On the Weight of Edge Geminates*

Nina Topintzi (Aristotle University of Thessaloniki) and Stuart Davis (Indiana University)

Edge geminates (EGs) are of a different nature than intervocalic geminates. They are rarer and structurally different; they emerge – at least superficially – as tautosyllabic within an onset (word-initial geminate) or coda (word-final geminate), as opposed to the typically heterosyllabic intervocalic geminates. In this paper, we present an initial typology of the weight properties of EGs and make observations that may predict whether an EG patterns as heavy or light. For the latter part, we consider the relationship between EGs and edge consonant clusters in the languages under consideration and investigate the existence of correlations. An initial finding suggests that if EGs are unique in a language, i.e. the language possesses no edge clusters, then the geminate is more likely to pattern as moraic (cf. Trukese and Pattani Malay in initial position and Hadhrami Arabic finally). Additionally, weightless EGs seem to co-occur with weightless clusters at the same edge. Weight asymmetries between initial and final geminates in the same language are also attested. We provide tentative thoughts as to why the typology is shaped the way it is.

Keywords: edge geminates, weight, edge clusters, typology, asymmetries

1 Introduction

Initial and final geminates, collectively referred to as edge geminates (EGs), differ from intervocalic geminates in various ways. A predominant difference relates to their structural representations. While intervocalic geminates are typically heterosyllabic, EGs are tautosyllabic – at least superficially – either within an onset (word-initial geminate) or a coda (word-final geminate).

Three aspects of the typology of geminates have received so far the most attention: (i) the typology of geminate positioning, (ii) the typology of geminate consonant preference, and (iii) the typology of geminate weight. With regard to positioning (Thurgood 1993; Pajak 2010; Kraehenmann 2011; Dmitrieva 2012), the general finding is that intervocalic is

* We are grateful to two anonymous reviewers for helpful comments and interesting ways to re-state the universals we propose here, as well as to the audience of the MIT Linguistics Colloquium, and especially Donca Steriade, for useful feedback. All remaining errors are our own.
the preferred location for geminates with edge geminates being rarer. For the latter in particular, the consensus is that word-final geminates are more common that word-initial geminates (Thurgood 1993). In an experimental study with Russian, English, and Italian speakers though, Dmitrieva (2012) found that word-initial geminates were more perceptually distinct than word-final geminates. Interestingly, word-final geminates seem to presuppose medial geminates, but no such correlation exists between initial and medial geminates (Dmitrieva 2012 citing Taylor 1985 and Muller 2001). Languages like Pattani Malay (§3.1.1) and Leti (§3.2) famously possess initial geminates in the absence of geminates elsewhere. The same holds for Ngada, Yapese, and Nyaheun (Kraehenmann 2011). In other languages, e.g. Trukese (Davis, this volume), both initial and medial geminates are attested, while final ones are banned.

Findings on the typology of geminate consonant type preference are often contradictory (Thurgood 1993, Morén 1999; Kawahara 2007; Kraehenmann 2011). For example, Kawahara (2007) points to the avoidance of geminate sonorants in many languages while Morén (1999) maintains that there are no universal preferences (i.e. implicational universals) regarding geminate consonant type. In terms of frequency of occurrence, Ladefoged and Maddieson (1996) observe that geminate nasals and geminate voiceless stops are the most frequently occurring typologically. In word-initial position geminate stops are the most frequently occurring (Kraehenmann 2011: 1129, using Muller’s 2001 data).

The typology of geminate weight, discussed in detail in Davis (2011) (who assumes moraic theory (Hayes 1989)) reveals three major patterns: (a) Languages in which geminates pattern like other coda consonants with respect to syllable weight in either being both moraic (Latin, Lake Miwok) or nonmoraic (Selkup, Tübatulabal). This is what Tranel (1991) has called the Principle of Equal Weight for Codas. The two other patterns have been discovered in later work and rebut this principle. (b) Languages in which geminates always pattern as moraic, even when other codas may not, as in Koya, Seto, San’ani Arabic, or Cahuilla (Davis 2003, 2011 and references cited therein) and (c)

1 By medial geminates, we refer to word-medial intervocalic geminates spanning syllable boundaries. It is beyond the scope of the present work to discuss other interesting, but much rarer, types of medial geminates, especially of the potential EG-type, as in Fennoswedish [venn.da] ‘to turn’ (Kiparsky 2008) or Marshallese [ji.bbun] ‘morning’ (Topintzi 2008).
Languages in which geminates are not weight-bearing even when other coda consonants are moraic. The best example of this type to date is illustrated by Ngalakgan (Baker 1997). Crucially in that language moraic elements must have their own place features, a fact that excludes geminates.

Evidently, the focal point for the typological work on geminates has been intervocalic geminates and occasionally to a lesser degree edge geminates. The weight typology of EGs however has not been sufficiently explored (but see Muller (2001) and Ham (2001) for some general work on initial and final geminates, respectively). This is an endeavour we presently undertake. In addition, we consider any potential relationship between EGs and edge clusters, a task that becomes all the more meaningful, given proposals that have treated geminates as two C-slots on the consonantal tier (Ringen and Vago 2011), and for that reason, comparable to CC-clusters. Driven by such proposals, we will in particular compare the weight behaviour of EGs to that of edge clusters, since in the moraic theory of weight (Hayes 1989) – which here and elsewhere we have adopted (e.g. Davis 1999, 2003; Topintzi 2008, 2010) – geminates are defined through their possession of underlying weight. Thus, while strictly speaking, clusters – as a reviewer correctly points out – have no special status in moraic theory, the examination of their weight properties in relation to the corresponding properties of geminates offers a more encompassing account that simultaneously takes into consideration various, often competing, components in the analysis of geminates.

The questions we seek to answer are summarized as follows.

(1) **Typology of Edge Geminates**

(A) With respect to weight, do EGs pattern as moraic (heavy) or not?

(B) What is the relationship between EGs and edge clusters? Languages with initial geminates may or may not have initial consonant clusters. Similarly, languages with final geminates may or may not have final clusters.

   i. In languages that have EGs and edge clusters, do the geminates and clusters pattern the same way with respect to weight or can geminates be special?

   ii. In languages that have EGs but no corresponding edge clusters, are the edge geminates more likely or less likely to be moraic (heavy)?

   iii. In languages that have both initial and final geminates, how do the EGs pattern with respect to weight?
In the remainder of the paper we discuss the weight typology of final (Section 2) and initial (Section 3) geminates in conjunction to clusters and we exemplify with several languages and data. Notice that while our survey contains – what we think is – a representative sample of languages, not all languages with EGs reported in the literature are included (see for example Muller 2001 or Dmitrieva 2012 for a list of relevant languages). We believe that other languages displaying EGs merely exemplify one of the patterns otherwise discussed. Nonetheless, our findings should be taken as preliminary. In Section 4, we summarize the results of the typology and offer two implicational universals that can be succinctly stated as follows (“edge” refers to the same edge, either right or left edge, respectively): (I) If edge clusters do not exist, then edge geminates are moraic, and (II) If edge geminates are non-moraic, then edge clusters are non-moraic. We then discuss further implications and present some preliminary thoughts as to how we can understand the generalizations we have uncovered more formally. Finally, in Section 5, we offer brief concluding remarks.

2 Final geminates

We begin our discussion of final geminates with an instantiation chart that summarizes the weight status of the geminate (heavy vs. light), whether it can co-occur with a cluster and if so, what type of cluster in terms of weight. We then examine each pattern briefly.

(2) Instantiation chart: final geminates

<table>
<thead>
<tr>
<th>Final Geminates</th>
<th>Final Clusters</th>
<th>Moraic</th>
<th>Non-moraic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moraic</td>
<td>(I) Baghdadi Arabic, Hadhrami Arabic</td>
<td>(II) Swiss German, Cairene Arabic, Ponapean, Wolof</td>
<td>(III) Amharic San'ani Arabic</td>
</tr>
<tr>
<td>Non-moraic</td>
<td>(IV) ---</td>
<td>(V) ---</td>
<td>(VI) Hungarian, Tashlhiyt Berber</td>
</tr>
</tbody>
</table>

The first three patterns refer to languages with moraic geminates. Pattern (I) comprises languages with final moraic geminates but lacking final consonant clusters. Pattern (II) includes languages that have both final moraic geminates and final consonant clusters that pattern as moraic. Pattern (III), where there is a co-occurrence of moraic final
geminates and non-moraic final clusters, does not seem to be well-attested, though Amharic and San'ani Arabic are possible examples. For final non-moraic geminates the situation seems much more restricted with only one pattern arising, that of geminates and clusters that are both non-moraic.

2.1 Combinations of moraic final geminates and clusters

2.1.1 Pattern I: Languages with moraic final geminates and no final clusters

Pattern I is illustrated by some Arabic dialects such as Hadhrami Arabic (Bamakhramah 2009) and Baghdadi Arabic (Blanc 1964, Youssef 2013). The moraicity of geminates is evinced through quantity-sensitive stress and the bimoraic word minimum. Consonant clusters are avoided in final position and repaired through vowel epenthesis. Word-final geminates on the other hand remain intact. Relevant data from the Hadhrami dialect as spoken in the town of Ghayl Bawazir near the south coast of Yemen illustrate (Bamakhramah 2009, p.c. to second author).

(3) Hadhrami Arabic: Stress facts and final epenthesis in clusters vs. lack thereof in geminates

a. /gird/ [gírid] ‘monkey’ (cf. [gírd-i] ‘my monkey’)
b. /bint/ [bínit] ‘girl’ (cf. [bínt-i] ‘my girl’)
c. [rább] ‘Lord’
d. [ʔaxáff] ‘lighter/lightest’

The data item in (3d) is indicative of the moraicity of the final geminate since a final CVC syllable would not typically be stressed. A similar effect is reported for the Muslim and Christian varieties of Baghdadi Arabic (Blanc 1964).

(4) Baghdadi Arabic: Stress facts and final epenthesis in clusters vs. lack thereof in geminates

a. /ward/ [wáred] ‘flowers’  b. /sabt/ [sábet] ‘Saturday’
c. [sádd] ‘he shut’  d. [yendázz] ‘he will be sent’

2.1.2 Pattern II: Languages with moraic final geminates and moraic final clusters
In Pattern II, final geminates co-occur with final clusters, both of which carry weight. A handful of languages illustrate this pattern. In Swiss German (Kraehenmann 2001: 113) monosyllabic words must be bimoraic. To fulfil this requirement, /CVC/ words undergo vowel lengthening and become [CV:C], as opposed to /CVCC/ words which surface unchanged. This happens whether a final cluster or a geminate is involved.

Contrary to Hadhrami and Baghdadi Arabic discussed in the previous section, in Cairene Arabic (e.g. Davis and Ragheb 2014), word-final consonant clusters exist and pattern alongside word-final geminates in being moraic, as a wide range of phenomena demonstrate. Single final consonants, however, lack weight.

(5) Cairene Arabic: Stress
a. Final clusters and final geminates
   [maf.hímʃ] ‘he didn’t understand’   [ka.ta bouts] ‘I wrote’
   [ʔaż.xáff] ‘lightest’              [ʔa.máll] ‘most boring’

b. CVC syllables
   [mak.táb.na] ‘our office’          [ká.tab] ‘he wrote’
   [mu.hán.dis] ‘engineer’

As (5a) illustrates, final clusters and geminates always attract stress on the syllables that host them, whereas CVC syllables (b) do so in penultimate but not in final position (compare [ka.ta.buts] with [ká.tab]). This can easily be accounted for if final CVCs are treated as light monomoraic, but CVCC and CVG as bimoraic, thus heavy. This idea is further corroborated by the word minimality data. Cairene bans monomoraic CVC content words, but admits CVCC or CVG words due to their bimoraicity.

(6) Cairene Arabic: Minimal Word
   [ʔult] ‘I said’                     [ʔism] ‘name’               [nuSS] ‘half’

A similar effect is highlighted by the adaptation of monosyllabic loanwords from English. When these contain a final consonant cluster, then they are imported unchanged (7a). Loanwords of the CVC type in the source language though, are often modified and borrowed with a final geminate (7b). Such adaptation is not applicable in longer loanwords, cf. [bi.lás.tik] for ‘plastic’.
(7) **Cairene Arabic: Loanword adaptation**

a. CVCC loanwords

<table>
<thead>
<tr>
<th>English</th>
<th>Cairene</th>
</tr>
</thead>
<tbody>
<tr>
<td>film</td>
<td>film</td>
</tr>
<tr>
<td>bank</td>
<td>bank</td>
</tr>
</tbody>
</table>

b. CVC loanwords

<table>
<thead>
<tr>
<th>English</th>
<th>Cairene</th>
</tr>
</thead>
<tbody>
<tr>
<td>(seven) up</td>
<td>?abb</td>
</tr>
<tr>
<td>book (purse)</td>
<td>bukk</td>
</tr>
</tbody>
</table>

Ponapean (Goodman 1995, Kennedy 2003) has both final clusters and geminates, albeit of specific varieties, namely homorganic nasal obstruent (NC) clusters and sonorant geminates. Their shared behaviour as weighted elements is highlighted by the fact that long vowels are admitted before word-final singletons (8a), but not before geminates and NC clusters (Goodman 1995: 67-68), suggesting that maximally bimoraic syllables are permitted (8b).

(8) **Ponapean: Syllable weight maximum**

a. CV:C

<table>
<thead>
<tr>
<th>[koos]</th>
<th>‘bent, bumped’</th>
</tr>
</thead>
</table>

b. CVCC or CVG

<table>
<thead>
<tr>
<th>[mall]</th>
<th>‘clearing in the forest’</th>
</tr>
</thead>
<tbody>
<tr>
<td>[romʷmʷ]</td>
<td>‘calm’</td>
</tr>
<tr>
<td>[mʷanč]</td>
<td>‘late’</td>
</tr>
</tbody>
</table>

Wolof (Bell 2003) demonstrates a pattern similar to Ponapean, the difference being that the only final clusters allowed are of the nasal+voiceless obstruent type. Bell considers both nasal+voiceless obstruent sequences and geminates to be underlyingly moraic in Wolof. Like in Ponapean, long vowels in Wolof never occur before final geminates nor before nasal+voiceless obstruent clusters (at least in monosyllabic words), though long vowels can occur before a single consonant. This can be understood as an avoidance of trimoraic syllables. Also, neither geminates nor nasal+voiceless obstruent sequences can occur word-initially, which is consistent with their common patterning. Bell (2003) distinguishes nasal+voiceless obstruent sequences from prenasalized consonants, which always have a voiced release in Wolof. The prenasalized consonant can occur at the beginning of a word, and, when in word-final position, can be preceded by a long vowel. It has to be noted though that stress assignment is indifferent to the presence of a geminate or a cluster. While stress normally falls on the initial syllable, it shifts to the second
syllable if that contains a long vowel and is preceded by a syllable with a short vowel. As Bell argues, in Wolof, stress is only sensitive to vocalic moras.

2.1.3 Pattern III: Languages with moraic final geminates and non-moraic final clusters

Unlike the previous pattern, this one is not as robustly attested. Nonetheless, possible exemplifications are to be found in Amharic and San'ani Arabic. While Amharic stress is described as not being that prominent, Sande and Hedding (2014) present evidence for stress – shown in bold – being assigned by trochaic footing from the left edge of the word, with no stress on the final syllable (9a). In words with geminates though, stress is attracted to the syllable that hosts the geminate as part of the coda. This holds for all geminates, including those found word-finally (9b). The stress-attracting effect however does not arise when the syllable ends in a consonant cluster (9c). Hannah Sande generalizes this to syllable final clusters at the end of the word, but has not provided a specific example (p.c. to first author, 27/11/2014). Consequently, Sande and Hedding (2014) analyse geminates as weight-bearing in contrast to final consonant clusters.

(9) Amharic: Stress

a. (mät'.fat) ‘to vanish’ (do.ro) ‘chicken’
   (mät.räf)räf ‘to overflow’ (k'o.fi)ja ‘hat’
   (mäʃ.k'ä)(da.däm) ‘to race’ (as.da)(ka.käl)ku ‘I arranged (my schedule)’

b. mä(tʃäm.mär) ‘to add an ingredient to something’
   tä(gag.ga)(räf.tä)(wal.šit) ‘she will bake them’
   (bäl)(laʃ.tʃi) ‘y’all ate’
   sej.(tʃuʃ) ‘women’

This presupposes a distinction between vocalic and consonantal moras, an idea made explicit in Davis and Torretta (1998). The idea is tacitly shared by Topintzi (2010) in the discussion of Pirahã weight-sensitive stress, where it is proposed that the weight of nuclei matters more than overall weight. More generally, Gordon (2006) discusses several cases where different weight criteria may be applicable for different phenomena. For example, in Lhasa Tibetan both stress and tone are weight-sensitive. For stress purposes, only (C)VV syllables count as heavy, whereas for tone both (C)VV and (C)VC syllables with sonorant codas do.
Another language in which final geminates are distinguished from final clusters in being inherently moraic is San'ani Arabic, as described by Watson (2002). As in other dialects of Arabic, in San'ani when words end in a geminate consonant stress always falls on the final syllable, e.g. [ʔa.hámm] ‘important (comparative)’. With respect to words ending in a final consonant cluster, stress only falls on a final syllable if the word does not otherwise have heavy syllables. Compare [da.rást] ‘I/you (m.s.) studied’ having final stress with [dáw.wart] ‘I/you (m.s) looked for’ and [sá..fart] ‘I/you (m.s) travelled’ where the final cluster does not act as heavy. Normally, if a word has two heavy syllables, the rightmost one attracts the stress. This is why a word like [dáw.wart] ‘I/you (m.s) looked for’ is revealing. We thus consider San'ani Arabic as an example where final geminates are always moraic but final clusters need not be.

2.2 Combinations of non-moraic final geminates and clusters (Patterns IV-VI)

While we find languages in which the moraicity of final geminates are robustly attested regardless of the patterning of final clusters, our preliminary survey reveals fewer patterns of non-weight-bearing final geminates. In particular, we discovered no instances of languages that lack final clusters but have final geminates that are non-moraic, i.e. weightless. Moreover, we found no languages where final clusters act as moraic without the final geminates exhibiting the same behavior. Although we found no such language in our survey, in the light of Ngalakgan (Baker 1997), where, recall from Section 1, medial geminates are weightless although other coda consonants are moraic, such a pattern cannot be eliminated. Consequently, given our typological chart in (2), we found no languages that instantiate Patterns IV and V. Obviously, further research is needed to determine whether these gaps are systematic or accidental.

With respect to Pattern VI – languages with non-moraic final geminates and non-moraic final clusters – Hungarian is a reasonable candidate given the analysis of Ringen and Vago (2011) (also Siptár and Törkenczy 2000). Evidence comes from the observation that neither final geminates nor final clusters attract stress. Further, final geminates and final consonant clusters pattern together in that both trigger epenthesis when a suffix initial coronal consonant is added. The epenthesis process would be more difficult to
capture if geminates were assumed to be moraic. Consequently, there does not seem to be concrete language-internal evidence that requires such consonants to be moraic and as such we deem Hungarian fit for this Pattern.

An even more complicated case is that of Tashlhiyt Berber which possesses edge geminates and edge clusters in both initial and final position. This Berber language has been the subject of a series of studies by Dell and Elmedlaoui (1985, 1988, 2002) and more recently by Ridouane (2007, 2008) and Ridouane et al. (2014). Tashlhiyt is known for the occurrence of words containing only obstruents as shown below in (10a,b) and where obstruents can be syllabic. Given that words can begin and end in a geminate and that even an obstruent can be syllabic, the question arises as to how the edge geminates pattern with respect to edge clusters in Tashlhiyt. The answer to this question is not simple. We suggest here that Tashlhiyt falls into the pattern where neither the edge cluster nor the EG is inherently moraic. That is, with respect to final clusters it is like Pattern VI and with respect to initial clusters it is like Pattern XII (cf. (11)).

This is consistent with Ridouane (2007) who argues against a moraic analysis of Berber geminates. In Tashlhiyt words containing a string of obstruents, one (or more) of the obstruents is pronounced as syllabic. There is an algorithm for assigning syllabic status to such a consonant. The algorithm is fairly well understood both in the rule-based approach of Dell and Elmedlaoui (1985) and in the constraint-based analysis of Prince and Smolensky (1993/2004). What has not been formally addressed is how geminate consonants (especially obstruents) interact with the syllabification algorithm. Consider the data in (10) drawn from Dell and Elmedlaoui (1985, 1988), Ridouane (2007, 2008), Ridouane et al. (2014), and Ridouane (p.c. to second author, 10/12/2014).

(10) **Tashlhiyt Berber** (. indicates syllable boundary; caps indicate a syllabic consonant)

a. tF.s\*t
   ‘you cancelled’
b. tF.tKt
   ‘you suffered a sprain’
c. tQs.sf
   ‘it shrunk’
d. tK.kSt
   ‘you took off’
e. ta.zN.k”Tt
   ‘female gazelle’
f. ttsX.xan
   ‘dip (in sauce)’
g. a.lS
   ‘repeat’
h. ass
   ‘day’
i. G.li
   ‘guide’
The data in (10a-b) demonstrate that in words with strings of obstruents, certain obstruents are made syllabic depending on their location in the string in accordance with the analysis in Dell and Elmediaoui (1985) or Prince and Smolensky (1993/2004). The syllabic obstruent can be a fricative or a stop. (10c-d) show that in words with a string of obstruents that include a geminate in medial position, the geminate can either surface as nonsyllabic (10c) or as syllabic (10d) depending on the nature of the consonant sequence. Note that in (10d) the first part of the geminate surfaces as syllabic. When an obstruent becomes syllabic it acquires a mora since it is in the peak position of the syllable; otherwise, an obstruent is not moraic, not even a geminate one as Ridouane (2007) has maintained for Tashlhiyt.

The data items in (10e-f) show the behaviour of an edge obstruent geminate when it is adjacent to another obstruent. As seen in (10e), the final edge geminate can become syllabic, but an initial geminate (10f) resists becoming syllabic. The remaining data (10g-j) demonstrate the behaviour of an edge geminate or cluster when next to a vowel. In this situation, the geminate – unlike the cluster – does not seem to be syllabic (Rachid Ridouane, p.c.). More specifically, when the vowel is followed by a CC-cluster word finally, the second consonant becomes syllabic (10g), but if it is followed by a word-final geminate (10h), then the geminate remains nonsyllabic. Similarly, at the left edge of the word, if two consonants precede the vowel, the first one becomes syllabic as in (10i), but an initial geminate in the same context does not become syllabic (10j). Initial geminates then resist becoming syllabic in general but final geminates can become syllabic if preceded by an obstruent (as seen in (10e)).

In relating the Berber pattern to our proposed typology, we would tentatively put Berber with the language patterns where neither geminates nor edge clusters act as inherently moraic. Geminates in medial and final position can become syllabic in a way similar to other consonants, but the geminate is somewhat more resistant to becoming syllabic. Initial geminates never seem to become syllabic, while final geminates do not become syllabic when immediately following a vowel. We do not see these differences between the syllabification of the EG with that of an edge consonant cluster as reflecting on their possible moraicity, but rather we would maintain that the algorithm or constraint ranking that determines the syllabicity of consonants needs to take into consideration the resistance of the geminate to become syllabic. Thus, there may be a
high-ranked constraint that prevents geminate consonants from becoming moraic. That said, we leave for future research whether there is independent evidence for the possible moraicity of geminates or edge clusters in Berber.

3 Initial geminates

Turning now to initial geminates, we provide a similar instantiation chart to that in (2) for final geminates and consider the combinations arising between initial geminates and initial clusters.

(11) **Instantiation chart: initial geminates**

<table>
<thead>
<tr>
<th>Initial Geminates</th>
<th>Initial Clusters</th>
<th>Moraic</th>
<th>Non-moraic</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>Moraic</td>
<td>(VII) Pat. Malay, Trukese, Woleaian, Tedumuni Okinawan, Luganda</td>
<td></td>
</tr>
<tr>
<td>Moraic</td>
<td>(VIII) Cypriot Greek, Ponapean</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-moraic</td>
<td>(IX) Shuri Okinawan</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(X) ---</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(XI) ---</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(XII) Leti, Swiss German, Baghdadi Arabic, Tashlhiyt Berber</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.1 Combinations of moraic initial geminates and clusters

3.1.1 Pattern VII: Languages with moraic initial geminates and no initial clusters

Pattern VII comprises languages with moraic initial geminates but no initial consonant clusters. This pattern is well-attested. In particular, two of the four languages discussed presently (Pattani Malay and Trukese), have received increased attention in recent years. Pattani Malay (Yupho 1989; Hajek and Goedemans 2003; Topintzi 2008, 2010) only allows geminates in initial position (12). At the same time, no consonant clusters are allowed in that position. The moraicity of geminates is highlighted by both compensatory lengthening and stress.

<table>
<thead>
<tr>
<th>Singletons</th>
<th>Geminates</th>
</tr>
</thead>
<tbody>
<tr>
<td>tido</td>
<td>‘to sleep’</td>
</tr>
<tr>
<td>labo</td>
<td>‘to profit’</td>
</tr>
<tr>
<td>yato</td>
<td>‘comprehensive’</td>
</tr>
<tr>
<td>sepa?’</td>
<td>‘to kick’</td>
</tr>
</tbody>
</table>

Initial geminates may also arise as a product of a process whereby the initial syllable reduces and leads to a free variant where the – original – second onset geminates as in e.g. *buwi ~ wi ‘give’, sidadu ~ d:adu ‘police’, pimatɔ ~ matɔ ‘jewellery’* (Yupho 1989: 130). The same effect appears in words with a /Ci-/ derivational or verbal prefix followed by the stem (13). This phenomenon can be understood in terms of compensatory lengthening; in [bidiyi] for example, the mora of the first syllable bi is preserved after the deletion of its segmental material and emerges through gemination of the following onset consonant, i.e. d:. Effectively, the total number of moras between the two variants remains the same, i.e. [bii’diiyi] ~ [diiyi].

(13) Pattani Malay: *Geminates in prefixed forms* (Yupho 1989)

<table>
<thead>
<tr>
<th>Unprefixed form</th>
<th>Prefixed form</th>
<th>Geminate variant</th>
</tr>
</thead>
<tbody>
<tr>
<td>jale ‘road, path’</td>
<td>bijale</td>
<td>~</td>
</tr>
<tr>
<td>buwɔh ‘fruit’</td>
<td>bibuwɔh</td>
<td>~</td>
</tr>
<tr>
<td>diyi ‘self’</td>
<td>bidiyi</td>
<td>~</td>
</tr>
<tr>
<td>kaij ‘no gloss’</td>
<td>miŋai</td>
<td>~</td>
</tr>
</tbody>
</table>

Turning to the stress data, primary stress normally is assigned to the last syllable, with all preceding syllables receiving secondary stress (14i.a), unless they contain the weak schwa-like vowel [i] which remains stressless (14i.b). In words with geminates though, primary stress is word-initial, whereas secondary stress is assigned to any syllables following that (14ii.a). Most interestingly, this occurs even if the first vowel is the weak i, as in (14ii.b).


i) Words lacking geminates

a. jale ‘road path’
i) Words with initial geminates

   a. mátò ‘jewellery’
   jálè ‘to walk’

   b. kídà ‘to the shop’

A straightforward analysis of these facts, as put forward in Topintzi (2008), takes Pattani stress to be partly quantity- and partly quality-driven, with the quantity requirements being prioritized. When all syllables are light, primary stress is aligned to the right edge and secondary stress reveals the quality effect, i.e. all full vowels receive stress, save the weak i. When the first – the only possibility in Pattani – syllable is heavy though due to the presence of a geminate, primary stress appears in that position overriding the quality demands that otherwise require i to be stressless.

Trukese or Chuukese also displays initial geminates and lacks initial clusters. That geminates are weightful can be concluded through the examination of the word minimality facts as well as through a complex pattern of compensatory lengthening dubbed geminate throwback (Davis and Torretta 1998, Davis 1999, Muller 1999). Since Trukese is discussed at length in Davis (this volume), we presently only mention the word minimality facts and point the interested reader to that paper for further information.

(15) Trukese: Word minimality

<table>
<thead>
<tr>
<th>Form</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. CV: words</td>
<td></td>
</tr>
<tr>
<td>ma:</td>
<td>‘behaviour’</td>
</tr>
<tr>
<td>o:</td>
<td>‘omen’</td>
</tr>
<tr>
<td>té:3</td>
<td>‘islet’</td>
</tr>
<tr>
<td>nú:</td>
<td>‘unripe coconut’</td>
</tr>
<tr>
<td>b. GV words</td>
<td></td>
</tr>
<tr>
<td>t:o</td>
<td>‘clam sp.’</td>
</tr>
<tr>
<td>k:a</td>
<td>‘taro sp.’</td>
</tr>
<tr>
<td>č:a</td>
<td>‘blood’</td>
</tr>
</tbody>
</table>

3 ē=[ʌ], ú=[ı] (Davis and Torretta 1998: 112, fn. 2).
Trukese permits minimal words with a single long vowel (CVː) or with a short vowel preceded by a geminate (GV), but no CV or CVC words. This distinction can easily be explained if Trukese sets a bimoraic word minimum and treats CVː and GV words as bimoraic, whereas CV and CVC ones as monomoraic.

Woleaian, a language closely related to Trukese, also exemplifies Pattern VII. Kennedy (2003) observes that there is a morphological process that marks the denotative form of many verbs by geminating the initial consonant resulting in a word-initial geminate. Kennedy specifically argues that this geminate is moraic because of a high-ranked constraint in the language that aligns foot structure with morphological structure (see Kennedy 2003, Chapter 7 for details).

Tedumuni Okinawan also displays initial moraic geminates as evinced by pitch accent patterns and word minimality (Shinozaka and Fujimoto 2011), while it lacks clusters (Shigeko Shinozaka, p.c. to first author, 25/9/2015). In Tedumuni a pitch fall – marked here with ‘–‘ is normally located on the 2nd mora of the word; if the word consists of light syllables, that will be the second syllable, as in [a.mi] ‘rain’ or [ka.ta.na] ‘knife’, but if the first syllable is a heavy, then it appears at the end of that syllable, as in [suu.ru] ‘head’. In words with initial geminates, we get e.g. [ssu.ru] ‘medicine’, a clear indication that geminates must be moraic and render the syllables they are contained in heavy. As for word minimality, the facts are comparable to other cases already examined; CV words lengthen to CVː, but GV words do not.

A final language that we will briefly mention that has initial moraic geminates but otherwise lacks initial consonant clusters is Luganda (though the language has prenasalized consonants word-initially which Muller 2001 considers as single nonmoraic segments). Muller (2001) argues in detail that Luganda initial geminates pattern as moraic with respect to allomorphy selection and tone spreading. Concerning allomorphy selection, she observes that the present perfect suffix of verb stems that have /l/ as a final consonant is [-ze] if the stem is two moras, but it is [-dde] if the stem contains more than two moras. Verb stems with two vowels (i.e. two vocalic moras) and an initial geminate take the suffix [-dde], thus patterning as trimoraic. With respect to tone spreading, Luganda has a phrasal tone spreading process whereby an intonational high tone appears on all but the first mora of the final word of a phonological phrase if that word does not have a lexical high tone. Normally, the initial vowel of a word does not surface with the intonational high tone; however, if the word begins with a geminate consonant, the initial
vowel does surface with a high tone. As Muller (2001) argues, this strongly suggests that the initial geminate is adding a mora to the word.

3.1.2 Pattern VIII: Languages with moraic initial geminates and moraic initial clusters

Cypriot Greek and Ponapean are used here to illustrate Pattern VIII, where initial geminates and clusters arise and where both are moraic.

Cypriot is a language with initial (as well as medial) geminates and a variety of initial clusters. Geminates are aspirated (in the case of stops) and longer than singletons, e.g. [pěfti] ‘Thursday’ vs. [ph:ěfti] ‘s/he falls’ or [kafė] ‘brown’ vs. [kh:afė] ‘café’. The analysis of geminates is still a matter of debate; in much research (Malikouti-Drachman 1987, 2003; Muller 2001, 2002; Arvaniti 2010), geminates are represented as single root nodes associated to X-slots or C-slots and are considered weightless. On the other hand, Christodoulou (2007) assumes a moraic representation, while Coutsougera (2003) – unlike previous work – proposes that Cypriot geminates are tautosyllabic. These two ideas are combined in Armosti (2011), who conducted 5 experiments on Cypriot (one acoustic, one articulatory and three perceptual) and ends up proposing that Cypriot geminates are better understood as moraic onsets (cf. Topintzi 2010). Moreover, he shows that geminates pattern like (most) clusters with respect to certain processes.

Armosti’s arguments in favour of the moraicity of geminates are both phonetic and phonological. Phonetically speaking, Armosti (2011: §2.3.1.1.3) finds that Cypriot geminates both word-initially and word-medially behave exactly in the way predicted by Ham (2001), who claims that geminates should exhibit more durational stability across different places of articulation (POA) than singletons. Ham’s explanation of that makes crucial reference to the mora as an inherent property of a geminate. This mora is realized by means of some minimum duration in geminates and overrides any durational effects due to POA. In singletons however, there is no such mora to start with, hence durational effects caused by POA are free to surface. As Armosti (2011: 270) puts it: “In that respect, durational stability across POAs can be used as a diagnostic for geminate moraicity”.4 Following this line of

---

4 Besides POA effects, Armosti finds a number of other measurements of absolute and relative timing that serve as acoustic cues for the geminate-singleton contrast, including longer duration of closure and longer aspiration for stops in geminates (absolute timing), as well as V2:HV2 ratio (relative timing). The latter refers to the ratio of the following...
thought and given that Cypriot geminates fulfil the phonetic criterion of moraicity, they consequently render heavy the syllable that hosts them.

With regard to phonology now, Armosti (2011: 271) acknowledges that the major objection expressed in the literature against a moraic account of Cypriot geminates is that they do not actively participate in phenomena that commonly employ syllable weight, e.g. word minimality or stress. But, as he correctly points out, this does not necessarily imply that the language does not distinguish between heavy and light syllables; such distinction might be identified on a phonetic level, as argued for extensively in his thesis.

Nonetheless, Armosti argues that two processes, that of /n/-deletion and of /i/-epenthesis, can be viewed as sensitive to syllable weight. More concretely, in the article *tin* ‘the- sg. acc. fem’ and the particle *en* ‘not’, the final nasal may delete or an epenthetic vowel [i] may be inserted right after it, depending on the environment. Both processes are triggered by following geminates and clusters, but not singletons.\(^5\) In this context, singletons cause nasal place assimilation instead. This could be reformulated in terms of syllable weight (Armosti 2011: 278); the nasal deletes if followed by a heavy syllable, i.e. one with a moraic onset coming from a geminate or an onset cluster. In the same environment, but under different morphophonological conditions, /i/-insertion is preferred instead.

(16) Cypriot Greek: *Geminates and clusters vs. singletons in phonological processes* (Armosti 2011: 273)

<table>
<thead>
<tr>
<th>onset</th>
<th>UR</th>
<th>/n/-assimilation</th>
<th>/n/-deletion</th>
<th>/i/-epenthesis</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>singletons</td>
<td>/ɛn ˈpeːzːɔ/</td>
<td>['ɛmbɛzːɔ]</td>
<td>N/A</td>
<td>N/A</td>
<td>‘I don’t play’</td>
</tr>
<tr>
<td>clusters</td>
<td>/ɛn ˈpsinːɔ/</td>
<td>N/A</td>
<td>['ɛniˈpsinːɔ]</td>
<td>‘I don’t bake’</td>
<td></td>
</tr>
<tr>
<td>geminates</td>
<td>/ɛn ˈpʰːɛftɔ/</td>
<td>N/A</td>
<td>['ɛniˈpʰːɛftɔ]</td>
<td>‘I don’t fall’</td>
<td></td>
</tr>
</tbody>
</table>

While it is true that these processes group together geminates and clusters to the exclusion of singletons, they do not unequivocally speak for weight, thus one may wish to challenge our decision to classify Cypriot as a Type VIII language. A cursory comment made elsewhere in Armosti (2011: 283) leads us however to suspect that our Pattern VIII classification of Cypriot is on the right track. He mentions that /s/+stop clusters can be replaced by geminate aspirated stops in the case of nickname formation. Thus the feminine names [ðɛspɪˈnu] and [xristːˈlːu] become [pʰ:iˈnu] and [tʰːlːu], respectively. That the cluster is replaced by a moraic geminate and not a singleton is potentially indicative of the fact that it is moraic itself.

If one nonetheless still treats the moraicity of clusters in Cypriot with scepticism, then the question remains: should we classify Cypriot as Pattern VIII? In fact, some readers might raise comparable concerns for a few of the other languages already examined, e.g. Hungarian or Tashlihyt Berber. Typically in those cases, we find clear evidence that geminates and clusters pattern together segmentally, but have little or unclear evidence regarding their weight behaviour. So the question is: are we justified in grouping those languages the way we do?

We think, provisionally yes; while the suggested classifications are not unambiguous, they are (i) fully compatible with the other facts of the language and (ii) the most parsimonious classifications in the current typology. Taking Cypriot as an example, we classified both initial geminates and clusters as moraic (Pattern VIII) – although moraicity evidence exists only for the former – on the grounds that they exhibit common segmental behaviour. Such an assumption does not create conflicts with other facts of the language, while it is the most coherent of all. What are the alternatives? If it is undisputable that initial geminates are moraic, then we are left with Patterns VII and IX. Obviously Cypriot could not be type VII (moraic geminates and no clusters), since it evidently possesses clusters; IX (moraic geminates and non-moraic clusters) would be conceivable, but it would generate the question: “if geminates and clusters are prosodically different, why are they segmentally comparable?”, which in turn would create an unnecessary internal inconsistency. Given the lack of positive evidence in support of such an inconsistency, we

---

6 We thank an anonymous reviewer and Donca Steriade (p.c. to first author) for raising this issue and for inspiring the ensuing discussion.
opt for the more parsimonious classification, here Pattern VIII. Comparable reasoning applies to Hungarian or Berber.

Another language that exemplifies Pattern VIII is Ponapean, which recall, also possesses final geminates (Goodman 1995, Kennedy 2003). Word-initially we find nasal geminates (17a) as well as nasal-stop clusters where the nasal is syllabic (17b).

(17) Ponapean: Initial geminates and NC clusters

a. [m̥met] ‘full’
   [ŋ̥net] ‘to pant’

b. [mp̥ek] ‘to look for lice’

Kennedy (2003) justifies the moraicity of these initial sequences through reduplication. The durative reduplicant appears as either monomoraic or bimoraic. In monosyllabic stems, the reduplicant appears bimoraic if the stem is monomoraic, but it shows up monomoraic if the stem is bimoraic. This is what Kennedy (2003: 78 and references therein) dubs Quantitative Complementarity. In the examples below, the reduplicant is underlined.

(18) Ponapean: Durative for monosyllables

<table>
<thead>
<tr>
<th>Stem</th>
<th>Reduplicated Form</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>i.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pa</td>
<td>[paapa]</td>
<td>‘weave’</td>
</tr>
<tr>
<td>dod</td>
<td>[dondod]</td>
<td>‘frequent’</td>
</tr>
<tr>
<td>tep</td>
<td>[tepitep]</td>
<td>‘begin’</td>
</tr>
<tr>
<td>ii.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>duup</td>
<td>[duduup]</td>
<td>‘divide’</td>
</tr>
<tr>
<td>miik</td>
<td>[mimiik]</td>
<td>‘suck’</td>
</tr>
<tr>
<td>pei</td>
<td>[pepei]</td>
<td>‘fight’</td>
</tr>
</tbody>
</table>

Bimoraicity in the reduplicant (18i) is ensured through a single heavy syllable of the type [CVV] or [CVC] or two light [CV.CV]’s. Note that independent evidence is available that final singleton codas are not moraic, whereas medial ones are. In polysyllabic stems, facts are quite complex, but for our purposes it is sufficient to observe that LL polysyllabic stems of the type [CV.CV(C)] receive a bimoraic reduplicant as shown in (19).
(19) Ponapean: Bimoraic CVC or CV.CV. reduplicant for LL bimoraic stems

<table>
<thead>
<tr>
<th>Stem</th>
<th>Reduplicated Form</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>dune</td>
<td>[dundune]</td>
<td>‘attach in a sequence’</td>
</tr>
<tr>
<td>dilip</td>
<td>[dindilip]</td>
<td>‘mend thatch’</td>
</tr>
<tr>
<td>siped</td>
<td>[sipisiped]</td>
<td>‘shake out’</td>
</tr>
</tbody>
</table>

In stems that begin with a geminate (20i) or a CC-cluster (20ii), the reduplicant is similarly bimoraic, e.g. [mmi] or [mpi], both with an epenthetic vowel /i/. This might seem surprising; a stem like [mmed] is presumably comparable to [miik] (18ii), that is, a bimoraic monosyllable. But had that been the case, then we would expect a monomoraic reduplicant, e.g. *[mmmed]. A solution would be to consider both nasals in geminates and NC clusters syllabic, hence [m̩med] and [m̩pek]. That would render these strings bisyllabic LL and as such, subject to the reduplication pattern of (19).

(20) Ponapean: Bimoraic GV or CCV reduplicant (Kennedy 2003: 98)

<table>
<thead>
<tr>
<th>Stem</th>
<th>Reduplicated Form</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>i.</td>
<td>mmed</td>
<td>[mmimmed]</td>
</tr>
<tr>
<td></td>
<td>ṭŋet</td>
<td>[ṃŋiŋŋet]</td>
</tr>
<tr>
<td>ii.</td>
<td>mpek</td>
<td>[ṃpimpek]</td>
</tr>
<tr>
<td></td>
<td>nda</td>
<td>[ndinda]</td>
</tr>
</tbody>
</table>

3.1.3 Pattern IX: Languages with moraic initial geminates and non-moraic initial clusters

Shuri Okinawan (Shimoji 2012), displays Pattern IX in the chart in (11) whereby initial geminates are moraic but other initial clusters are not. The language allows for initial geminates and a limited set of initial clusters all beginning with a glottal stop. It also

7 A reasonable alternative that, however, proves wrong is to hypothesize that geminates (and NC clusters) are non-moraic. That would then render the stem [mmed] truly monosyllabic and monomoraic along the lines of (18i) and consequently require a bimoraic reduplicant. But such demand can no longer be satisfied by the reduplicant [mmi], since the geminate is weightless. We would thus expect a reduplicant like *[mmii] or *[mmim], none of which arises.
imposes a bimoraic word minimum that is satisfied by CCV words that begin with a
geminate, but not with clusters.

(21) Shuri Okinawan: Word Minimum

<table>
<thead>
<tr>
<th>GV words:</th>
<th>[ccu] ‘person’</th>
<th>[kk”a] ‘child’</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCV words:</td>
<td>*[ʔwa]</td>
<td>but [ʔwaa] ‘pig’</td>
</tr>
</tbody>
</table>

3.2 Combinations of non-moraic initial geminates and clusters

Comparable to the situation with final geminates (§2.2), initial moraic geminates too are
robustly attested regardless of the occurrence and patterning of initial clusters. Likewise,
our preliminary survey finds fewer patterns involving non-moraic initial geminates. More
specifically, we detected neither Pattern X nor XI of the chart in (11), that is, neither the
combination of non-moraic initial geminates and no initial clusters nor the combination
of non-moraic initial geminates and moraic initial clusters. As in the case of final
geminates, future research should be able to establish whether these gaps are systematic
or accidental.

Contrary to Patterns X and XI, the remaining Pattern XII that refers to languages with
non-moraic initial geminates and non-moraic initial clusters, seems well-attested.
Languages such as Leti (Hume et al. 1997) and Swiss German (Kraehenmann 2001; Ringen
and Vago 2011) fit the profile for this pattern. While both support a bimoraic minimal
word constraint, no CCV words are allowed in either language regardless of whether the
CC is a cluster or a geminate.

An additional possible candidate is Tashlhiyt Berber (see Section 2.2 for more details
regarding the behaviour of both final and initial geminates), while even more probable is
the case for the Arabic dialects that allow for initial consonant clusters with very few
restrictions, including initial geminates. This includes Baghdadi Arabic (Blanc 1964,
Youssef 2013) and Hadhrami Arabic (Bamakhramah 2009), but the possible weight
properties of these initial sequences have not been systematically discussed in the
literature on Arabic phonology. Nonetheless, if we consider Baghdadi Arabic, we observe
that a word-initial syllable that begins with a geminate does not attract stress to that
syllable. This is significant given that the stress rule is quantity-sensitive and, as we
indicated in Section 2.1.1, a word-final geminate always attracts stress to the final
syllable. Also, the word-initial geminate in Baghdadi Arabic seems to pattern as non-
moraic with respect to the bimoraic minimal word condition, since, as far as we are aware, the language lacks words consisting solely of an initial geminate followed by a short vowel. The patterning of initial geminates in Arabic dialects that have them is a ripe area for future research.

4 Discussion: Asymmetries and gaps in the typology of edge geminates

The tables in (22) repeat our findings regarding the emergence of final and initial geminates, as well as their co-existence, if applicable, with clusters. The names of the languages illustrating each pattern are also included.

(22) (i) Instantiation chart: final geminates (cf. (2))

<table>
<thead>
<tr>
<th>Final Geminates</th>
<th>Final Clusters</th>
<th>Moraic</th>
<th>Non-moraic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moraic</td>
<td>(I) Baghdadi Arabic, Hadhrami Arabic</td>
<td>(II) Swiss German, Cairene Arabic, Ponapean, Wolof</td>
<td>(III) Amharic, San'ani Arabic</td>
</tr>
<tr>
<td>Non-moraic</td>
<td>(IV) ---</td>
<td>(V) ---</td>
<td>(VI) Hungarian, Tashlhiyt Berber</td>
</tr>
</tbody>
</table>

(ii) Instantiation chart: initial geminates (cf. (11))

<table>
<thead>
<tr>
<th>Initial Geminates</th>
<th>Initial Clusters</th>
<th>Moraic</th>
<th>Non-moraic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moraic</td>
<td>(VII) Pat.Malay, Trukese, Woleaian, Tedumuni Okinawan, Luganda</td>
<td>(VIII) Cypriot Greek, Ponapean</td>
<td>(IX) Shuri Okinawan</td>
</tr>
<tr>
<td>Non-moraic</td>
<td>(X) ---</td>
<td>(XI) ---</td>
<td>(XII) Leti, Swiss German, Baghdad Arabic, Tashlhiyt Berber</td>
</tr>
</tbody>
</table>

The state of affairs is better summarized in (23), where the focus is placed on the typological gaps found. Moraicity or lack thereof is signalled through $\mu$ and $-\mu$, respectively.
The first thing one may notice is that the picture at both edges of the word is highly comparable. In fact, the patterns arising are identical: non-moraic EGs only occur in languages in which clusters at the same edge exist and moreover pattern as non-moraic too. Consequently, the distribution of non-moraic geminates is overall much more restricted than the corresponding distribution of moraic geminates.

Given the chart in (23) we can identify two interesting generalizations. In particular, as the darkly shaded cells reveal, there seem to be no languages possessing a non-moraic EG while lacking a consonant cluster on that edge. Second, there seem to be no languages possessing a non-moraic EG while having a moraic cluster on that edge (light shaded cells). At this point, we cannot tell if these gaps in the typology are accidental or not.

While we believe that, at least, most of them are not accidental – for reasons to be mentioned next – this is an issue that has to be resolved through future documentation of additional languages with edge geminates. For the time being, we tentatively state the generalizations above by means of implicational universals.

(24) Implicational Universal 1

If a language has an edge geminate but no consonant clusters on that edge, then the edge geminate patterns as moraic (or “If edge clusters do not exist, then EGs are moraic”)

(25) Implicational Universal 2

If a language has an edge geminate that patterns as nonmoraic and allows for consonant clusters on that edge, then that cluster must pattern as nonmoraic too (or “If EGs are non-moraic, then edge clusters too are non-moraic”)

(23) Summary of EGs and edge clusters

<table>
<thead>
<tr>
<th>Final Geminates</th>
<th>Final Clusters</th>
</tr>
</thead>
<tbody>
<tr>
<td>G_µ #</td>
<td>(I) YES</td>
</tr>
<tr>
<td>G_¬µ #</td>
<td>(IV) NO</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Initial Geminates</th>
<th>Initial Clusters</th>
</tr>
</thead>
<tbody>
<tr>
<td># G_µ</td>
<td>(VII) YES</td>
</tr>
<tr>
<td># G_¬µ</td>
<td>(X) NO</td>
</tr>
</tbody>
</table>
If indeed the observed generalizations reflect true universals, a reasonable question to ask is whether these can be somehow explained. A partial answer may possibly lie in the nature of geminates. The following discussion is largely based on medial geminates and singleton codas; since these are attested in several languages, it is both the case that they have been extensively explored, but also that a comparison of their properties has been conducted.

According to many theorists (Hayes 1989; Ham 2001; Davis 2003, 2011; Topintzi 2008, among others), geminates are inherently moraic, i.e. weightful. As discussed in Section 1, we find languages where both geminate and singleton codas are weightful (Latin, Lake Miwok), or where both are weightless (Selkup, Malayalam and Tübatulabal in Tranel 1991). While the latter set of languages may at first glance contest the underlying weight of geminates, Davis (2003, 2011: 890) has argued that these data too can be re-analyzed in a manner compatible to the inherent moraicity of geminates.

Moreover, Davis has offered a strong argument against Tranel’s Principle of Equal Weight for Codas (Tranel 1991 and Section 1 here) and in favor of the moraic analysis of geminates. In several papers (Davis 2011 and references therein), he has shown that languages with moraic geminates and non-moraic singletons exist, that is, cases where the weight of codas is non-uniform. These include West Swedish, Koya, Seto, Fula, and Cahuilla. The opposite pattern, where geminates are non-moraic, but singleton codas are moraic is clearly unattested, with one possible exception, that of Ngalakgan ($1, §2.2.2). In Ngalakgan however, the place of articulation of consonants is instrumental (geminates and codas in homorganic clusters are non-moraic; codas in heterorganic clusters are moraic) and could perhaps help us understand the language’s unlikely pattern. At present, we consider Ngalakgan a deviation that does not disturb the general pattern.

This state of affairs finds analogues in the case of EGs and clusters. Patterns II and VIII, with uniform moraicity, are analogous to Latin; patterns VI and XII, with uniform non-moraicity, are analogous to Malayalam. Interestingly, patterns III (Amharic) and IX (Shuri Okinawan) are comparable to data from languages such as Cahuilla, since moraicity is non-uniform. Geminates are moraic, but clusters are not.

Two attested patterns remain: I and VII, i.e. languages with moraic EGs but no clusters whatsoever. These can be best accounted for by making reference to edges and assuming the representational account of Topintzi (2008, 2010). In that work, initial and final moraic geminates look like singleton onsets and codas, respectively, but differ in
bearing a mora that the real singletons lack. Since moraic edge geminates are largely representationally identical to singletons, it follows that they can co-exist in the absence of clusters. The latter would require a different representation, e.g. double linking to an onset or coda.

(26) i. Word-final distinction between singletons and geminates

a. singleton CVC]  
\[\sigma\] \\
\[\mu\] \\
\[\nu \ c]\n
b. geminate CVC:\n\[\sigma\] \\
\[\mu\] \\
\[\mu\] \\
\[\nu \ c:\]\n
ii. Word-initial distinction between singletons and geminates

a. singleton CV  
\[\sigma\] \\
\[\mu\] \\
\[\nu \ c\]\n
b. geminate C:\nu
\[\sigma\] \\
\[\mu\] \\
\[\mu\] \\
\[\nu \ c\]\n
This leaves us with the alleged gaps of languages with weightless geminates and no clusters (IV, X), as well as languages with weightless geminates and weightful clusters (V, XI). If one adopts a moraic analysis for all geminates, then weightless geminates can be derived under the assumption that “geminates may be underlyingly moraic, [but] do not [have to] surface as moraic” (Davis 2011: 892). In practice this means that weightless geminates could start life as moraic but end up weightless on the surface due to, for example, a constraint that bans consonant moraicity. If that is the case though, all consonants, including those in clusters, should be equally affected, effectively enforcing weightlessness for all consonants in geminates or clusters. In turn, this would suggest a neutralization of contrast between singletons and geminates and would consistently produce the structures in (26.i.a) and (26.ii.a). As a result, we would generate a language with weightless singletons (or geminates neutralizing to singletons) and weightless clusters. Such languages obviously occur, but are irrelevant to the typology of geminates.

What this suggests is that we may wish to maintain the structures in (26) for moraic edge geminates, but suggest that weightless geminates are to be represented differently. For example, they might involve double-linking to two root nodes (Selkirk 1990) or to two
timing slots (Tranel 1991) under the onset or coda nodes. Empirical evidence from some of the languages that display weightless geminates is in fact supportive of this position. For example, in Baghdadi Arabic the *weightless* initial geminates are all heteromorphemic, brought about through morpheme concatenation or vowel deletion over a morpheme boundary. This contrasts with final moraic geminates in Arabic which are almost always tautomorphemic. In Leti, no phonological restrictions are placed on what the two consonants can be at the beginning of a word. Any cluster or – what looks as – a geminate is possible. Davis (1999: 97-98) analyses this fact by means of an optional adjoined C-slot at the beginning of a word before the single onset that can be filled by anything (or not at all if the word begins with a single consonant). It remains to be seen whether all cases of weightless initial EGs are similar to either the Baghdadi Arabic or the Leti case and whether a comparable analysis is available for final non-moraic geminates, whose behaviour as explained in §2.2 is much less understood.

Assuming such a representation of weightless geminates, and in particular one that brings them closer to clusters, we can better comprehend the gaps we find in our typology. If a language has weightless geminates it should necessarily also exhibit (weightless) clusters (cf. attested patterns VI and XII), leaving no space to describe a language with weightless geminates and no clusters (Patterns IV and X) since these two present comparable structure, nor a language with moraic clusters and weightless geminates (Patterns V and XI), since as before, weight assignment, should equally apply to all consonants.

This solution is also in line with a potential re-statement of the proposed implicational universals, as suggested to us by an anonymous reviewer. The two universals are merged here into a single universal, stating that *If a language has nonmoraic edge geminates then it must have consonant clusters at that edge and the clusters must be nonmoraic.* Its advantage is that it captures Patterns VI and XII, as well as all four gaps in one go. At the same time, it makes no prediction for the correlation between moraic geminates and clusters. This might be a good thing, as it parallels the behaviour of medial geminates and singletons: recall that moraic medial geminates may combine with both moraic singleton codas (Latin) or non-moraic ones (Cahuilla), unlike non-moraic medial geminates, which, with the exception of Ngalakgan, only co-exist with non-moraic coda singletons (Malayalam). The finding that a language must have clusters if weightless geminates are present may in turn suggest that weightless geminates are essentially
clusters themselves, and consequently, as indicated above, receive a different representation from the ‘real’ moraic geminates.

Although we acknowledge the appeal of this re-statement of facts, we presently refrain from fully endorsing it for two main reasons. First, its adoption would require us to expressly accept the split representation of geminates. While we think that the present work may in fact prove significant in resolving the debate on geminate representation demonstrating that in fact both the weight-based (Hayes 1989, Davis 2011, Topintzi 2008) and length-based (Ringen and Vago 2011) accounts have been right all along, but for different types of geminates, it is evident that our typological survey is still rather limited. It is thus preliminary to commit ourselves on this matter. Second, the alleged parallel between EGs and edge clusters and medial geminates and medial clusters is not exactly a parallel, simply because at edges, clusters – and possibly geminates – are arguably tautosyllabic, whereas medially they are necessarily heterosyllabic (so that C₁ is found in a coda singleton), indicating that the position within the word may be a more important parameter to consider.

We close this section by noting that besides the asymmetries arising in the typology of EGs across languages, we also find asymmetries within the same language. In particular, some languages present both initial and final geminates, but their behaviour is not necessarily uniform. In Swiss German and Baghdadi Arabic, final geminates are moraic while initial geminates are not. On the other hand, EGs of Ponapean (and probably Puluwat; Elbert 1974) seem to be consistently moraic. Given the rarity of languages that demonstrate geminates at both word edges, it is impossible to determine whether the lack of moraic initial geminates and non-moraic final geminates is random or systematic. For Dmitrieva though (2012: 217), this gap is not coincidental. She ascribes it to the fact that onset and coda moraicity cannot be simultaneously available to the language. While this largely seems to be empirically supported, EGs in Ponapean contest this view. Moreover, without any further elaboration, it still does not follow why a language should not combine initial weightful geminates with final weightless ones.

---

8 Note that Dmitrieva (2012: 170) accepts initial geminates in Ponapean, but appears unsure about final ones. Another possible counter-example is Karo (Topintzi 2010), a language that lacks geminates. In that analysis, Karo voiceless and sonorant onsets bear weight and so do sonorant codas.
5 Conclusion

In this paper we have explored the typology of edge geminates. Although much rarer than medial geminates, many languages nevertheless employ them. In our survey, we have considered languages with weightful or weightless geminates, weightful or weightless clusters, as well as languages that lack clusters altogether. We have then investigated the co-occurrence possibilities between these categories and we have found that most combinations are allowed, however certain gaps in the typology emerge. While a larger database of languages would need to be consulted before we reach any safe conclusion, certain implicational universals can be enunciated at this point.

According to the first implicational universal (24), if a language has an edge geminate, but no consonant clusters on that edge, then the EG patterns as moraic. The second universal (25) states that if a language has an EG that patterns as nonmoraic and allows for consonant clusters on that edge, then that cluster must pattern as nonmoraic too. We have also observed that a few languages demonstrate geminates at both edges of the word, but their behaviour does not have to be uniform in terms of weight, i.e. weightful geminates at one edge may co-occur with weightful geminates at the other edge, but do not have to.

As mentioned, certain asymmetries that arise in the typology of geminates can be understood through already available theoretical machinery. A fuller survey however, besides examining more languages (ideally newly documented ones), would also need to consider additional parameters. For example, does it matter whether the EG is derived or underlingly present? How do partial geminates (i.e. nasal+homorganic clusters) behave? Do they pattern like EGs or can they exhibit distinct behaviour? Future research should seek to answer these questions too and examine the implications, if any, for the typology of edge geminates.

The somewhat restricted database is partly due to the inherently limited number of languages that possess edge geminates. In some cases the situation is additionally hindered by lack of or insufficient information relevant to EGs and/or clusters. Some languages are reported to also possess EGs, but have not been included in this survey (cf. §1), e.g. Estonian final geminates (Dmitrieva 2012) or Circassian initial geminates (Muller 2001). We believe this omission does not disturb the main findings outlined earlier. Inclusion of those languages however, as well as documentation of new languages with EGs would likely benefit and enrich the present typology.
REFERENCES


Davis, Stuart. this volume. ‘Geminates and weight manipulating phonology in Chuukese (Trukese)’. 

29


Ladefoged, Peter and Ian Maddieson. 1996. The sounds of the world’s languages. Oxford: Blackwell


Sande, Hannah and Andrew Hedding. 2014. ‘Geminates and weight in Amharic’, LSA, Minneapolis, Jan 2-5.
