A new phonological analysis of geminates in Cypriot Greek (and beyond)

Nina Topintzi  
Aristotle University of Thessaloniki, Thessaloniki, Greece  
topintzi@enl.auth.gr

Abstract

Cypriot Greek geminates have been the subject of much past research and controversy. One major challenge has been their proper syllabification. Another involves their weight status. Taking stock of recent phonetic evidence that supports an inherent mora for Cypriot Greek geminates and tautosyllabicity on the basis of durational effects on the vowel that follows the geminate (Armosti 2011), I build a formal phonological analysis of Cypriot Greek lexical geminates as moraic onsets. The analysis employs WCL (‘weight-and-concomitant-length’), a new theory of gemination designed to handle both the weight and length properties of lexical geminates in comparison to singletons, as attested cross-linguistically. WCL manages to produce both heavy and light geminates on the surface, and dispenses with double linking as a means to express a geminate’s increased length. These features of WCL prove crucial in the analysis of Cypriot Greek. Finally, WCL’s theoretical and typological superiority is furnished through its extensions to other patterns.

Keywords

lexical geminates – Cypriot Greek – onsets – weight – clusters – typology
1 Introduction

A famous trait of Cypriot Greek (CyGr) that brings it on a par with some other Greek dialects in the Dodecanese, Chios and West Cyclades (Κοντοσόπουλος 2000: 21, 43) is that it possesses geminate consonants. These appear medially as well as initially; also, in the case of plosives and affricates, geminates appear aspirated, as shown in (1).\(^1\)

(1) CyGr: singleton vs. geminate plosives (Armosti 2011: 16)

<table>
<thead>
<tr>
<th>Singleton</th>
<th>Geminates</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ˈpefti ‘Thursday’</td>
<td>‘pʰ:efti ‘s/he falls’</td>
</tr>
<tr>
<td>b. kaˈfe ‘brown’</td>
<td>kʰ:aˈfe ‘café’</td>
</tr>
<tr>
<td>c. xtiˈpun ‘they hit’</td>
<td>paˈpʰ:us ‘grandfather’</td>
</tr>
<tr>
<td>d. ˈkokos ‘Kokos (a name)’</td>
<td>‘kokʰ:os ‘grain’</td>
</tr>
</tbody>
</table>

The phonetics and phonology of CyGr geminates has received substantial attention in the literature (Malikouti-Drachman 1987, 2003; Arvaniti 2001, 2010; Tserdanelis & Arvaniti 2001; Muller 2001, 2002; Coutsougera 2003; Christodoulou 2007; Armosti 2011; Topintzi & Davis 2017; among others). Despite individual particularities and differences, some consensus has been reached in terms of phonetics. Thus, theorists agree that geminates are longer than singletons (by about 1.5–2 times) and the placement of stress as well as the geminate’s own position within the word account for the durational differences found across studies.

Turning to phonology now, things are less clear; while most researchers nowadays accept that CyGr geminates are true geminates, in the sense that they are best represented as single root nodes associated to two timing slots—instead of clusters of identical consonants (see e.g. Arvaniti 2010 for relevant discussion)—the timing units involved and the repercussions for syllable structure and weight are still debated. The majority of work (e.g. Malikouti-Drachman 1987, 2003; Muller 2001, 2002; Arvaniti 2010) takes CyGr geminates to be light, i.e. non-moraic, with just a few assigning weight to them (Christodoulou 2007; Armosti 2011). Similarly, for most, geminates straddle syllable

---

\(^{1}\) The focus of this study is the underlying/lexical geminates of CyGr. The language also possesses derived postlexical geminates, which are not presently handled (see §3.1, 3.3 for discussion).
boundaries, with few exceptions that consider them tautosyllabic (Coutsougera 2003; Armosti 2011).

To understand these discrepancies, one however needs to take into consideration the analysis of geminates in a wider, cross-linguistic context. Two are the most influential models of geminate representation. In the former (Hayes 1989; Davis 1994, 1999, 2003; Topintzi 2008, 2010, among others), dubbed the ‘weight’ approach, a geminate consonant is *underlyingly* heavy, thus bears a mora $\mu$. On the surface, the (intervocalic) geminate appears doubly linked so that its first half links to the (moraic) coda of one syllable and its second half directly as an onset of the second syllable (2a). In the latter, which I will term the ‘length’ approach (e.g. Tranel 1991; Ringen & Vago 2011), the geminate is both underlingly and on the surface a doubly linked consonant, only this time an additional timing tier is needed. Weight is not inherently present here, but can be added on codas, through the process of Weight-by-Position (2b). There are merits to both accounts and also important differences in their predictions. The interested reader is pointed to Davis (2011) for a summary. Here, I only focus on specific aspects of each proposal relevant to the discussion.

(2) The ‘weight’ (a) vs. ‘length’ (b) approaches

<table>
<thead>
<tr>
<th></th>
<th>Underlying</th>
<th>Intervocalic</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>$\mu$</td>
<td>$\sigma$ $\mu$ $\mu$ $\sigma$ $\mu$</td>
</tr>
<tr>
<td>b.</td>
<td>$\alpha$ $\mu$ $\mu$ $\alpha$ $\sigma$ $\mu$ $\mu$ $\sigma$ $\sigma$</td>
<td></td>
</tr>
</tbody>
</table>

Under the assumption that CyGr geminates are light (but see Section 3.1), these are seemingly better catered for by a ‘length’ approach, since this view is inherently designed to capture the increased length of geminates as opposed to singleton counterparts, even in the absence of weight. Indeed, this idea underlies the majority of analyses of CyGr geminates. Obviously, light geminates pose difficulties for the ‘weight’ approach, where weight is part and parcel of being a geminate.

On the flip side, though, and looking beyond CyGr, the ‘weight’ account has proven more influential, for a number of reasons, both empirical and theo-
On the empirical side, geminates in most languages are indeed heavy, regardless of whether singleton codas likewise are (as in e.g. Latin) or are not (as in e.g. Koya). This latter case in fact is impossible to produce in the ‘length’ approach, where syllable weight is a derived property. If Weight-by-Position needs to apply to produce heavy coda geminates, then it will necessarily render singleton codas heavy too. On the theoretical side, the ‘weight’ account is more economical in the sense that it merely employs a structure—namely, moras—which is anyway required, as opposed to the ‘length’ view which requires an extra timing tier, on top of the mora tier. It is thus no surprise that even for CyGr, a language where syllable weight may not be at stake, moraic approaches to geminates have been proposed, as mentioned above.

The preponderance of the ‘weight’ approach is remarked upon even by proponents of the lightness of CyGr geminates, such as Arvaniti (2010). She further comments that a possible solution to the light geminates problem the ‘weight’ approach faces is the adoption of the shared-mora view of Broselow et al. (1997), in which the mora stemming from the geminate consonant is shared on the surface between the geminate and the preceding vowel. Effectively, then, mora sharing is used to produce a light geminate. Crucially however, mora sharing also entails phonetic shortening of the preceding vowel, as happens in Hindi or Levantine Arabic. Arvaniti points out, on the basis of much previous work, that there is nonetheless no evidence that such shortening occurs in CyGr (see also Section 3.2) and concludes that “the overall phonetic behaviour of CyGr geminates could reflect the fact that they may not be fully-fledged geminates, or may be atypical geminates in some ways” (2010: 121).

To make matters worse, recall that CyGr geminates also appear word-initially. Albeit somewhat rare, other languages too exemplify a similar pattern. Trukese is a well-known example (Davis 2017 and references therein), and in fact quite analogous to CyGr, since it admits geminates in exactly the same positions: initially and medially. Pattani Malay is even more startling, because geminates exclusively appear in word-initial position and nowhere else (Yupho 1989; Hajek & Goedemans 2003). A common denominator is that initial geminates cannot be accounted for by either model of (2). Regardless of the specific details, geminates are doubly linked to higher structure, which intervocally means straddling of syllable boundaries. But, carrying over this representation initially is no trivial task.

Have we then reached a stalemate, as Arvaniti seems to suggest, and give up on CyGr geminates altogether due to their presumed atypical nature? The response I entertain here is negative; CyGr geminates are no less typical than other ‘typical’ geminates. I also offer an analysis that captures the CyGr facts comprehensively, but is also theoretically sound and typologically plausible. I
show that this requires a hybrid model that considers all underlying geminates to be moraic (following the ‘weight’ account) and longer on the surface than their non-geminate counterparts (following the ‘length’ account).

Before proceeding, I wish to clarify that the nature and scope of this article lie on the formal-analytical side, building on the empirical findings of previous research. This is a deliberate decision; in the last 20 years or so, we have learnt much about the phonetics (mostly) and the phonology of CyGr geminates on a primarily descriptive level, but these findings have not been more systematically modelled in a formal fashion, especially one that takes into consideration the main typological patterns of gemination, as occurring across languages. This is a task I undertake here. However, for reasons explained in Section 3.3, this first step tackles lexical geminates and—to some extent—their relationship to clusters. A fuller account that also considers post-lexical geminates is left for future work.

The remainder of the paper is structured as follows. Section 2 outlines the proposed theory, which is in line with Topintzi & Zimmermann (In prep.). Section 3 summarizes the data under consideration, supplies the main findings of Armosti (2011) which serve as the empirical basis for the analysis, and justifies the focus on lexical geminates. Section 4 presents an optimality theoretic account of CyGr lexical geminates adopting the WCL model of gemination from Section 2, whereas Section 5 addresses emergent issues and outlines the model’s typological predictions. Section 6 provides concluding remarks.

2 A new model of geminates

The proposal I outline here (for details and extensions, see Topintzi & Zimmermann In prep.), couched within Optimality Theory (Prince & Smolensky 1993/2004), endorses the standard ‘weight’ theory of geminates, which is to say that all a lexical geminate (G) is taken to be underlyingly a moraic consonant. However, it comes with a twist; namely, that the underlying mora may not eventually surface as such. This latter suggestion was hinted in Davis (2011)—in view of light geminates—but without further concrete elaboration. This task is undertaken here. In anticipation of the ensuing discussion, I call this theory the ‘weight-and-concomitant-length’ (WCL) account of geminates.

An important conceptual attribute of WCL then is that it views the increased length found in lexical geminates as a concomitant feature of their underlying weight, and not as an isolated, independently arising property. Technically, this is achieved by assuming that the specifics of the mora’s prosodic integration give rise to different effects. Thus, the mora’s integration under a syllable
(a higher node) relates to phonological weight, whereas its association to a segment (a lower node) relates to phonetic length.

This means that for an underlying intervocalic geminate, i.e. /VμCμVμ/, there are typically three possible outputs, as demonstrated in (3) (see § 3.2 for further discussion). Also, following standard assumptions (but see Vaux 2003 or Elfner 2006 for counter-evidence and related discussion), syllabification is considered to be lexically non-contrastive, thus added on the surface. In that sense then, both the syllables and the association lines to other phonological elements (e.g. moras, segments) are in reality epenthetic. This is marked by grey shading around syllables and dotted association lines. This point becomes crucial later on, in the consideration of singleton vs. geminate contrasts.

(3) Possible outputs for an underlying intervocalic geminate: first take

<table>
<thead>
<tr>
<th>Underlying:</th>
<th>Output 1</th>
<th>Output 2</th>
<th>Output 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>VμCμVμ</td>
<td>[VμCμVμ]</td>
<td>[VμC(μ)Vμ]</td>
<td>[VμCμVμ]</td>
</tr>
<tr>
<td>abbreviation:</td>
<td><img src="image_url" alt="Diagram" /></td>
<td><img src="image_url" alt="Diagram" /></td>
<td><img src="image_url" alt="Diagram" /></td>
</tr>
</tbody>
</table>

(3a) depicts the typical intervocalic geminate, abbreviated as [VμCμVμ], which is both weightful and long, as in Latin or Koya. Heaviness comes about through the underlying mora’s epenthetic link to the higher syllable node—which is responsible for weight—whereas increased length is due to the presence of an underlying link between the mora and the segment it dominates. The fact that the geminate straddles syllable boundaries through its double linking is thus not due to length requirements, but a by-product of syllabification considerations, i.e. to render the coda moraic and provide an onset for the following syllable (cf. Ham 2001 for a similar proposal). (3b) contains the representation of a weightless geminate [VμC(μ)Vμ], as in Selkup and Ngalakgan (§ 5). As before, the geminate is interpreted as long, since it sustains its underlying mora-to-segment link. Due to the absence of an epenthetic link from the mora to higher prosodic structure, however, the geminate cannot be interpreted as heavy, leading to a weightless geminate. Finally, (3c) presents a short weightless consonant, where both the mora and its underlying link have been erased. In other words, we get a light singleton. At first, this might be a strange output to consider for an underlying geminate, but in reality, it is an important one under the tenet of Richness of the Base (Prince & Smolensky 1993/2004), according to which no restrictions should be placed on inputs. Thus, even in a language, e.g.
Standard Modern Greek, which lacks surface geminates altogether, we should be able to consider an input geminate and map it to some output. Presumably, in such a case, the underlying geminate neutralizes to a weightless singleton.

The patterns just illustrated can be captured by means of the permutation of the constraints in (4). The exemplification of each pattern follows.

(4)  \textbf{MAX-}\(\mu\) \quad \text{Do not delete an input mora in the output}

\(\mu>\sigma\) \quad \text{Every mora should be dominated by a syllable}

\textbf{*C}\(\mu\) \quad \text{A consonant should not be dominated by a prosodically integrated mora}

The order of candidates in the tableaux of this section is maintained identical, so that comparison across tableaux is facilitated. Candidate (a) is the long weightful G, candidate (b) is the long weightless G and (c) is the neutralized short singleton. Tableau (5) generates (a) as the winner. Both \(\mu>\sigma\) and \textbf{MAX-}\(\mu\) dominate \textbf{*C}\(\mu\). (5b) is eliminated due to a violation of high-ranked \(\mu>\sigma\), because its mora is not dominated by a syllable; similarly, (5c) fatally violates \textbf{MAX-}\(\mu\) as a result of the underlying mora’s deletion. (5a) wins, since it only incurs a violation of low-ranked \textbf{*C}\(\mu\). Note that candidate (b) escapes violation of the same constraint, because although it retains the underlying mora in the output, this is prosodically unintegrated, i.e. it has no linking to the higher syllable node.

(5)  \textbf{Weightful geminate:} \(\mu>\sigma\), \textbf{MAX-}\(\mu\) \(\gg\) \textbf{*C}\(\mu\)

\begin{center}
\begin{tabular}{c|c|c}
\(\mu\) & \(\mu\) & \(\mu\) \\
V & C & V \\

\(\mu>\sigma\) & \textbf{MAX-}\(\mu\) & \textbf{*C}\(\mu\) \\

\(\sigma\) & \(\sigma\) & \(\sigma\) \\
V & C & V \\

\(\sigma\) & \(\sigma\) & \(\sigma\) \\
V & C & V
\end{tabular}
\end{center}
Generation of the weightless G merely requires a swap in the positions of $\mu>\sigma$ and $^*C^\mu$, as shown in tableau (6). Finally, neutralization to a short singleton requires the ranking: $\mu>\sigma$, $^*C^\mu >> \text{MAX-}\mu$.

(6) **Weightless geminate:** $^*C^\mu$, $\text{MAX-}\mu >> \mu>\sigma$

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>$\mu$</td>
<td>$\mu$</td>
</tr>
<tr>
<td>V</td>
<td>C</td>
<td>V</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma$</td>
<td>$\sigma$</td>
<td>$\sigma$</td>
</tr>
<tr>
<td>V</td>
<td>C</td>
<td>V</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
</table>
| $^*C^\mu$ | MAX-| $\mu>\sigma$
|      | $\mu$ |      |

To sum up to this point, the claim has been that underlying geminates surface as heavy when their mora remains in the output and is prosodically integrated, and surface as light when the mora is retained but unintegrated. Notably, both types of geminates emerge as phonetically long and this is due to their linking to a mora via an *underlying* association line. Of course, this is not the only way a consonant may be rendered long; Topintzi & Zimmermann (In prep.) discuss other types of long consonants, such as the ‘fake’ geminates of English (cf. *unnamed*) or geminates derived through assimilation as in Hungarian, and provide a fuller set of principles regarding the phonetic interpretation of geminates. The interested reader is pointed to this work for further details.

In the traditional approaches of ‘length’ and ‘weight’, geminates found at word edges, namely initially or finally, are quite difficult to handle. This is either because double linking in these positions is not usually readily available (problem for the ‘length’ approach) or because the alternative single linking may lead to the inability to distinguish between e.g. final heavy singletons from final heavy geminates in languages like Hungarian (Curtis 2003) (problem for the ‘weight’ approach). In WCL though, as previously explained, double link-
ing is not a marker of increased duration, but rather a result of syllabification
requirements. Instead, it is the consonant’s linking to a mora via an underlying
association line that marks geminate length, which in turn means that espe-
cially at word edges, we may find consonants with just single linking and these
can still be interpreted as geminates by virtue of their association to an under-
lying mora. For initial geminates in particular, which is our focus in CyGr, this
means that these are to be analyzed as moraic onsets (cf. Topintzi 2008, 2010).
I explore this point in more detail in § 4.

There is one more important component to the core theory of WCL; this has
to do with the fact that geminates in a language co-exist with singletons, whose
weight behaviour is not necessarily uniform to that of Gs. Thus, we find lan-
guages where geminates and singletons are both uniformly heavy (e.g. Latin)
or light (Selkup), and others where weight is non-uniform. In Koya for exam-
ple, geminates are heavy and singletons are light. The opposite is also possible;
Ngalakgan is quite remarkable, because geminates there are light and single-
tons are heavy. I will sketch an analysis of this intriguing pattern in the current
model (in § 5), as it is not derivable in either theory of gemination mentioned
before. For now, it suffices to say that our theory, so far oriented towards Gs,
needs to account for singleton weight too in the same language and capture
the diversity attested.

To that end, WCL assumes that morphological affiliation is detectable in
the phonology. This idea is not novel; for instance, earlier work has pursued
the concept of morphological colours (cf. van Oostendorp 2006 or Revithi-
adou 2007), according to which all elements in an underlying morpheme are
coloured, i.e. have their own index. Consequently, epenthetic elements are
colourless. This is relevant to the empirical issue just raised, because it effec-
tively means that we can always tell apart underlying moras—the hallmark of
being a geminate—from epenthetic ones. The latter are inserted as a response
to a high-ranking Weight-by-Position constraint (Rosenthal & van der
Hulst 1999, Morén 2001), and render a singleton consonant heavy, which oth-
wise remains light, since by definition the singleton bears no mora underly-
ingly. This difference between the two types of moras is always visible in the
phonological structure and the interplay between them generates the differ-
ent combinations of singleton and geminate weight. In other words, geminate
weight does not rely on singleton weight and vice versa. More on that in § 5.
3 Cypriot Greek geminates: the data

Having set the necessary background for WCL, we can now turn to CyGr geminates. Section 3.1 offers an overview of the data. This is followed by a summary of Armosti’s (2011) main phonetic findings alongside their theoretical interpretation (§ 3.2), as these will form the basis of the analysis in Section 4. Finally, § 3.3 briefly addresses the intricacies that the different types of geminates (lexical, assimilated, concatenated) involve and justifies the analytical choice made here, which is to focus on lexical geminates only.

3.1 Data overview

As mentioned at the beginning of the paper, CyGr geminates appear either initially or medially. Most consonants of the language present contrastive singleton and geminate versions, i.e. \([p(ʰː), t(ʰː), k(ʰː), f(ː), v(ː), ðf(ʰː), ð(ː), s(ː), f(ː), x(ː), m(ː), n(ː), l(ː)]\), and in many cases, their allophones too, e.g. \([c–eʰː]\). Geminate plosives and affricates are accompanied by aspiration. In the case of the rhotic, the singleton vs. geminate contrast involves the tap /ɾ/ vs. trill /ɾː/, respectively. Some consonants are reported to only occur as geminates, namely \([ʦʰː]\) and \([zː]\); others, have singleton and geminate versions, but are not deemed contrastive, e.g. \([ð]\) and \([ðː]\); still, others only emerge as singletons, but these are always allophonic variants, e.g. the voiced \([b, d, ɡ, ɹ]\) variants of the voiceless plosives (Armosti 2011: 3). Elsewhere, a slightly different picture is presented. For instance, Coutsougera (2003) treats geminate fricatives as derived through assimilation, that is, non-lexical. For our purposes, it is sufficient to notice that the phonemic contrast between singletons and geminates is well-accepted and well-established for a large set of CyGr consonants. The fact that gemination is not contrastive for all is not surprising; it is in fact typical for many languages with geminates, e.g. Italian (Bertinetto & Loporcaro 2005; see also Morén 2001: 73 for a list of several languages). Some examples with lexical geminates appear in (7).

—

2 A reviewer reminds me that what Armosti calls ‘inherent’ geminates, i.e. \([ʦʰː]\) and \([zː]\), may appear word-final in loanwords, as in \([eɪʦʰː]\) ‘AIDS’ (Armosti 2011: 7). The status of \([ʦʰː]\) as a true lexical geminate, at least word-finally, is questionable ([zː] is not really discussed again in his work), but is taken up again in § 4.1.
Lexical Geminates (Newton 1967; Coutsougera 2003; Arvaniti 2010; Armosti 2011)

<table>
<thead>
<tr>
<th>Singletons</th>
<th>Geminates</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Initially</td>
<td>'pefti 'Thursday'</td>
</tr>
<tr>
<td></td>
<td>ka'fe 'brown'</td>
</tr>
<tr>
<td></td>
<td>'lia 'Lia (a name)'</td>
</tr>
<tr>
<td></td>
<td>'maθca 'cloaks'</td>
</tr>
<tr>
<td>b. Medially</td>
<td>xti'pun 'they hit'</td>
</tr>
<tr>
<td></td>
<td>'kokos 'Kokos (a name)'</td>
</tr>
<tr>
<td></td>
<td>'mila 'apples'</td>
</tr>
<tr>
<td></td>
<td>'ðoman 'flat roof'</td>
</tr>
</tbody>
</table>

Beyond lexical geminates, CyGr also presents derived, post-lexical geminates that come in two flavors. The first type consists of sequences of two identical consonants resulting from morphological concatenation (8a). These have often been referred to as ‘fake’ geminates (Hayes 1986). The second type involves geminates derived from total assimilation between two consonants (8b).

3.2 Armosti (2011) as the empirical basis of analysis
Recall that the literature on CyGr geminates, although relatively rich, has not reached consensus. For some, geminates are light and for others, heavy. Similarly, in some analyses they are treated as heterosyllabic, but elsewhere as tautosyllabic. In this study, I largely adopt the descriptive generalizations that Armosti (2011) suggests. This is primarily done for two reasons; first, because Armosti’s work remains to this date the most comprehensive study of CyGr geminates. Second, because it reflects the state of the language as still spoken. Many earlier phonological accounts, e.g. Malikouti-Drachman (1987) or Muller (2001, 2002), primarily based themselves on the landmark, but impressionistic, work of Newton (1967, 1972).

Armosti’s main conclusion is that CyGr geminates are heavy and, in fact, can be best viewed as moraic onsets, in the spirit of Topintzi (2010). To support this
idea, he points to relevant phonetic evidence on the basis of five experiments (one acoustic, one articulatory and three perceptual). One major finding is the higher durational stability across different places of articulation (POA) that CyGr geminates display in comparison to singletons both word-initially and word-medially. According to Ham (2001), this is exactly how Gs should behave given the presence of an inherent mora, which needs to materialize by means of some minimum duration thus taking precedence over any durational effects due to POA. Singletons bear no such mora to start with, which in turn allows for the emergence of durational effects due to POA.

Besides POA effects, additional acoustic cues for the geminate-singleton contrast can be identified. In terms of absolute timing, Armosti—in line with other previous phonetic studies on CyGr—finds longer duration of closure and longer aspiration for stops in geminates in comparison to singletons. Relative timing is also informative, and according to many (cf. Armosti 2011: §2.3.1.3.1 for discussion), provides a more reliable cue than absolute timing for durational contrasts, since it remains unaffected by changes of speaking rate. Unlike languages such as Japanese, Swiss German, or Hungarian, where it is the preceding vowel (V1) that is durationally affected by gemination, in CyGr, it is the following one (V2). Thus, while GV2 sequences are durationally longer than CV2 sequences, V2 is compensatorily shortened after a G, which again speaks in favor of the geminate’s moraicity. With regard to syllabification, Armosti argues that the moraic geminate must wholly syllabify as an onset (as V1.GV2) and not as a heterosyllabic coda-onset string (V1C.G.C.GV2), because in that case “the duration of the resulting geminate would have been regulated by both the V1C and CV2 syllable—but it was shown that only the CV2 syllable played such a role” (Armosti 2011: 92).

While it is true that the bulk of evidence for the moraicity of CyGr geminates is of phonetic nature, Armosti points to two post-lexical phonological pro-

---

3 Recently, Lee & Aso (2019) have suggested that onsets are moraic only when the closure in geminates is lengthened or when VOT is lengthened. In their view, Marshallese has moraic onsets due to gemination, Hateruma Yaeyama—their empirical case study—due to strong aspiration, and Cypriot Greek due to both.

4 Of the various ratios Armosti has measured, the most reliable for the singleton vs. geminate contrast has been the V2:HV2 quotient, followed by V2:CV2. V2:HV2 expresses the ratio of the following vowel to the aspiration-plus-vowel sequence and V2:CV2 the ratio of the following vowel to the stop-plus-vowel sequence. Broadly, when V2 occupies 80% of the HV2 sequence, it is preceded by an unaspirated segment, but when it occupies 50% of HV2, then it is preceded by an aspirated one. For V2:CV2, when V2 occupies 50% of the syllable it is preceded by a singleton, when it occupies 40% it is preceded by an unaspirated geminate and when it occupies 30% of the syllable, it is preceded by an aspirated geminate (Armosti 2011: 93).
processes that may lead to the same conclusion. Roughly speaking, these involve /n/-deletion and /i/-epenthesis in certain particles. Both processes are triggered by following geminates and clusters, but not by singletons, as shown in (9). Singletons instead trigger nasal place assimilation.

(9) 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>a. singletons</td>
<td>/en ˈpezːo/</td>
<td>[ˈɛmbɛzːo]</td>
<td>N/A</td>
<td>N/A</td>
<td>‘I don’t play’</td>
</tr>
<tr>
<td>b. clusters</td>
<td>/en ˈpsinːo/</td>
<td>N/A</td>
<td>[ˈeˈpsinːo]</td>
<td>[ˈeniˈpsinːo]</td>
<td>‘I don’t bake’</td>
</tr>
<tr>
<td>c. geminates</td>
<td>/en ˈpːefto/</td>
<td>N/A</td>
<td>[ˈeˈpʰːefto]</td>
<td>[ˈeniˈpʰːefto]</td>
<td>‘I don’t fall’</td>
</tr>
</tbody>
</table>

Several details have been reported with regard to these processes. First, as Armosti (2011: 272) states, both clusters and geminates trigger the deletion of a preceding nasal. The application of /i/-epenthesis is more restricted; it may occur with certain functional words ending in /n/ (and some ending in /s/), provided these are followed by verbs beginning with a geminate or a cluster. If the verb begins with a singleton, the option is unavailable. Συμεωνίδης (2006: 189–190) clarifies which functional words induce /i/-epenthesis, namely the negative particle /en/, but also the conjunction /an/ and the pronouns /ton/, /tin/, /mas/, /tus/ etc. The articles /tin/ and /ton/ do not trigger /i/ epenthesis, as syntactically they are not found immediately before verbs.

An alternative way to state this pattern is by means of syllable weight (Armosti 2011: 278): nasal deletion or vowel epenthesis—depending on specific morphophonological conditions—occur before heavy syllables. A reviewer wonders how a mora would cause a preceding nasal to delete or a vowel to be epenthized, presumably having in mind that weight sensitivity is normally associated to prosodic processes or phenomena. However, recent work has revealed interactions between segmental phonology too and syllable weight. For instance, Cardoso & Honeybone (To appear) report that in Liverpool

---

5 An exception to this are stop+sonorant clusters, as well as the cluster [pc], which behave like non-moraic singletons, e.g. [en ˈprepi] → [eˈmprepi] ‘it mustn’t’. Interestingly, the aspirated [pcʰ] resists post-nasal voicing and triggers nasal deletion instead (Armosti 2011: 285).

6 I am grateful to a reviewer for bringing these details to my attention and providing the necessary references.
English a back dorsal palatalises into [ç] following long front high vowels, but remains dorsal after their short congeners.

Finally, note that when the nasal precedes a word-initial geminate /lː/, as in /en lːios/ → [elːios], it adds extra duration to the geminate creating a ‘super-geminate’ (Payne & Eftychiou 2006). Such super-geminates can also be created under concatenation with geminate stops and affricates, thus /to ntʃip ’pːefti/ → [tondʒipʰːefti] ‘the jeep is falling’. For Armosti (2011) the attested aspiration on the stop is evidence for its underlying geminate status and the extra duration, a result of concatenation.

3.3 Focus on lexical geminates: justification
In the ensuing analysis, I focus on the CyGr lexical geminates of (7) and set aside derived geminates. Let me justify this analytical choice. I start with general issues that pertain to the phonetics-phonology mapping of geminates and then move on to CyGr-specific considerations.

One first point to consider is that the phonetic properties of lexical and derived geminates do not (have to) converge. To clarify, we know of cases where all types of Gs, i.e. lexical-concatenated-assimilated, are phonetically comparable (Bengali in Lahiri & Hankamer 1988), cases where lexical and assimilated Gs are phonetically similar to one another to the exclusion of concatenated ones (Tashlhiyt Berber in Ridouane 2010) and also cases where lexical and concatenated Gs are phonetically similar to one another to the exclusion of assimilated ones (Libyan Arabic in Issa 2015).

Turning to phonology now, we observe that in most accounts, these three types of geminates receive different representations. This is also true for the WCL model advanced here, where, broadly, lexical geminates come equipped

---

7 Note that these are not the only typologically occurring cases of derived geminates. Some languages demonstrate prosodically-derived gemination, where a consonant may lengthen to fulfil a desirable prosodic profile (cf. loanword adaptation in Japanese, Kubozono et al. 2009) or morphological gemination, where a consonant’s lengthening emerges as the exponence of a morphological category (as in the Woleaian denotative formation, Inkelas 2014).

8 It needs to be said that this mapping is in fact a complex stand-alone issue, which is well beyond the scope of the present paper. Here, only brief remarks are made. To my knowledge, no study has managed to offer a uniform proposal for such a mapping that works across geminate types and languages, particularly when prosodic and morphological gemination are also taken into consideration. While this is certainly a project for investigation, a one-size-fits-all approach is probably unlikely, especially since the increased length of geminates in comparison to singletons is the primary gemination cue, but not the only one. Several other secondary ones exist (e.g. vowel duration, F0, RMS amplitude, etc.), depending on the language under examination.
with an underlying link to a mora (regardless of whether this has a weight effect on the surface; compare (5) vs. (6)), which derived geminates lack. In turn, this suggests that they do not have to behave uniformly for phonological purposes. This prediction is borne out. For instance, Tigrinya post-vocalic velars and uvulars spirantize, but only when the consonant in question is a singleton or the first member of a false geminate (Kirchner 2001: 104). If it is a lexical or assimilated G, spirantization fails. This suggests that the two behave similarly. But, as Kirchner (2001: 108) observes, in Tiberian Hebrew a related spirantization process is blocked before fake geminates too. Again, this suggests that a different grouping of geminate patterning is possible phonologically speaking, where all of them behave uniformly, regardless of their source.

The consequence of the above is that there is no necessary correlation between phonological structure and phonetic interpretation, nor between phonological structure and phonological behaviour. These may differ on a language-specific basis.

So what about CyGr? Phonetically, all three geminate types are, according to Armosti, durationally the same, although lexical and assimilated Gs are further enhanced by means of length and intensity of aspiration (2011: 64), albeit not to the same degree. Thus, assimilated Gs present partial enhancement in comparison to lexical Gs (2011: 291). Phonologically, the situation is dubious; Armosti (2011: 274) is being explicit that the lexical and assimilated Gs pattern similarly for the purposes of palatalisation—so that the whole G palatalises, unlike a consonantal cluster where only the second member does—but does not clarify the situation for false Gs. However, elsewhere (2011: 9), he lists another possibility for the example in (8b), namely [ˈɛːɪ̯kçipɾɪaˈkon], which he considers to involve a false geminate, exactly because palatalisation affects its second half only, as it happens in clusters too. But this argument is rather circular; does it show that false geminates behave differently in palatalisation or since there is a difference in palatalisation this must be evidence for a false geminate?

---

9 The common patterning of these has led many people to ascribe a single term to lexical and assimilated geminates, namely ‘true’ geminates, which suggests that these should (always?) pattern together. In the light of Tiberian Hebrew mentioned in the text, I contend that this may well be a possibility, but not a necessary one. In the WCL model outlined earlier, the difference between an underlying and a derived geminate is always visible, but whether these will end up being treated by the phonology differently or not seems to be down to the nature of the process in question (see also Kirchner 2001: 131 who suggests that an abstract distinction between ‘true’ and ‘fake’ geminates does not seem required, at least on the basis of lenition processes affecting geminates) and/or attributed to specific markedness considerations at play each time.
Moreover, we have no information as to how postlexical Gs behave with respect to the /n/-deletion and /i/-insertion processes that we discussed earlier with regard to lexical Gs.

Ultimately, the task of making concrete claims, phonologically speaking, about derived Gs in CyGr is at the moment risky, and consequently, best avoided. One reason, as explained immediately above, is that not much about their phonological behaviour is known; a second reason is that although derived Gs have received some attention phonetically-speaking (Payne & Eftychiou 2006, Armosti 2011, among others), this phonetic information is insufficient for broader phonological generalizations. Arvaniti (2010) comments that she does not discuss “a phonological account that would take the results of Payne & Eftychiou (2006) into account” since they “have not yet been shown to apply to the Cypriot system at large”. The same point carries over to the present work too, given Armosti’s admission that “most of the post-lexical geminates used in this study were concatenated geminates, while only one was an assimilated geminate” (2011: 59).

In sum, due to the reasons stated, I take a conservative approach and confine myself to the analysis of CyGr lexical geminates, which we have sufficient information for. This also serves well for the typological perspective of the paper, as it makes it directly comparable to the cross-linguistic studies available in the literature, where lexical geminates dominate (see Hayes 1989, Tranel 1991, among others).

4 Analysis

4.1 Main analysis

Using Armosti’s (2011) meticulous work as the empirical backbone of the present study, this section supplies the core analysis of CyGr Greek singletons and lexical geminates across all positions. Recall that geminates are considered heavy and emerge intervocally, as well as initially, but (almost) never finally. Singletons on the other hand are light.

To achieve this result, we need to employ the constraints in (4), repeated in (10) for convenience, enriched with some additional ones. These are typically either well-known constraints, such as NoCoda or Max-C or positional variants of the constraints in (10) tailored to edge positions (cf. (14)). Less familiar constraints are defined as they are introduced.
(10) Max-$\mu$  
$\mu > \sigma$  
$^*C^\mu$  

Do not delete an input mora in the output  
Every mora should be dominated by a syllable  
A consonant should not be dominated by a prosodically integrated mora

Let us start with the consideration of simple singletons, as emerging in a coda position. As explained earlier, these are by definition non-moraic in the input and, in CyGr, remain so also in the output. Tableau (11) illustrates how.

(11) Non-moraic consonants emerge as weightless singletons

<table>
<thead>
<tr>
<th>/V$^\mu$C.CV$^\mu$/</th>
<th>Max-C</th>
<th>NoCoda</th>
<th>$^*C^\mu$</th>
<th>WBP</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. V$^\mu$C.CV$^\mu$</td>
<td>*</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>b. V$^\mu$C$^\mu$.CV$^\mu$</td>
<td>*</td>
<td>*</td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>c. V$^\mu$.CV$^\mu$</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Candidates (11a) and (11b) both violate NoCoda, unlike (11c) which escapes such violation through deletion, but in doing so, it violates high ranked Max-C.\textsuperscript{10} Codas that end up heavy (11b) do so at the expense of $^*C^\mu$. Originally weightless consonants that surface as light codas (11a) violate WBP, the well-known Weight-by-Position constraint, which in the context of WCL is defined as follows:

(12) WBP  
Assign a violation mark for every coda consonant that is not dominated by a prosodically integrated mora

Given the ranking argument $^*C^\mu >> WBP$, (11a) is thus the proclaimed winner. How about geminates? I begin by examining word-initial geminates. In a manner comparable to the situation in (3) for intervocalic geminates, initial geminates too start life as moraic, but may adopt one of the following three output representations.

\textsuperscript{10} A [V.CCV] candidate would also escape such a violation, but I have left it out, since my interest here is on geminates and their relationship towards singletons, rather than syllabification of singletons/clusters per se. More on tautosyllabic clusters in § 4.2.
Possible outputs for an underlying initial geminate

<table>
<thead>
<tr>
<th>Underlying:</th>
<th>( \mu )</th>
<th>V ( \nu )</th>
<th># C V</th>
<th>abbreviation:</th>
<th>( C^\mu V^\nu )</th>
<th>( C^{(\mu)}V^\nu )</th>
<th>CV( ^\nu )</th>
</tr>
</thead>
<tbody>
<tr>
<td>interpreted as:</td>
<td>G</td>
<td>G</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>contributes to ( \sigma ) weight:</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(13a) is interpreted as a heavy long geminate in the guise of a moraic onset; the mora’s upward relationship to a syllable node guarantees heaviness, while its downward underlying link to a segment ensures increased length. (13b)—abbreviated as \( C^{(\mu)}V^\nu \)—is still interpreted as a long geminate by virtue of its underlying mora-to-segment link, but lacks weight, as its mora is prosodically unintegrated. Finally, (13c) illustrates neutralization to a singleton, since the underlying mora has been entirely removed. CyGr, like Pattani Malay or Trukese (Topintzi 2010), possesses initial geminates of the type in (13a), in contrast to e.g. Leti or Swiss German (Hume et al. 1998, Kraehenmann 2001, 2003), which exhibit weightless initial geminates. Obviously the majority of languages, including Standard Greek, present the pattern in (13c).

To generate the CyGr pattern in (15a), we require the positional constraints \#\( \mu \)\( > \sigma \) and Max-\#\( \mu \) (see (14) for definitions) to both dominate \( ^* \text{Moraic Onset} \) (represented with the shorthand \( ^* \text{e}_\mu (C^\mu) \)) which naturally militates against moraic onsets. Obviously, the generic constraints \( \mu \)\( > \sigma \) and Max-\( \mu \) are violated too, but are presently omitted so that the emphasis is placed on the positional aspects of the ranking. We include them later in the full tableaux.

### (14) L-Edge-specific constraints

- \( \text{Max-}\#\mu \)  Do not delete a word-initial input mora in the output
- \( \#\mu \)\( > \sigma \)  Every word-initial mora should be dominated by a syllable

### (15) Initial moraic consonants emerge as weightful geminates (moraic onsets)

<table>
<thead>
<tr>
<th>( /C^\mu V^\nu/ )</th>
<th>( #\mu )( &gt; \sigma )</th>
<th>( \text{Max-}#\mu )</th>
<th>( ^* \text{e}_\mu (C^\mu) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( ^* )</td>
<td>a. ( #C^\mu V^\nu )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. ( #C^{(\mu)}V^\nu )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. ( #CV^\nu )</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\( ^* \)! indicates violation.
Next, we consider the medial intervocalic geminates of CyGr. The possible outcomes for a /\textit{V}\text{μ}C\text{μ}V\text{μ}/ input are presented in (16). These include the representations in (3) with one addition. Thus, so far we have seen that intervocalic underlying geminates may emerge as weightful or weightless (Output 3), and they may even neutralize to singletons (Output 4). The weightful variety typically involves geminates that straddle syllable boundaries (Output 1), i.e. [\textit{V}\text{μ}C\text{μ}V\text{μ}]. The main advantage of this is that the lexical mora dominates a coda, and not an onset consonant, which is universally dispreferred (cf. *_{σ}[C\text{μ}]). Satisfication of ONSET is simultaneously achieved through the epenthetic link to the following syllable. Indeed, the vast majority of languages opt for this structure. However, that does not mean that [\textit{V}\text{μ}C\text{μ}V\text{μ}] is flawless.

\begin{center}
\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
Underlying: & Output 1 & Output 2 & Output 3 & Output 4 \\
\hline
\textit{V}\text{μ}C\text{μ}V\text{μ} & G & G & G & C \\
\textit{V}\text{μ}C\text{μ}V\text{μ} & yes (σ 1) & yes (σ 2) & & \\
\hline
interpreted as: & add σ weight & & G & no \\
\hline
\end{tabular}
\end{table}
\end{center}

Crucially, it violates the constraint *_{σ}S_{σ}, as defined below.\textsuperscript{11}

\begin{equation}
\ast_{σ}S_{σ}
\end{equation}
Assign a violation mark for a pair of syllable nodes that dominate the same segment

If this constraint is highly ranked and preservation of the lexical mora is imperative, then Output 2 in (16) becomes a better option. Effectively then, the geminate emerges as a moraic onset \textit{medially}, i.e. [\textit{V}\text{μ}.C\text{μ}V\text{μ}]. Such representation for medial heavy geminates, albeit rarely occurring, is argued to exist in languages like Marshallese (Topintzi 2008), Ndjébbana (Wolf 2008), Tedumuni Okinawan (Shinhara & Fujimoto 2011), Ngigua, as spoken in San Marcos Tlacoyalco, Mexico (Bell 2015) and possibly, Puluwat and Eastern Popoloca, as spoken in San Juan Atzingo, Mexico (Topintzi 2010).

\textsuperscript{11} This is pretty much equivalent to Itô & Mester’s \textsc{CrispEdge}[σ] constraint (1995: 38). As I do not wish to commit myself regarding the \textsc{CrispEdge} family of constraints they propose, I stick to the more neutral *_{σ}S_{σ}.
Given Armosti’s findings, and in particular the fact that durational effects were observed on the vowel after the geminate and not on the one before—as happens in languages with a $[\text{VC}^\mu\text{V}^\mu]$ representation—CyGr Greek renders itself a strong candidate for a moraic onset representation for intervocalic geminates, along the lines of Output 2. The tableau below exemplifies the required ranking.

(18) Intervocalic moraic consonants emerge as weightful geminates (moraic onsets)

<table>
<thead>
<tr>
<th></th>
<th>$\star \sigma S^\sigma$</th>
<th>$\mu &gt; \sigma$</th>
<th>MAX-$\mu$</th>
<th>$\star [\text{C}^\mu]$</th>
<th>$\star \text{C}^\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. $\text{V}^\mu\text{C}^\mu\text{V}^\mu$</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. $\text{V}^\mu\text{C}^\mu\text{V}^\mu$</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>c. $\text{V}^\mu\text{C}(\mu)\text{V}^\mu$</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. $\text{V}^\mu\text{C}(\mu)^\mu$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

High-ranking $\star \sigma S^\sigma$ eliminates the heterosyllabic weightful geminate (18a), $\mu > \sigma$ dispenses with the weightless geminate (18c) and MAX-$\mu$ knocks out the neutralized-to-a-singleton candidate (18d). Candidate (18b), the intervocalic moraic onset geminate, wins by virtue of the low ranking violations that militate against prosodically integrated moraic consonants generally ($\star \text{C}^\mu$), and onsets, particularly ($\star [\text{C}^\mu]$).

I have so far considered underlyingly light consonants, i.e. singletons, as well as geminates in initial and medial position. This leaves us with one more pattern; final geminates. Again, the basic structures predicted to occur are a heavy final geminate (19a), a light final geminate (19b) and a light neutralized singleton (19c).

(19) Possible outputs for an underlying final geminate

<table>
<thead>
<tr>
<th>Underlying:</th>
<th>a. $\text{V}^\mu\text{C}^\mu$</th>
<th>b. $\text{V}^\mu\text{C}(\mu)$</th>
<th>c. $\text{V}^\mu\text{C}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>annotation</td>
<td>G</td>
<td>G</td>
<td>C</td>
</tr>
<tr>
<td>contributes to $\sigma$ weight:</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>

Downloaded from Brill.com10/26/2022 08:35:39AM via free access
CyGr lacks final geminates on the surface, with one possible exception, that of final [ʦʰː], to be discussed momentarily. The implication here is that any such underlying geminates neutralize to singletons in the output (cf. 19c/21c). To ensure that, we need to employ the constraint in (20) in addition to the constraints we have already examined.

\[(20) \text{R-Edge-specific constraint} \]
\[^*C_\mu \# \quad \text{A word-final consonant should not be dominated by a prosodically integrated mora} \]

Crucially, both $\mu > \sigma$ and $^*C_\mu \#$ should dominate $\text{Max-}\mu$ so that deletion of the underlying mora is preferred over maintaining it under the syllable node (as in (21a)) or maintaining it but leaving it prosodically unintegrated (as in (21b)). One more reasonable candidate to consider is the one in (21d), where the consonant sponsoring the underlying mora is deleted altogether, but this violates high-ranking $\text{Max-C}$.

\[(21) \text{Final moraic consonants neutralize to weightless singletons} \]

<table>
<thead>
<tr>
<th>/V$\mu$C$\mu$#/</th>
<th>$\mu &gt; \sigma$</th>
<th>$^*C_\mu #$</th>
<th>$\text{Max-C}$</th>
<th>$\text{Max-}\mu$</th>
<th>$\text{NoCoda}$</th>
<th>$^*C_\mu #$</th>
<th>$\text{WbP}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. V$\mu$C$\mu$#</td>
<td>$^!$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. V$\mu$C($\mu$)#$^!$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. V$\mu$C</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>d. V$\mu$</td>
<td></td>
<td>*$^!$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Returning to the issue of the final [ʦʰː], this was found to be much longer than other final singletons. Armosti (2011) takes this as evidence that it is indeed a geminate consonant. There are at least two observations that render this conclusion questionable, both of which are acknowledged by him (2011: 86). The first is that the alveolar affricate lacks a singleton counterpart. Given the absence of phonological contrast, should this then be interpreted phonologically as gemination? Perhaps yes, since word-initially—where gemination

\[\text{There is at least one more reasonable candidate to consider, on the basis of typological considerations, namely [V}_\mu(\langle C \rangle)], where the final consonant is extrametrical and gets to be integrated directly under a higher prosodic node like the foot or word node. This can be ruled out by a constraint such as C$>\sigma$ which asks that consonants be dominated by syllables. To avoid overburdening the tableau unnecessarily, I omit this here.} \]
contrasts are well-established for other consonants—[ʦʰː] has a comparable duration to (other) geminates in the same position. The second observation is possibly more challenging. In the case studied, [ʦʰː] appeared word-finally, but also utterance-medially before a vowel-initial word. In effect, it could thus have resyllabified as an intervocalic geminate. The upshot is that the status of [ʦʰː] as the only final lexical geminate of CyGr is at best dubious. If it is however validated, then the ranking in (21) would have to be slightly modified; one possible solution could involve a top-ranking contextual Max-μ constraint—sensitive to moras associated to the alveolar affricate—which would protect /ʦː/ from losing its mora, even at the expense of emerging as a geminate, unlike other underlying geminates in the same position.

The tableaux just drawn can be summarized by means of a Hasse diagram, as in (22). To demonstrate which rankings are formed on the basis of which tableau, the lines and the numbered examples are colored in orange, green, red and blue, for tableaux (11), (15), (18) and (21), respectively.

(22) Rankings established

Finally, (23) puts together all the patterns discussed in one grand tableau, which provides the core CyGr grammar for geminates and singletons in the WCL approach.
(23) Overall ranking for all patterns

<table>
<thead>
<tr>
<th></th>
<th>*σSσ</th>
<th>μσ</th>
<th>*Cμσ</th>
<th>Max-μ</th>
<th>Max-μ</th>
<th>*μ[σμ]</th>
<th>*Cμ</th>
<th>NoCoda</th>
<th>WbP</th>
</tr>
</thead>
<tbody>
<tr>
<td>/VpCCV#/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>i. Non-moraic C: Weightless singleton</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. Vp.C.CV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Vp.C[μ]V</td>
<td>*</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. Vp.CV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| /#CpV#/ |     |     |       |       |       |        |     |        |     |
| ii. Initial moraic C: Weightful geminate (moraic onset) |
| a. #CpV | * |     |  *    |     |       |        |     |        |     |
| b. #C[μ](p)V | * |     |  *    |     |       |        |     |        |     |
| c. #CV |       |     |       |     |       |        |     |        |     |

| /VpCpV#/ |     |     |       |       |       |        |     |        |     |
| iii. Intervocalic moraic C: Weightful geminate (moraic onset) |
| a. Vp.CpV | * |     |  *    |     |       |        |     |        |     |
| b. Vp.C[μ]V | * |     |  *    |     |       |        |     |        |     |
| c. Vp.C[μ](p)V | * |     |  *    |     |       |        |     |        |     |
| d. Vp.CV |       |     |       |     |       |        |     |        |     |

| /VpCp#/ |     |     |       |       |       |        |     |        |     |
| iv. Final moraic C: (Neutralized) weightless singleton |
| a. Vp.C[p] | * |     |  *    |     |       |        |     |        |     |
| b. Vp.C[μ](p) | * |     |  *    |     |       |        |     |        |     |
| c. Vp.C |       |     |       |     |       |        |     |        |     |
| d. Vp |       |     |       |     |       |        |     |        |     |

4.2 Outstanding issues: onset clusters and weight

In the previous discussion, I assumed that a medial cluster in a string such as /VCCV/ was syllabified heterosyllabically as [VC.CV] and that the coda so produced bore no weight (cf. (11) and (23i)). Although it is accepted that CyGr allows for final codas, as in words like [ˈfo.ɾos] ‘tax’ (final codas are mainly /s/, but also /n/ and more peripherally /ɾ/), the situation on medial codas is debatable. For example, Malikouti-Drachman (2000, 2001) takes a sonority-driven approach, where effectively, CC clusters are complex onsets if they present rising sonority, and otherwise are split into two syllables, thus [plex.ˈto] or [kar.ˈca]. Coutsougera (2003) on the other hand adopts a proposal where virtually all clusters in CyGr make complex onsets satisfying Onset Maximization. This holds true even for most CC clusters of falling sonority, thus [ple.ˈxto]. A notable exception is the medial /ɾ/; while it is “not firmly established as a coda or onset element in ɾ+obstruent clusters word-internally”, it tends to be pushed
in the coda position (Coutsougera 2003: 23). The interested reader may consult the aforementioned works for arguments relevant to the distribution and application of phonological processes in favor of one or the other position.

While I do not take any firm stance on this issue—partly because it is too complex to address here and partly because it is tangential to the issue at hand, namely that CyGr possesses geminates, which are best analysed as moraic onsets, whether initial or medial—it seems that regardless of the specific details, we must accept that CyGr admits both heterosyllabic and tautosyllabic clusters. The former, as discussed, lead to weightless singleton codas (also word-finally). But what about the latter?

Recall from (9) that initial Gs and some onset clusters trigger postlexical n-deletion or i-epenthesis, unlike onset singletons and some other clusters which instead trigger nasal place assimilation. I have followed Armosti (2011) in interpreting this discrepancy as a weight effect, so that heavy syllables trigger deletion/epenthesis and light ones do not. How this weight comes about in geminates has been the topic of much of the earlier discussion. With regard to onset clusters now, we need to say that there are two types of onset clusters in CyGr: (a) obstruent+obstruent ones13 which are rendered heavy on the surface, vs. (b) obstruent+sonorant ones which are deemed light.14 Note that there is some preliminary work by Χριστοδούλου (1967), cited in Armosti (2011: 284) in favour of elongation of the second consonant in consonant clusters, which may be compatible with this idea.15

Such a situation is not unprecedented. Topintzi (2010) argues that besides moraic onset geminates like in Pattani Malay and Trukese or, as claimed here, CyGr, onset moraicity can be a derived property that affects singletons. In languages like Pirahã, Arabela, Karo (Topintzi 2010) or Tümpisa Shoshone (Dayley 1989), the quality of onset matters, so that voiceless or obstruent onsets behave as heavier than voiced or sonorant onsets with regard to stress. In Arrernte (Topintzi & Nevins 2017), it is the presence of the onset that is pivotal; onsetful syllables are considered heavy, whereas onsetless ones are light in a number of prosodic phenomena, including allomorph, stress and language games.

---

13 With the caveat mentioned in fn. 5. Note that such syllabification is compatible with Cout-
sougera, but not with Malikouti-Drachman.

14 For Coutsougera (2003: 16) even a sonorant+obstruent initial cluster is a branching onset, e.g. [ɾtõnno] ‘rise, achieve’ which is preferred over its alternative [oɾtõnno].

15 At the time of writing, I had no access to the original work by Χριστοδούλου (1967) and Armosti only provides examples with obstruents, but no obstruent+sonorant clusters. As a reviewer correctly points out, this is all impressionistic and should be treated with caution. Armosti (2011: 285) in fact calls for empirical confirmation of this elongation. To my knowledge, this has not yet been performed.
The case illustrated in CyGr, namely that (certain) onset clusters appear heavy but onset singletons as light, has until recently only marginally been argued by Gordon (2005) for Bislama and Nankina. Topintzi (2010) offers some additional cases from Damin, Luganda, and Supyire, but observes that these—alongside Bislama and Nankina—are subject to re-interpretation that does not involve onset weight. Note that all the above onset weight cases involved categorical effects. A major breakthrough that furnishes the potential importance of onset size—rather than its presence or quality—has been accomplished in the work of Ryan (2014) and, later, Garcia (2017) on gradient onset weight. For instance, Ryan (2014) shows that in both English and Russian a strong correlation between stress and onset size emerges in simplex words; in trisyllables for example, the larger the word-initial onset (where CCC>CC>C>∅), the more likely the initial position for stress.

CyGr thus emerges as a compelling case, since the heaviness of (some) complex onsets over light singletons is arguably an instance of categorical weight. Interestingly, at the time of writing, I have become aware of an additional instance of the sort. On the basis of stress data, Smith & Lubera (2021) argue that in Iron Ossetian onset complexity (and not merely the presence of an onset) determines syllable weight. Also, like CyGr, codas in Iron Ossetian are inconsequential for weight purposes. The cases of CyGr and Iron Ossetian certainly prod us to re-consider more closely Bislama and the other languages where onset complexity for categorical weight has been questioned. It may ultimately be the case that the possibility for heaviness in complex onsets has been right all along!

Going back to CyGr, let us briefly consider how the complex onset moraicity data can be incorporated into the approach developed so far. I will keep to the gist of the analysis, setting aside various details and technicalities which will need to be worked out in future work, also taking into account a more thorough examination of other languages with moraic complex onsets. An initial observation is that the weight of onset clusters is necessarily derived on the surface, since these involve a cluster of two singletons, which by definition are underlyingly light. Given the empirical necessity of surface weight assignment on both codas and onsets, Topintzi & Nevins (2017) suggest that the familiar WbP constraint is more accurately Weight-by-Position\textsubscript{CODA} which as we have seen in CyGr already, is very low-ranked (cf. (23)) and argue for a Weight-by-Position\textsubscript{ONSET} constraint with comparable effects for onsets. Claiming though that WbP\textsubscript{ONS} should be relatively high-ranked in CyGr is problematic. This is because it would induce onset weight on any non-geminate onset, including plain singletons, an undesirable result. One solution to this issue would be to argue in favor of a specific version of WbP\textsubscript{ONS} which only assigns weight to com-
plex onsets, namely $WbP_{\text{comp,ons}}$. Under a ranking such as $WbP_{\text{comp,ons}} >> \ast \sigma[C^\mu] >> WbP_{\text{ons}}$, complex onsets would thus acquire weight, but not the singleton ones. The updated ranking for CyGr appears in (24).

\[
\begin{array}{c}
\ast \sigma
\end{array}
\]

\[
\begin{array}{c}
\mu > \sigma, \ast C^\mu \# \text{, MAX-C} \\
\text{MAX-}\mu
\end{array}
\]

\[
\begin{array}{c}
\ast [C^\mu, C^\mu, \text{NoCODA}]
\end{array}
\]

\[
\begin{array}{c}
WbP_{\text{ons}} \\
WbP_{\text{coda}}
\end{array}
\]

16 Obviously, further adjustments are in order to capture the fact that only some complex onsets appear weightful. Others pattern with singletons in being light. Interestingly though, the heavy complex onsets involve obstruent+obstruent clusters as opposed to the light obstruent+sonorant ones. Typologically speaking, it is indeed obstruent (or voiceless) consonants that make better heavy onsets (see Topintzi 2010, Ryan 2019).

17 This raises the question of whether comparable weight asymmetries may emerge in singleton vs. complex codas. They surely do, but their theoretical interpretation is a debatable issue. In one pattern for instance, final CVCC syllables attract stress, but final CVC fail to do so. Many scholars interpret this as final consonant extrametricality, especially in various Arabic dialects (see e.g. Gordon et al. 2010 and references therein), i.e. $\text{CVC}(\text{C})$ vs. $\text{CV}(\text{C})$ where only the former syllable can be rendered heavy due to $WbP$. Others, such as Shaw (1985) on Stoney Dakota, claim that CVCC is genuinely heavy as opposed to light CVC or CV. Another pattern is the one between heavy CVC and superheavy CVCC (and typically also heavy CVV vs. superheavy CVVC), as in German, which for Hall (2002: 384) contrast in being bimoraic and trimoraic, respectively.
5 Discussion and typological extensions

In this section, I address some of the issues WCL and/or its implementation in CyGr raises when seen from a broader or typological perspective. Due to length limitations, the discussion necessarily is brief. The interested reader is however directed to Topintzi & Zimmermann (In prep.), which includes a much more thorough investigation of these and other typological concerns, as well as a factorial typology that illustrates most of the points outlined below.

One first issue has to do with an asymmetry pertaining to edge geminates. In particular, why do we need separate positional constraints (e.g. *Cμ# and *σ[Cμ alongside the generic *Cν) and why are the ones devoted to final geminates more restricted, so that the left-edge-sensitive #μ>σ exists, but its right edge counterpart μ#>σ does not? In the analysis of CyGr for instance, at least some of the positional constraints advocated before could be eliminated at no detrimental effect, although not all; in (23iv), for example, use of positional *Cμ# is crucial so that by outranking Max-μ >> *Cμ, neutralization to a singleton (23iv.c) obtains.

There are two main considerations here: (a) the various geminate combinatorial possibilities and (b) the uniformity of geminate weight vs. lack thereof. Recall that according to the present proposal, CyGr singleton codas are light, whereas geminates—initial or medial—heavy. Final geminates are virtually absent, which entails that underlyingly moraic consonants in a final position neutralize to singletons. Wolof (Diouf 2017) is sort of the mirror image, since it too has light singleton codas, but heavy medial and final geminates. The absent initial geminates point to singleton neutralization in that position. And of course, the majority of languages only allows for one type of geminates, namely the medial intervocalic ones, whether weightful or weightless, suggesting that underlyingly moraic consonants initially and finally neutralize to singletons. Additional combinations are obviously available, but the above suffice to illustrate that admittance of geminates in certain positions paired with their ban in other positions within the word speak for the need of positional constraints relevant to syllable structure.

Further corroboration is provided by the (non-)uniformity of consonantal weight. Focusing for a moment on geminates only (setting singletons aside), in CyGr, or Wolof for that matter, geminate weight has been uniform in that geminates across different positions are systematically heavy. But this is by no means a necessity cross-linguistically. The most celebrated example of the sort is probably Thurgovian Swiss German where geminates are found in all positions (initial, medial, and final). On the basis of stress and word minimality data, Kraehenmann (2001, 2003) offers convincing evidence that initial gemi-
nates are light, whereas medial and final ones are heavy. Again, to account for the different behaviour of geminates depending on the position they are found in, corresponding positional constraints are in order. Turning to singletons now, these may also be weightful or not and can co-occur with both weightful and weightless geminates. Thurgovian is a fitting example in that respect too. In addition to geminates, it also possesses coda singletons that are heavy medially, but extrametrical finally. *C# in particular and its positional variants, here *Cμ#, are thus also relevant for singletons, since these constraints militate against consonantal moraicity, whatever its source, underlying or surface.

Having established the necessity of several positional constraints on mora markedness and faithfulness on the grounds of empirical adequacy, I now turn to the issue of a positional asymmetry; a (positional) faithfulness effect for the preservation of initial moras arises through the existence of \( \text{MAX-}\#\mu \) and \#\( \mu > \sigma \), but no comparable constraints are advocated for final moras. This decision is also driven by empirical considerations. For example, this choice of constraints captures the implicational universal according to which “final contrastive geminates always imply intervocalic geminates in the language” (Dmitrieva 2012: 166). Effectively, we know of no language that only admits final geminates. No such correlation exists however between initial and intervocalic geminates. Languages where the sole type of geminates is the intervocalic ones exist (e.g. in Latin), as do languages where the only geminates present are the initial ones (e.g. in Pattani Malay).

One more feature that emerges in the analysis of CyGr and is pivotal for WCL is the independence of singleton vs. geminate weight. Lexical geminates, as we have seen, are underlyingly moraic; lexical singletons on the other hand are not, but may acquire weight on the surface as a response to WbP satisfaction. That means that lexical geminates can always be teased apart from singletons, even if representationally they look very similar. A prime example of that is the case of final heavy geminates vs. final heavy singletons, as reported to contrast in languages like Hungarian (Curtis 2003). Both are heavy due to the presence of a mora, but in singletons (25b), this mora is epenthetic, as indicated by the mora’s gray shading. Further, the consonant in (25a) is interpreted as a long geminate by virtue of its underlying link to a mora. This representational distinction is a welcome result on empirical grounds, since previous accounts that dealt with final geminates, such as Ham (2001), failed to capture the contrast in question.

---

18 The pattern is further complicated by sonorancy factors in both Gs and Cs that are set aside here.
Roughly speaking and abstracting away from positional effects, a singleton will remain light if \( ^*C^\mu > WbP_{coda} \) but will gain weight under the reverse ranking \( (WbP_{coda} >> ^*C^\mu) \). A geminate on the other hand will remain heavy under \( \mu > \sigma \), \( \text{Max} - \mu, ^*C^\mu \), but will emerge as light if \( \text{Max} - \mu, ^*C^\mu >> \mu > \sigma \). It should be evident that the interplay between these two parts of the analysis produces all possible—and simultaneously attested—combinations; one category involves languages where geminates and singletons are uniformly heavy (e.g. Latin), another where both are light (e.g. Selkup). A third one yet emerges when geminates are heavy, but singletons are light. In §1, I offered Koya as a relevant example. CyGr, the empirical focus of this study, is another, and indeed at its core, the analysis includes all these components, as sketched in (26).

(26) Bare bones of CyGr analysis

a. Top-ranked \( ^*C^\# \rightarrow \) blocks heaviness from final consonants of any type

b. \( \mu > \sigma, \text{Max} - \mu, ^*\# > \sigma, \text{Max} - ^*\# > ^*\sigma \rightarrow ^*C^\# \rightarrow \) allows for heavy initial and medial geminates

c. lowest ranked \( WbP_{coda} \rightarrow \) ensures that singletons are light

While rare, the fourth logical possibility is also attested. This involves languages, e.g. Ngalakgan, with weightless geminates and weightful singletons. The main evidence for this pattern comes from quantity-sensitive stress (see Baker 2008). For example, in trisyllabic words with light open syllables only, stress is placed on the first syllable, schematically \([ˈLLL]\). Stress shifts to the second syllable if it contains a coda, provided the consonant in that position is heterorganic to the next of the following onset, thus \([LˈHL]\). If it is the first member of a homorganic NC cluster or a geminate, then no stress shift applies, suggesting a \([ˈLLL]\) form. On the basis of their single articulatory gesture, Baker (2008: 188) syllabifies geminates (and homorganic NC clusters) as monogestural onsets, a
practice adopted here too. In other words, medially Ngalagkan is argued to possess heavy singleton codas and light geminate onsets.

The Ngalagkan pattern—also found in neighbouring Australian languages such as Marra, Rembarrnga, and Kuninjku (Baker 2008: 181)—is notorious in the geminates literature, mainly because neither the ‘weight’ nor the ‘length’ approach of geminates may derive it. The former falls short because it views geminates as heavy and singletons as light or heavy depending on whether $\text{WbP}_{\text{coda}}$ is applicable or not. Ngalagkan constitutes a problem since its geminates are light. The length approach also fails, because it ties weight with application of $\text{WbP}_{\text{coda}}$. If $\text{WbP}_{\text{coda}}$ is responsible for the weight of Ngalagkan singletons, then it should also posit weight on Ngalagkan geminates, against what actually happens. WCL instead, by treating geminate and singleton weight independently from one another, is in a unique position to account for Ngalagkan. The tableau in (27) demonstrates how.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|}
\hline
 & $\text{WbP}_{\text{coda}}$ & $^*C_\mu$ & $\text{Max}-\mu$ & $\mu>\sigma$ \\
\hline
\hline
$V^\mu CCV^\mu$ & i. Weightful medial singleton & & & \\
\hline
\hline
a. $V^\mu C.CV^\mu$ & $^*$ & & & \\
\hline
\hline
$V^\mu CV^\mu$ & ii. Weightless medial geminate & & & \\
\hline
\hline
\hline
a. $V^\mu C^\mu V^\mu$ & $^*$ & & & \\
\hline
\hline
b. $V^\mu C(\mu)V^\mu$ & & & $^*$ & \\
\hline
\hline
c. $V^\mu C.V^\mu$ & & & $^*$ & \\
\hline
\end{tabular}
\end{table}

(27) Ngalagkan heavy singleton codas vs. light onset geminates

The ranking $\text{WbP}_{\text{coda}} \gg ^*C_\mu$ will enforce an epenthetic mora on a coda singleton so that (27i.b) is the proclaimed winner. Underlying moras are however treated differently; $^*C_\mu$ blocks a heavy straddled geminate from emerging (27ii.a), while $\text{Max}-\mu$ militates against the underlying mora’s deletion as in (27ii.c). (27ii.b), the light onset geminate, wins. By retaining its underlying mora as prosodically unintegrated in an onset position, it escapes the violation of $\text{WbP}_{\text{coda}}$ at the mere cost of low-ranked $\mu>\sigma$. Such ranking then permits the emergence of derived weight in coda singletons, while masking underlying weight in lexical geminates. Geminate length is still identifiable, though, by the preservation of a link to an underlying mora.\textsuperscript{19}

\textsuperscript{19} A complete discussion of Ngalagkan which takes into consideration the lack of geminates.
One final concern, raised by one of the reviewers, is whether WCL predicts that consonantal length is always underlying. The answer here is negative. WCL treats the increased length seen in **lexical** geminates as a consequence of their underlying mora. But, this does not mean that lengthening cannot be a derived property (see also end of § 2 and § 3.3). Indeed, ‘fake’ and assimilated geminates, among other derived geminates, emerge as long. They are however representationally different from lexical geminates, and phonology may or may not treat them on a par. While Topintzi & Zimmermann (In prep.) suggest a fuller set of principles regarding the phonetic interpretation of geminates (lexical and derived), the fact is that this issue remains open to further investigation, subject to a better understanding of the phonetics-phonology mapping in gemination, an understanding we still lack.

6 Conclusion

Building on Armosti’s (2011) phonetic work on CyGr, which establishes moraicity for geminates and (some) onset clusters, I have put forward here a phonological analysis of CyGr weight that follows the ‘weight and concomitant length’ (WCL) approach, according to which lexical geminates are linked to an underlying mora. The mora’s links to higher and lower structure are associated to weight and length, respectively. The ‘upward’ link to a syllable node on the surface ensures that the geminate will be interpreted as heavy. Otherwise it will be deemed light. The mora’s downward link to the consonant ensures that it is interpreted as long. If this underlying link is severed, then neutralization to a singleton obtains. Underlying singletons on the other hand may acquire weight on the surface as a response to a WbP constraint. Effectively then the weight of geminates and that of singletons are independent from one another, allowing for the production of intricate patterns of geminates and singletons across different positions within the word.

This system can thus generate the CyGr pattern which involves light onset and coda singletons, (some) heavy onset clusters, heavy initial and medial onset geminates, and final neutralized geminates to singletons. If this interpretation of the data is correct, then CyGr exemplifies a unique case of onset weight, as it is—to my knowledge—the first language to date to exhibit both lexical and derived heavy onsets, in the guise of geminates and complex onsets, at edges, as well as the presumed weightlessness of final singletons can be found in Topintzi & Zimmermann (In prep.).
respectively. In fact, if there is anything ‘atypical’ in CyGr—to borrow from Arvaniti (2010)—it is not its geminates per se, but rather their co-existence with derived heavy onsets. This fascinating result certainly enriches the steadily expanding typology of onset weight systems and asks us to reconsider the issue of categorical weight on the basis of onset complexity.

Acknowledgments

I am grateful to two anonymous JGL reviewers for instructive and helpful feedback, much of which has been incorporated in the paper. Thanks also to the JGL editors Anna Roussou and Brian Joseph for their help at various steps of the submission-to-publication process. Any remaining errors or misconceptions are my sole responsibility.

References


Κοντοσόπουλος, Νικόλαος. 2000. Διάλεκτοι και ιδιώματα της Νέας Ελληνικής. Αθήνα: Εκδόσεις Γρηγόρη.


Shinohara, Shigeko & Masako Fujimoto. 2011. Moraicity of initial geminates in the Tedu-


