.3   *Languages in which geminates repel stress*

A third type of geminate behavior is witnessed in languages where stress is attracted to a closed syllable, but not to one with a geminate. The Australian language Ngalakgan, discussed by Baker (1997, 2008), serves as a major example. Consider

the pattern of primary stress in (25), taken from Baker (2008) who notes that other geographically proximate languages have a similar stress pattern.

(25)   Ngalakgan

a.	pu'ruʈci	'water python'	k.	'calapir	'red ant (species)'
b.	ki'piʈkuluc	'frogmouth (bird)'	l.	'kupuj	'sweat (N)'
c.	mi'ʈarppuʔ	'crab'	m.	kaɲ'ʈalppuru	'plains kangaroo'
d.	pu'ʈolkoʔ	'brolga (bird)'	n.	'cakanta	'macropod sp.'
e.	ˌmaca'purka	'plant sp.'	o.	'ɲurunʈuc	'emu'
f.	'ḻaʈkurca	'vine sp.'	p.	'ɲoloŋkoʔ	'eucalyptus'
g.	'calpurkic	'fish sp.'	q.	'ɲamucˌculo	'subsection term'
h.	'ciwi	'liver'	r.	'capaʈʈa	'tortoise sp.'
i.	'ceraʈa	'women's ceremony'	s.	'moloppoḻ	'catfish sp.'
j.	'paʈaˌmunu	'sand goanna'	t.	'ɲanaʔpaj	'and moreover'

(25a)–(25l) show that primary stress in Ngalakgan falls on the leftmost (non-final) heavy syllable; if there is no heavy syllable, it falls on the initial syllable. From these data, it can be surmised that a coda consonant is moraic in making a syllable heavy, be it an obstruent (as in (25a) and (25b) or a sonorant (as in (25c)–(25g)). The data in (25m)–(25t) show that the leftmost closed syllable (underlined) fails to attract primary stress. Note that the coda in the leftmost closed syllable in (25m)–(25t) belongs to one of three types: in (25m)–(25p) the coda is a nasal homorganic with the following onset, in (25q)–(25s) it is the first part of a geminate consonant, and in (25t) the coda is a glottal stop. Key to our discussion is the fact that CVG syllables in (25q)–(25s) resist stress. However, the comparison of the stress-resistant nature of CVG syllables with the other instances of stress-resistant closed syllables in (25m)–(25p) and (25t) points to the fact that a common property these syllables have is that they do not possess their own place features: either the place features are shared with the following onset, or, in the case of the glottal stop in (25t), there is a lack of place features altogether. Thus, Ngalakgan seems to divide closed syllables into two types: those in (25a)–(25l), in which the coda has independent place features and attracts stress, and those in which the coda does not have its own independent place features and fails to attract stress. This suggests that Ngalakgan is best analyzed as having a requirement that moraic elements have independent place features (i.e. not shared with a following onset), as advocated in Baker (1997). It follows that the stressed closed syllables in (25a)–(25l) would be bimoraic and attract stress, whereas the CVG syllables in (25m)–(25p) would be monomoraic and not attract stress.

   Languages like Ngalakgan seem to present a challenge for the underlying weight representation of geminates in (10a), since not only do syllables with geminates not attract stress, but they are not even equal in weight to CVC syllables, as in (25a)–(25l), which do attract stress. Baker (2008) offers an articulatory gestural analysis of the difference. He observes that the apparent CVC syllable attracts stress only if the postvocalic coda consonant has an articulatory gesture distinct from that of the following onset. That is, in a CVCCV sequence, the first syllable counts as heavy only if the two intervocalic consonants have distinct articulatory gestures. When the intervocalic sequence involves a geminate or a nasal that is homorganic to a following consonant (or a glottal stop, which does not have a distinct articulatory gesture) there is only one articulatory gesture, and stress is

not attracted onto CVG (or CVN where N is homorganic to a following con-
sonant). Baker (2009) adopts a composite view of geminate representation that
incorporates a gestural tier along with root nodes and moras. Under this view,
stress in Ngalakgan is characterized as sensitive to the gestural tier. Ringen and
Vago (forthcoming), in contrast, take the Ngalakgan data as supporting the length
representation of geminates where the stress rule treats linked structures like those
in (10b) as light. Davis (2003), who maintains the underlying weight representa-
tion of geminates, suggests that the Ngalakgan stress pattern in (25) does not
necessarily argue against the underlying moraic representation of geminates as
in (10a); rather, the language has a high-ranked constraint that requires moraic
elements to have their own place features. Thus, while geminates may be under-
lyingly moraic, they do not surface as moraic.[11]

     To conclude this section, we have surveyed languages demonstrating three types
of behavior of CVG syllables in stress systems: (i) cases where CVG and CVV
pattern together; (ii) cases where CVG patterns with other CVC syllables; and
(iii) cases where CVG syllables are specifically resistant to stress. We have tried
to maintain the underlying moraic weight representation of geminates despite appar-
ent evidence to the contrary. In the concluding section we will further discuss
representational issues.

4   Representational issues and conclusion

In this overview, we have focused on the cross-linguistic patterning of geminate
consonants while trying to maintain the representational view of geminates in (10a),
in which geminates are marked as being underlyingly moraic, over the length
representation in (10b). In §2, we provided evidence for the underlying moraic
representation of geminate consonants by considering a variety of phonological
patterning pertinent to geminates. This included the moraic analysis of initial
geminates in Trukese in §2.1 and the cross-linguistic avoidance of CVVG syllables
in §2.3. We made it clear in §2.2 that geminates do not always behave like a
sequence of two C-slots in prosodic patterning, thereby contradicting the length
representation of geminates in (10b). In §3, we surveyed geminate patterning in
stress systems identifying three types of behavior. Despite the differences, we still
argued for the underlying moraic view of geminate consonants.

     The issue of the representation of geminate consonants has been a controversial
matter and will most likely remain so in future investigations. This is because
geminates do not display uniform behavior, as we have illustrated. It seems that
the very nature of the data under examination determines what type of repre-
sentation must be appropriate. For example, the parallel patterning between final
CC sequences and final geminates in Hungarian seems to be supportive of the
length representation, while the difference between final CC sequences and final

---

[11]   It should be noted that Baker (2008) actually considers the intervocalic geminates in (25q)–(25s)
and the intervocalic homorganic nasal clusters in (25m)–(25p) to be syllabified completely as onsets
rather than as heterosyllabic. This differs from his earlier work, Baker (1997), where a heterosyllabic
parse of geminates and homorganic nasal clusters is maintained. One shortcoming of Baker's (2008)
onset analysis is that geminates and homorganic nasal clusters do not occur word-initially. Nonethe-
less, even if geminates could be analyzed as syllabifying as onsets, they would not add weight to a
syllable and thus would be different from the initial geminates of Trukese discussed earlier.

geminates in Arabic dialects such as Hadhrami Arabic appears to be better accounted for by the weight analysis of geminates. Such differences lead to three possibilities for a representational view of geminates, in a general sense. The first possibility is that geminate representation is language-specific. In a language of the Trukese type, geminates are represented to be underlyingly moraic (10a), as argued for by Davis and Torretta (1998) and Davis (1999b), but in a language of the Hungarian type, geminates have a length analysis (10b), as argued for by Ringen and Vago (forthcoming).[12]

The second possibility – one maintained in the body of this chapter – is that there is one specific universal representation of geminates (i.e. the underlying moraic weight representation), and apparent counterexamples to it are explained by constraints that act on surface forms. This strategy was emphasized especially in §§3.2 and 3.3. For example, in the Selkup-type stress pattern in (24), which ignores all CVC syllables, including those closed by a geminate, there is an independent constraint that restricts pitch realization to vocalic elements. The constraint would choose to place stress on any CVV syllable over any syllable closed by a consonant, even if that consonant were part of a geminate. Thus, the lack of CVG syllables that attract stress in Selkup need not reflect on the underlying moraicity of geminate consonants. Given that geminates show a varied degree of sensitivity to stress, as seen in §3, any single universal analysis of geminates would have to make use of such flexible strategies to account for the apparent problematic cases.

A third possibility for a representational account of geminates is to maintain that there is a universal representation of geminates, but one that combines different aspects of the various representations discussed in §1 and elsewhere in this chapter. Hume *et al.* (1997) combine a length representation with moraic theory, and Curtis (2003) integrates the two-root node representation presented in (5) with a moraic view. (See Curtis 2003 for a detailed discussion and comparison of different representational views of geminate consonants.) A more recent composite view is implicit in Baker (2009) in which a geminate is represented on both a timing (length) tier and a gestural tier. It is also viewed as having a moraic representation if it functions as heavy. Baker's (2009) representation of a geminate [k] is shown in (26).

(26)   (μ)
         |
        [k]   Gestural tier
        /\
       X   X   Timing (Length) Tier
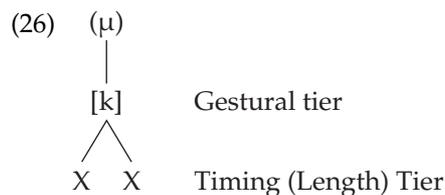
Baker's analysis is quite attractive as a universal, especially if we assume the underlying moraic nature of geminates. Languages display variable patterning for geminates because the phonology (i.e. the constraints or rules) makes reference

---

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

to different aspects of the representation. For example, in Trukese, the moraic aspect of the geminate representation is crucial, while it may or may not surface with two X-slots, depending on the constraints. In Hungarian, given the parallel behavior of word-final clusters and geminates, the X-tier representation is effective, while the geminate may or may not surface as moraic, depending on the constraints. In Ngalakgan, the behavior of geminates with respect to stress can be understood through the gestural tier, as proposed by Baker (2008, 2009). Geminates and homorganic nasal + consonant clusters have a single gesture. Consequently, they can pattern as single consonants despite having two X-slots, as long as there is a high-ranked constraint that requires a moraic element not to share place features. This would provide an explanation for the repulsion of geminates to stress. It follows that while geminates may have one underlying universal representation (as in (26)), its surface realization may vary cross-linguistically, e.g. as non-moraic in Ngalakgan, but moraic in Trukese.

The composite view, or some version of it, may ultimately be the best universal representation for an underlying geminate. For example, one criticism of a purely weight account of geminates is that it cannot distinguish between a geminate that syllabifies entirely in a coda from a single coda consonant in a language in which Weight-by-Position applies. This matter comes up in the Palestinian Arabic dialect described by Abu Salim (1980) and mentioned in Rose (2000) whereby a coda singleton in a word like [bit.na] 'our house' is representationally indistinguishable from a coda geminate in a word like [sitt.na] 'our grandmother' on a strictly moraic view of geminates as in (10a). Unless a length tier (or two root nodes) is assumed, there is no obvious way to distinguish the two cases. Although such examples are probably rare, the occurrence of this type of contrast indeed favors a composite analysis, especially given the language-specific evidence for Arabic moraic structure presented in various parts of this chapter. That said, it may still be possible to argue for the underlying moraic weight representation as universal, but with the understanding that the surface realization of geminates may vary across languages because of the interaction of relevant constraints.

In conclusion, there is much about the phonology of geminates that remains to be investigated. Geminates do not all pattern the same way across languages. Consequently, geminate phonology will remain an area of theoretical controversy for the foreseeable future.

ACKNOWLEDGMENTS

REFERENCES

Abu Salim, Issam. 1980. Epenthesis and geminate consonants in Palestinian Arabic. *Studies in the Linguistic Sciences* 10. 1–11.

Baker, Brett. 1997. Edge crispness: Segment to mora isomorphism. *Proceedings of the West Coast Conference on Formal Linguistics* 16. 33–47.

Baker, Brett. 2008. *Word structure in Ngalakgan*. Stanford: CSLI.

Baker, Brett. 2009. Monogestural clusters as onsets: The Australian evidence. Paper presented at the 83rd Annual Meeting of the Linguistic Society of America, San Francisco.

Bamakhramah, Majdi. 2009. Syllable structure in Arabic varieties with a focus on super-heavy syllables. Ph.D. dissertation, Indiana University.

Broselow, Ellen. 1992. Parametric variation in Arabic dialect phonology. In Ellen Broselow, Mushira Eid & John J. McCarthy (eds.), *Perspectives on Arabic linguistics IV*, 7–45. Amsterdam & Philadelphia: John Benjamins.

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper & Row.

Churchyard, Henry. 1991. Compensatory lengthening and "gemination throwback" in Trukese and Puluwat as evidence for Rime in moraic phonology. Ms., University of Texas.

Churma, Donald & Yili Shi. 1996. Glottal consonants and the "sonority" hierarchy. *Proceedings of the Eastern States Conference on Linguistics (ESCOL) 1995*. 25–37.

Clements, G. N. & Samuel J. Keyser. 1983. *CV phonology: A generative theory of the syllable*. Cambridge, MA: MIT Press.

Curtis, Emily. 2003. Geminate weight: Case studies and formal models. Ph.D. dissertation, University of Washington.

Davis, Stuart. 1994. Geminate consonants in moraic phonology. *Proceedings of the West Coast Conference on Formal Linguistics* 13. 32–45.

Davis, Stuart. 1999a. On the moraic representation of underlying geminates: Evidence from prosodic morphology. In René Kager, Harry van der Hulst & Wim Zonneveld (eds.), *The prosody–morphology interface*, 39–61. Cambridge: Cambridge University Press.

Davis, Stuart. 1999b. On the representation of initial geminates. *Phonology* 16. 93–104.

Davis, Stuart. 2003. The controversy over geminates and syllable weight. In Féry & van de Vijver, 77–98.

Davis, Stuart. 2010. Quantity. In John Goldsmith, Jason Riggle & Alan Yu (eds.), *The handbook of phonological theory*, 2nd edn., 000–000. Malden, MA: Blackwell.

Davis, Stuart & Gina Torretta. 1998. An optimality-theoretic account of compensatory lengthening and geminate throwback in Trukese. *Papers from the Annual Meeting of the North East Linguistic Society* 28. 111–125.

Dyen, Isadore. 1965. *A sketch of Trukese grammar*. New Haven: American Oriental Society.

Féry, Caroline & Ruben van de Vijver (eds.) 2003. *The syllable in Optimality Theory*. Cambridge: Cambridge University Press.

Goodenough, Ward & Hiroshi Sugita. 1980. *Trukese–English dictionary.* Philadelphia: American Philosophical Society.

Gupta, Abha. 1987. Hindi word stress and the obligatory branching parameter. *Papers from the Annual Regional Meeting, Chicago Linguistic Society* 23. 134–148.

Halle, Morris & G. N. Clements. 1983. *Problem book in phonology*. Cambridge, MA: MIT Press.

Halle, Morris & Jean-Roger Vergnaud. 1980. Three-dimensional phonology. *Journal of Linguistic Research* 1. 83–105.

Ham, William. 1998. Phonetic and phonological aspects of geminate timing. Ph.D. dissertation, Cornell University.

Hart, Michele. 1991. The moraic status of initial geminates in Trukese. *Proceedings of the Annual Meeting, Berkeley Linguistics Society* 17. 107–120.

Hayes, Bruce. 1986. Inalterability in CV phonology. *Language* 62. 321–351.

Hayes, Bruce. 1989. Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20. 253–306.

Hayes, Bruce. 1995. *Metrical stress theory: Principles and case studies*. Chicago: University of Chicago Press.

Hume, Elizabeth, Jennifer Muller & Aone van Engelenhoven. 1997. Non-moraic geminates in Leti. *Phonology* 14. 371–402.

Idemaru, Kaori & Susan Guion. 2008. Acoustic covariants of length contrast in Japanese stops. *Journal of the International Phonetic Association* 38. 167–186.

José, Brian & Julie Auger. 2005. Geminates and Picard pronominal clitic allomorphy. *Catalan Journal of Linguistics* 4. 127–154.

Kager, René. 1989. *A metrical theory of stress and destressing in English and Dutch*. Dordrecht: Foris.

Kawahara, Shigeto. 2007. Sonorancy and geminacy. *University of Massachusetts Occasional Papers* 32. 145–186.

Keer, Edward. 1998. Icelandic preaspiration and the moraic theory of geminates. *Proceedings of the 10th Conference of Nordic and General Linguistics, Reykjavik*. (ROA-312.)

Keer, Edward. 1999. Geminates, the OCP and the nature of Con. Ph.D. dissertation, Rutgers University.

Kenstowicz, Michael. 1994. *Phonology in generative grammar*. Cambridge, MA & Oxford: Blackwell.

Kiparsky, Paul. 2003. Syllables and moras in Arabic. In Féry & van de Vijver, 147–182.

Kiparsky, Paul. 2008a. Fenno-Swedish quantity: Contrast in Stratal OT. In Bert Vaux & Andrew Nevins (eds.) *Rules, constraints, and phonological phenomena*, 185–219. Oxford: Oxford University Press.

Kiparksy, Paul. 2008b. Weight and length. Presented at the CUNY Conference on the Syllable, New York City.

Kirchner, Robert. 2000. Geminate inalterability and lenition. *Language* 76. 509–545.

Kraehenmann, Astrid. 2001. Quantity and prosody asymmetries in Alemannic: Synchronic and diachronic perspectives. Ph.D. dissertation, University of Konstanz. Published 2003, Berlin & New York: Mouton de Gruyter.

Krueger, John. 1961. *Chuvash manual: Introduction, grammar, reader, and vocabulary*. Bloomington: Indiana University Press.

Leben, William R. 1980. A metrical analysis of length. *Linguistic Inquiry* 11. 497–509.

Levin, Juliette. 1985. A metrical theory of syllabicity. Ph.D. dissertation, MIT.

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

McCarthy, John J. 1981. A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry* 12. 373–418.

McCarthy, John J. 1988. Feature geometry and dependency: A review. *Phonetica* 45. 84–108.

Morén, Bruce. 1999. Distinctiveness, coercion and sonority: A unified theory of weight. Ph.D. dissertation, University of Maryland at College Park.

Muller, Jennifer. 2001. The phonology and phonetics of word-initial geminates. Ph.D. dissertation, Ohio State University.

Pajak, Bozena. Forthcoming. Contextual constraints on geminates: The case of Polish. *Proceedings of the Annual Meeting, Berkeley Linguistics Society* 35. (ROA-1051.)

Paradis, Carole. 1988. On constraints and repair strategies. *Linguistic Review* 6. 71–97.

Prince, Alan. 1990. Quantitative consequences of rhythmic organization. *Papers from the Annual Regional Meeting, Chicago Linguistic* Society 26(2). 355–398.

Pycha, Anne. 2007. Phonetic vs. phonological lengthening in affricates. In Jürgen Trouvain & William J. Barry (eds.) *Proceedings of the 16th International Congress of Phonetic Sciences*, 1757–1760. Saarbrücken: Saarland University.

Pycha, Anne. 2009. Lengthened affricates as a test case for the phonetics–phonology interface. *Journal of the International Phonetic Association* 39. 1–31.

Ringen, Catherine & Robert M. Vago. Forthcoming. Geminates: Heavy or long? In Charles Cairns & Eric Raimy (eds.) *Perspectives on the syllable*. Leiden: Brill.

Rose, Sharon. 2000. Rethinking geminates, long-distance geminates, and the OCP. *Linguistic Inquiry* 31. 85–122.

Schein, Barry & Donca Steriade. 1986. On geminates. *Linguistic Inquiry* 17. 691–744.

Schmidt, Deborah. 1992. Compensatory lengthening in a segmental moraic theory of representation. *Linguistics* 30. 513–534.

Selkirk, Elisabeth. 1990. A two-root theory of length. *University of Massachusetts Occasional Papers* 14. 123–171.

Selkirk, Elisabeth. 1991. On the inalterability of geminates. In Pier Marco Bertinetto, Michael Kenstowicz & Michele Lopocaro (eds.) *Certamen Phonologicum II: Papers from the 1990 Cortona Phonology Meeting*, 187–209. Turin: Rosenberg & Sellier.

Sherer, Timothy. 1994. Prosodic phonotactics. Ph.D. dissertation, University of Massachusetts, Amherst.

Steriade, Donca. 1990. Moras and other slots. *Formal Proceedings of Formal Linguistics Society of Midamerica* 1. 254–280.

Thurgood, Graham. 1993. Geminates: A cross-linguistic examination. In Joel Ashmore Nevis, Gerald McMenamin & Graham Thurgood (eds.) *Papers in honor of Frederick H. Brengelman on the occasion of the twenty-fifth anniversary of the Department of Linguistics, CSU Fresno*, 129–139. Fresno: Department of Linguistics, California State University, Fresno.

Topintzi, Nina. 2006. Moraic onsets. Ph.D. dissertation, University College, London.

Topintzi, Nina. 2008. On the existence of moraic onsets. *Natural Language and Linguistic Theory* 26. 147–184.

Tranel, Bernard. 1991. CVC light syllables, geminates and moraic theory. *Phonology* 8. 291–302.

Tsujimura, Natsuko 2007. *An introduction to Japanese linguistics.* 2nd edn. Cambridge, MA: Blackwell.

Tyler, Stephen. 1969. *Koya: An outline grammar*. Berkeley: University of California Press.

Watson, Janet C. E. 2002. *The phonology and morphology of Arabic*. Oxford: Oxford University Press.