

A (not so) paradoxical instance of compensatory lengthening: Samothraki Greek and theoretical implications*

Nina Topintzi

University College London, University of Patras and University of Thessaloniki

Samothraki Greek onset-/r/ deletion with subsequent compensatory lengthening (CL) of the following vowel poses two major problems for current phonological theory. First, such a pattern should be impossible because in moraic theory onsets never bear weight, thus - under the assumption that CL is all about mora conservation (cf. Hayes 1989 and several others) - their deletion should not induce lengthening. Second, CL is an instance of opacity and opacity is the single biggest conundrum that Optimality Theory faces. This paper addresses both issues and suggests that CL should not be treated as mora conservation, but instead as position preservation through the presence of a mora. This move sidesteps the previous problems as no reference to the input moraicity of segments is required and consequently onsets, like codas, can cause CL. Moreover, concerns about OT's parallelism, single input-output mapping and Richness of the Base are taken into account ensuring that the resulting analysis observes all of them. The proposed solution neatly accounts for a number of CL cases and is successfully implemented in Samothraki Greek, an elaborate analysis of which is offered to capture the full range of facts.

Keywords: Compensatory lengthening, Samothraki Greek, position correspondence, Optimality Theory

1 Introduction

Compensatory lengthening (CL) refers to the process whereby the deletion of a segment is compensated for by the lengthening of a neighbouring segment, e.g. Latin *fideslia* > *fide:lia* 'pot'. In his seminal paper on CL, Hayes (1989) attributes this phenomenon to the fact that although a segment is lost, its timing unit - its mora - is not, but simply re-associates to an adjacent melodic unit. Thus, in our example *fideslia* above, the word initially syllabifies and receives moras. Then the coda *s* deletes, but its mora stays behind and subsequently re-links to the preceding vowel rendering it long bimoraic: /fideslia/ → [fi^u.de^us^u.li^ua^u] → [fi^u.de^u∅^u.li^ua^u] → [fi^u.de^{uu}.li^ua^u]. All in all, the total number of moras in the word remains constant.

Hayes (1989) moreover clearly states that no CL should occur after the deletion of an onset consonant. This is a direct consequence of moraic theory, which stipulates that onsets are never moraic. Consequently, their deletion cannot cause lengthening, i.e. re-association of their mora to a neighbouring segment, since there is no mora on an onset to start with. Nonetheless, Hayes considers two possible counter-examples to this generalization in his investigation of Samothraki Greek and Onondaga. Both involve *r*-deletion from an onset position and lengthening of the following vowel. Hayes proposes that both can be re-analysed differently so that they no longer present CL after onset-loss.

In this paper, I will re-examine Hayes' data of Samothraki Greek and offer additional facts which had gone unnoticed in his account. As we will see, consideration of the full range of data is imperative, as it proves that his account is flawed. I will show that Samothraki Greek is actually a genuine case of CL after onset loss

and will provide a full analysis. In addition, I will claim that the onset may not only be the *trigger* of CL - as in Samothraki Greek - but it can also be the *target* of CL, as data from Trique illustrate.

Apart from the empirical merits of this paper, certain technical proposals are also put forward. In particular, I attempt to provide a more general account that can handle cases of CL caused by the deletion of a consonant, while at the same time ensuring that it is compatible with the OT tenets of parallelism and Richness of the Base (Prince and Smolensky 1993). While CL can easily be accounted for in serial derivational frameworks, it is notoriously difficult to do so in an output-based theory like OT. The reason is that CL presents a case of counterbleeding opacity, whereby Weight-by-Position (or whatever constraint forces moraicity of the consonant) has to precede consonant deletion, so that lengthening occurs, i.e. /fideslia/ → [fi^μ.de^μs^μ.li^μa^μ] → [fi^μ.de^μ∅^μ.li^μa^μ] → [fi^μ.de^{μμ}.li^μa^μ] otherwise no lengthening obtains as the opposite ordering illustrates: /fideslia/ → [fi^μ.de^μ∅^μ.li^μa^μ] → Weight-by-Position is inapplicable → [fi^μ.de^μ.li^μa^μ].

Serial derivational models that admit intermediate stages and rule ordering face no difficulty with such cases, but classic OT, which in its original formulation only admits an input-output mapping without intermediate stages, has no adequate way to handle them. The proposed model shows that if we treat CL not as the preservation of a mora (cf. Hayes 1989), but as the preservation of a position through a mora, then a neat account of CL is produced while being faithful to the OT claims.

The paper is structured as follows: Section 2 considers the full range of the Samothraki Greek data and offers numerous examples. Section 3 explains why Hayes' (1989) influential analysis of CL fails to account for Samothraki Greek. Data like the one found in this Greek dialect considerably weaken his proposal and suggest that treating CL as mora preservation is flawed (§3.1). In fact, this analysis faces more general problems, since it proves impossible to translate into OT terms without clashing with some of the theory's basic tenets: Richness of the Base and parallelism (§3.2). An alternative is consequently in order. The present account attempts to resolve all these issues (§4) by arguing that CL is not about mora preservation but about position preservation through a mora (§4.1). The proposed model deals with CL without resorting to input moraicity and also offers good coverage on more difficult cases such as CL because of prenasalisation or glide formation (§4.2). Crucially, it can also account for Samothraki Greek CL, a case of CL which has so far resisted an adequate account. An elaborate analysis appears in Section 5.

Section 6 considers numerous alternative accounts of CL that practically cover the full array of approaches within OT as well as phonetically-based accounts such as Kavitskaya (2002). I claim that all of them exhibit certain flaws that render them non viable. Section 7 addresses some interesting remaining issues, such as: i) the asymmetry between coda and onset CL, whereby the first is increasingly common and the latter exceedingly rare (§7.1), ii) the existence of CL where onsets function not only as triggers of CL as in Samothraki Greek, but also as targets as in Trique (§7.2), and iii) the proper characterisation of CL with respect to segments or syllable constituents (§7.3). Section 8 offers a synopsis and some concluding remarks.

2 Samothraki Greek data

Samothraki Greek (SamG) is spoken in the island of Samothraki in North Eastern Greece. Like other Northern Greek dialects, it is typical in raising stressless *e*, *o* to *i*, *u* respectively, e.g. *péde* > *pédi* “five”, *potamós* > *putamós* “river” and losing underlying *i* and *u* (with some exceptions) *tiyáni* > *tyan* “frying pan”, *kufós* > *kfos*

“deaf”. As these processes are orthogonal to our main focus here, I will set them aside. Unfortunately, as is the case for many other dialects, SamG is in decline, spoken mainly by older generations. Presumably, due to social and educational circumstances as well as due to the increased accessibility to Standard Modern Greek through the media, younger generations have little or no knowledge of the dialect (Katsanis 1996: 27).

Unless indicated otherwise, the data below come from Katsanis (1996; henceforth K). Some extra data are due to additional sources and from personal communication with Maria Tsolaki, a native speaker of SamG and Marianna Ronga at the Aristotle University of Thessaloniki, who has knowledge of the data too.

Recall that the reason SamG is interesting is because it presents data with /r/-deletion only from an onset position with subsequent CL of the *following* vowel. However, /r/ in a coda position does not delete at all. Let us first consider the latter case¹.

(1) *Coda /r/ stays* [here and throughout all glosses are mine]

fanár	‘lantern’ (K: 48)	arpázu	‘I grab’ (K: 48)
figár	‘moon’ (K: 58)	karpós	‘seed’ (K: 48)

In most other instances, /r/ deletes. In the case of singleton onsets, /r/ deletion leads to vowel lengthening, but only word-initially (2). Word-medially no such lengthening occurs (3)².

(2) *Deletion of /r/ initially and lengthening* (K: 50-51)

ra > a:	rafts > á:fts	‘tailor (masc.)’
ri > i:	riɣaŋ > í:ɣaŋ	‘oregano’
ru > u:	rúxa > ú:xa	‘clothes’
re > e:	réma > é:ma	‘stream’
ro > o:	róɣa > ó:ɣa	‘nipple, berry (of a grape)’

(3) *Deletion of /r/ word-medially and no lengthening* (K: 52)³

aro > ao	θaró > θaó	‘I reckon’
iru > iu	léftirus > léftius	‘free’
are > ae	varéɻ > vaéɻ	‘barrel’
iri > ii	θiriða > θiða	‘pigeon-hole’
eru > eu	kséru > kséu	‘I know’
uri > ui	lurí > luí	‘strap, strip’
era > ia	méra > mía	‘day’
ara > aa	skára > skáa	‘grill’
ure > ue	kurévo > kuévu	‘I cut someone’s hair’

In complex onset clusters of the type *CrVC*, /r/ again deletes and lengthening occurs independently of the cluster’s position within the word.

(4) */r/ in onset CrVC: deletion and lengthening*

a) *biconsonantal clusters* (K: 54-56)

pr+o > po:	prótos > pó:tus	‘first’
vr+i > vi:	vrisi > ví:s’	‘tap’
fr+e > fe:	fréna > fé:na	‘brakes’
xr+o > xo:	xróma > xó:ma	‘colour’
kr+a > ka:	krató > ka:tó	‘I hold’

$\gamma r+a > \gamma a:$	$\gamma r\acute{a}f\acute{o} > \gamma \acute{a}:f\acute{u}$	‘I write’
$\theta r+o > \theta o:$	$\theta r\acute{o}n\acute{o}s > \theta \acute{o}:n\acute{u}s$	‘throne’
$\delta r+a > \delta a:$	$\delta r\acute{a}k\acute{o}s > \delta \acute{a}:k\acute{u}s$	‘dragon’
$br+e > be:$	$\gamma abr\acute{e} > \gamma ab\acute{e}:$	‘bridegroom’
$dr+u > du:$	$\delta \acute{e}d r\acute{o} > \delta \acute{e}d\acute{u}:$	‘tree’
$tr+u > tu:$	$metr\acute{u}n > mit\acute{u}:n$	‘they count’

b) *triconsonantal clusters* (K: 58-59)

$spr+a > spa:$	$\acute{a}spra > \acute{a}spa:$	‘white’
$xtr+a > xta:$	$\acute{e}xtra > \acute{e}xta:$	‘hostility’
$zdr+u > zdu:$	$sidr\acute{u}f\acute{c}\acute{a} > zdu:f\acute{c}\acute{a}$	‘company (of people)’
$ftr+a > fta:$	$r\acute{a}ftra > \acute{a}:fta:$	‘tailor (fem.)’

In contrast to these data, r-deletion does not happen in complex clusters of the type $CrVV$. More specifically, in all cases known to me, the first vowel in the sequence is a front vowel /i/ or /e/, thus this string is more accurately described as $Cr+i/e+V$. The output of such sequence is not $C + i:/e: + V$ as anticipated given the previous examples, rather it is $C+i+rjV$ without r-deletion or lengthening, but with metathesis and glide formation (cf. (5)). I have been unable to find any data where the first vowel is back, since standard Greek words like *fráula*, *krúo*, *práos*, *akróasi* are not used in the dialect (Ronga p.c.). It is thus impossible at present to test what happens in the sequences $C+r+a/o/u+V$. Despite the absence of empirical confirmation, we actually do not expect to get metathesis and glide formation for independent reasons. Given that the vowels involved are back, a back glide *w* would be required, but this is missing from the dialect, as it is from Standard Modern Greek too. With this caveat in mind, the ensuing analysis is designed to only account for $C+r+i/e+V$ sequences.

- (5) $C+r+i/e+V \rightarrow C+i+rjV$ Ronga (p.c.)
- | | |
|--|--------------------------------------|
| $pr\acute{a}k\acute{o}n\acute{i} > pirjak\acute{o}n$ | ‘jagged file used to sharpen knives’ |
| $al\acute{e}tr\acute{i}a > al\acute{e}tirja$ | ‘plough (plural)’ |
| $tr\acute{i}a > tirj\acute{a}$ | ‘three’ |

Katsanis describes exactly the same phenomenon but with reference to velars only, i.e. $C_{[Velar]}+r+i/e+V \rightarrow C_{[Velar]}+\hat{i}+rjV$. An additional process applies here, namely centralization of front vowels *i/e* to \hat{i} (or \hat{e} occasionally)⁴.

- (6) $C_{[Velar]} + r + i/e + V \rightarrow Velar + \hat{i} + rjV$ (K: 71)
- | | | | |
|--|--------|--|-------------|
| $\acute{a}\gamma r\acute{i}os > \acute{a}\gamma \hat{i}rj\acute{u}s$ | ‘wild’ | $axr\acute{i}ast\acute{o}s > ax\hat{i}rjast\acute{u}s$ | ‘unneeded’ |
| $kr\acute{e}as > k\hat{i}rj\acute{a}s$ | ‘meat’ | $\gamma r\acute{i}\acute{a} > \gamma \hat{i}rj\acute{a}$ | ‘old woman’ |

Centralization of front vowels in the environment $k/g/x/\gamma+r+i/e$ generally happens, i.e. even when a consonant follows (compare (7) with (4))⁵.

- (7) $C_{[Velar]} + r + i/e + C \rightarrow Velar + \hat{i}:/\hat{e}: + C$
- | | | | |
|--------------------|---|-------------|---------|
| $gr+i > g\hat{i}:$ | $gr\acute{i}z\acute{o}s > g\hat{i}:z\acute{u}s$ | ‘gray’ | (K: 56) |
| $kr+i > k\hat{i}:$ | $kr\acute{i}n\acute{u} > k\hat{i}:n\acute{u}$ | ‘I judge’ | (K: 56) |
| $kr+e > k\hat{e}:$ | $kremn\acute{o}s > k\hat{e}:mn\acute{u}s$ | ‘precipice’ | (K: 72) |

It should thus be evident that while vowel centralization is clearly related to velar consonants, the absence of /r/-deletion and the emergence of a glide in (5) and (6) is independent of the quality of the consonant; it relates to the presence of a $V_{\text{front}}+V$. In what follows, I will abstract away from centralization and focus on the /r/ loss and lengthening facts.

In a nutshell, deletion of /r/ from an onset position in SamG is almost always compensated for by the lengthening of the following vowel. This effect can be straightforwardly accounted for if we allow onsets to cause CL contra Hayes (1989). Any restrictions to the application of the process, as in e.g. word-medial position, will relate to other, more general principles of the language. Our first task however is to show why Hayes' account is not viable.

3 Problems for CL as “mora preservation” in SamG and in its OT implementation

3.1 Why Hayes (1989) fails when applied to SamG

As noted, Hayes' model attributes CL to mora preservation and stipulates that onsets never carry moras. Consequently, CL from onset loss is prohibited. The data in SamG however provide a potential counterexample to such a view, which is why Hayes is forced to offer a different account for them. More concretely, his analysis focuses on inputs with /CrV/ sequences which surface as [CV:]. First, he claims that the cluster is split by V-epenthesis. The result of this is [CVrV]. Subsequently, intervocalic /r/ loss applies and merger of [VV] to [V:] follows. Thus, lengthening is merely an epiphenomenon of vowel merging, rather than the direct result of CL after onset deletion.

Hayes bases this proposal on a suggestion made by Newton (1972), who claims that a similar process occurs in the development of *xartí* to *xaití* “paper” or *karpós* to *kaipós* “seed” whereby /i/ is an epenthetic vowel. Such cases would thus be analysed as [VrC] → [VriC] → [ViC]. While Katsanis (1996: 48) mentions that these examples are reported in Heisenberg (1921), whose credibility is unquestionable, he nonetheless claims that such forms are not accepted by modern speakers. Katsanis eventually accepts Newton's interpretation of the facts, but remains somewhat sceptic regarding its validity given that /i/ epenthesis is common in Imvros or Lesbos but does not emerge in Samothraki, as Newton himself observes with the exception of two examples, i.e. *tun-i-mikró* (MG: ton mikró) “the little one” and *babázim* (MG: babáz mu) “my dad” (K: 51)⁶. Apart from the complex onset data, Hayes also investigates the loss of word-medial /r/ and suggests that under his proposal, the absence of lengthening there is anticipated since there is no motivation for vowel epenthesis in a VrV sequence.

Despite the ambiguity of facts, Hayes' proposal is thus far tenable. What proves fatal for it though is the consideration of the word-initial data, which are totally absent from his discussion. Recall that word-initially, deletion of singleton /r/ is actually followed by vowel lengthening. In analogy to the complex onset proposal, one would expect vowel epenthesis as a possible solution here, but no motivation exists for V-insertion in this case, since no cluster is involved. In fact, there should not be any reason for /r/-deletion either here, since in Hayes' view, /r/-deletion only occurs intervocalically. Consequently, any attempt to diverge the interest to an alleged epenthetic vowel rather than the onset itself is futile. Reference to onsets is imperative given the behaviour of /r/ word-initially.

3.2 Why an OT analysis of CL as “mora preservation” fails

Apart from the specific problems that SamG CL poses for Hayes (1989), the “mora preservation” model faces general problems too. In particular, it proves impossible to translate into OT. To see why this is the case, let us take a hypothetical example of CL such as /kan/ → [ka:], where the coda deletes rendering the previous vowel long. If CL is all about mora preservation and given that OT involves just a single input-output mapping, we are forced to assume that the segment which eventually gets deleted is in the first place moraic, so that due to MAX-μ, its mora is preserved and realized on another segment. The tableau below illustrates.

(8) CL in OT --- moraic input⁷

	/ka ^μ n ^μ /	NOCODA	MAX-μ	MAX-SEG	DEP-μ
a.	ka ^μ		*!	*	
b.	ka ^{μμ}			*	

(8b) with a long vowel wins over (8a) with a short vowel, as it manages to preserve all input moras in the output too, whereas (8a) loses one mora. However, in order that we conform to Richness of the Base (ROTB), which states that no restrictions should be placed on inputs, we are also obliged to consider the same input, but without any moraic specification⁸.

(9) CL in OT --- non-moraic input

	/ka ^μ n/	NOCODA	MAX-μ	MAX-SEG	DEP-μ
a.	ka ^μ			*	
b.	ka ^{μμ}			*	*!

This time, MAX-μ is no longer applicable since both outputs retain all input moras. MAX-SEG is equally violated by both candidates, therefore the decision is passed onto DEP-μ. This chooses the candidate that lacks lengthening (9a), because it avoids mora insertion. As a matter of fact, given that the violations of (9a) are a subset of those incurred by (9b), (9a) harmonically bounds (9b), which means that any ranking of the constraints in (9) would favour the non-lengthened output.

Consequently, we run into a dilemma where both options are equally unappealing; we either need to impose input moraicity, which however violates ROTB (8), or we conform to ROTB, but then half of the time, we produce the incorrect results (9). Notably, serialist frameworks such as Hayes (1989) do not face this problem, because they permit intermediate stages whereby syllabification and moraification can happen prior to CL no matter what the moraicity of the input is, i.e. /kan/ → [ka^μn^μ] → [ka^μ∅^μ] → [ka^{μμ}].

The moral of the story is that due to the lack of intermediate stages, no OT analysis should build an account of CL as mora preservation, because the latter presupposes the existence of moras in the input (which is the only place where moras can be deposited in input-output OT), contra ROTB.

The proposal I will offer next bypasses such problems, as it works independently of the input moras assumed. It can also deal with the SamG facts for the same reason; given that it does not rely on input moras and their preservation, it does not need to make any explicit reference to *moraic* onsets which can consequently cause CL⁹ (unlike Topintzi 2005a¹⁰). Crucially though, an OT analysis of CL should not hinge on that point and

indeed the present approach does not, as it wishes to remain consistent with the OT tenets of parallelism and ROTB by avoiding reference to input moras or utilisation of intermediate stages. The next section outlines the present account.

4 The alternative: CL as “position preservation through a mora”

4.1 Outline of the proposal

The basic idea I will pursue is that CL is not about mora preservation, but about position preservation via a mora. An input segment needs to have an output correspondent. There are two ways it can achieve this. It may either survive in the output intact (or with some of its identity features altered) or it may delete, in which case CL occurs via the addition of a mora. CL is then just seen as an alternative strategy for a segment to show up in the output with lengthening acting as a ‘cue’ for the lost segment. This can be coded in the constraint termed POSITION CORRESPONDENCE (POSCORR).

- (10) POSCORR: An input segment must have an output correspondent either segmentally by means of a root node or prosodically by means of a mora

To illustrate, let us consider the hypothetical example utilized above. The root /kan/ may either surface on its own or be followed by vowel-initial (-a) or consonant-initial (-ta) suffixes.

- (11) a. /kan/ → [ka:]
b. /kan-ta/ → [ka:ta]
c. /kan-a/ → [kana]

(11) is a typical case of synchronic CL. (11a-b) show that when the root-final consonant appears in syllable final position, it deletes. On this occasion, CL occurs, producing lengthening of the preceding vowel. In (11c), no deletion occurs, since the consonant /n/ is no longer in syllable-final position. At the same time, no lengthening occurs either. How are we to account for these facts in a parallelistic account? The next tableaux will shed some light on this issue.

To facilitate their reading, I will be using indices to identify corresponding positions. To avoid too many indices, moras of input segments will have no indices. Only the inserted mora [in bold] due to CL will have an index to mark the mapping to the deleted segment.

(12) /kan/ → [ka:]

	/k ₁ a ₂ ^μ n ₃ /	NOCODA	POSCORR	DEP-μ
a.	..a ₂ ^μ n ₃	*!		
b.	..a ₂ ^μ n ₃ ^μ	*!		*
c.	..a ₂ ^μ		*!	
d.	..a ₂ ^{μμ3}			*

The first two candidates retain the coda consonant and thus violate NOCODA, even though the second also assigns a mora to this consonant. The third candidate deletes the coda, but fails to present any correspondent of the position held by /n/, either segmental or prosodic. As a result, POSCORR is violated. The winning candidate is therefore (d), because it manages to simultaneously satisfy NOCODA but also retain a correspondent of the position occupied by /n/ by lengthening the previous vowel (indicated by the added μ and the index 3 which identifies it with the deleted consonant). All this happens at the cost of DEP-μ only.

Of course, there is one more important candidate we should have considered, namely ka_{2,3}^μ. This has no lengthening, but presents coalescence of the deleted segment with the previous vowel. While this does not seem very likely, given that in such case we should also probably get nasalization of the vowel, we should consider this competitor too. By ranking the constraint against coalescence (UNIFORMITY) [McCarthy and Prince 1995] highly enough, this candidate will be easily excluded as in (14a)¹¹.

(13) UNIFORMITY: No element of the output has more than one correspondent in the input, i.e. no coalescence

(14) /kan/ → [ka:]

	/k ₁ a ₂ ^μ n ₃ /	UNIFORMITY	POSCORR	DEP-μ
a.	..a _{2,3} ^μ	*!		
b.	..a ₂ ^{μμ3}			*

This same analysis should work for /kan-ta/ → [kata]. Now, let us see whether this ranking also works in /kan-a/ → [kana].

(15) /kan-a/ → [kana]¹²

	/k ₁ a ₂ ^μ n ₃ -a ₄ ^μ /	UNIFORMITY	NoCODA	POSCORR	DEP-μ
a.	..a ₂ ^μ n ₃ a ₄ ^μ				
b.	..a ₂ ^{μμ3} n ₃ a ₄ ^μ				*!
c.	..a ₂ ^μ a ₄ ^μ			*!	

This time there's no trigger for /n/-deletion (15c), so there's no reason why the input segment /n₃/ should not have a segmental correspondent (15a). A prosodic correspondent by means of an inserted mora would satisfy POSCORR but only superfluously at the expense of DEP-μ (15b). Thus, we can see that the right results obtain. More generally, the pattern for CL involves the following:

- (16) *Pattern for CL* --- preliminary version:
- a trigger for deletion of segments in particular syllabic positions must exist (markedness constraint M)
 - POSCORR >> DEP- μ ensures that CL occurs, i.e. insertion of a mora identified with the position of the deleted input segment
 - UNIFORMITY >> DEP- μ : prosodic rather than segmental identification is preferred

Needless to say that the same results would also obtain had the input contained an underlyingly moraic consonant. The only difference would be that the DEP- μ violations would no longer apply given that the mora identified with the lost segment would already be present in the input. This system can also generate the absence of CL, as shown in (17). A hypothetical example of this sort also appears in /kan/ → [ka] illustrated in (18).

- (17) *Deletion of a segment not followed by CL* --- preliminary version:
- requires minimal re-ranking: DEP- μ >> POSCORR implying that no trace (segmental or prosodic) of the deleted segment is left behind

(18) *Lack of CL: /kan/ → [ka]*

	/k ₁ a ₂ ^μ n ₃ /	UNIFORMITY	NOCODA	DEP- μ	POSCORR
a.	..a ₂ ^μ n ₃		*!		
b.	..a ₂ ^μ n ₃ ^μ		*!	*	
c.	..a ₂ ^μ				*
d.	..a ₂ ^μ n ₃ ³			*!	
e.	..a _{2,3} ^μ	*!			

This is a good result, because obvious alternative analyses of CL face several problems. In an analysis à la Hayes, serial derivation has to be utilized under which WBYP acts before the coda consonant is deleted so that this receives a mora. After coda deletion, lengthening occurs. The absence of CL is attributed to the inactivity of WBYP or to its application after coda deletion. This is only possible however if the process is seen derivationally. The other possibility which would avoid this issue involves underlying moraic specification, but this is problematic too as we had seen in §3.2.

Note however that (c) is not the only conceivable winner in (18). As hinted in (16), high-ranking POSCORR dictates the need for a position correspondent, but due to the ranking UNIFORMITY >> DEP- μ , the prosodic option through a mora is chosen. Thus, CL occurs. But there is another possibility. If all else stays the same, but there is a minimal re-ranking so that DEP- μ >> UNIFORMITY, there will still be a position correspondent, but this time it will take the form of a segmental one. This ranking would select (18/19e) instead as illustrated in the tableau in (19). In effect, this is another way to produce the lack of CL.

(19) *Lack of CL: /kan/ → [ka]*

	/k ₁ a ₂ ^μ n ₃ /	NOCODA	POSCORR	DEP-μ	UNIFORMITY
a.	..a ₂ ^μ n ₃	*!			
b.	..a ₂ ^μ n ₃ ^μ	*!		*	
c.	..a ₂ ^μ		*!		
d.	..a ₂ ^{μμ3}			*!	
e.	..a _{2,3} ^μ				*

In most instances where CL fails to occur, it is difficult to decide between (18c) and (19e), so in principle, both options are available. There is however a reason why I am inclined to choose (18c) over (19e) whenever data do not suggest otherwise. (19e) involves fusion, which entails that it should be possible to recognise whether the segment is the product of coalescence. For instance, in the case of /kan/ becoming [ka], we could argue that fusion is involved had the resulting vowel shown nasalization, i.e. as in [kã]. This is indeed the case in French, as shown below. Consonant deletion is followed by nasalization, which suggests that the winning candidate in French would be the product of the ranking in (19).

(20) *French: [Y] ~ [Vn] alternations in masculine-feminine (Tranel 1987: 70-71)¹³*

bon [bɔ̃]	vs.	bonne [bɔ̃n]	‘good’
fin [fɛ̃]	vs.	fine [fin]	‘thin’
certain [sertɛ̃]	vs.	certaine [sertɛ̃n]	‘sure’

A similar example is found in Portuguese (Mateus and d’Andrade 2002), where it is argued that nasalized vowels are the product of an underlying sequence of a vowel and a nasal. For instance, the prefix /in-/ appears as [in-] before a vowel, but as a nasalized vowel [ĩ] before a consonant. In the latter case, presumably nasal consonant deletion has occurred, with the nasal feature being left behind and linked to the preceding vowel.

(21) *Portuguese nasalized vowels (Mateus and d’Andrade 2002: 22)*

a.	V-initial stems			
	<i>acabado</i>	‘finished’	<i>inacabado</i>	[inɛkɐbádu] ‘unfinished’
	<i>oportuno</i>	‘opportune’	<i>inoporuno</i>	[inopurtúnu] ‘inopportune’
b.	C-initial ítems			
	<i>capaz</i>	‘able’	<i>incapaz</i>	[ĩkɐpás] ‘unable’
	<i>posto</i>	‘put in place’	<i>imposto</i>	[ĩpóʃtu] ‘tax’

This kind of evidence however is not always available. Consequently, whenever such direct evidence is lacking, I will assume that the language prefers to delete the position altogether leaving no relic of it behind, i.e. as in (18c).

4.2 Fine-tuning of the proposal

Recall that an OT analysis of CL à la Hayes (1989) faces the difficulty that production of CL is not consistent. More concretely, if the segment-to-be-deleted bears an input mora, e.g. /CV^μC^μ/ then CL is generated as a response to MAX-μ satisfaction. On the other hand, if the input lacks this mora, e.g. /CV^μC/, MAX-μ is no longer

an issue and in fact DEP- μ is now going to penalise CL. The proposed analysis solves this problem, by replacing MAX- μ with POSCORR which is interested in retaining input *positions* in the output too. This can either occur segmentally (by retaining the segment at hand) or prosodically (by identifying a mora with the position held by the now lost segment). The core of the analysis then involves:

- (22) *Deletion of segment followed by CL: POSCORR >> DEP- μ*
Deletion of segment not followed by CL: DEP- μ >> POSCORR

This now raises two issues. First, we have established that CL is generated independently of the moraicity of the input assumed, but is the absence of CL also ensured no matter what the input is? I will show, that as DEP- μ stands, the absence of CL is not guaranteed, therefore some adjustment to the understanding of DEP- μ is in order. Second, we have shown that we can produce CL after the *deletion* of a segment, but can we also generate it when the segment is not fully lost, as is the case, for instance, in Bantu prenasalisation? Again, minimal modification ensures that such cases can also be accounted for.

4.2.1 Lack of CL independently of the assumed input

First let us tackle the lack of CL. To illustrate the problem, I will apply the two rankings in (22) on the moraic /CV ^{μ} C ^{μ} / and the non-moraic /CV ^{μ} C/ inputs and examine the predicted outputs.

- (23) *Deletion of segment followed by CL: POSCORR >> DEP- μ*

(i) non-moraic input

	/C ₁ V ₂ ^{μ} C ₃ /	POSCORR	DEP- μ
a.	V ₂ ^{μ2}	*!	
b.	V ₂ ^{μ3}		*

(ii) moraic input

	/C ₁ V ₂ ^{μ} C ₃ ^{μ} /	POSCORR	DEP- μ
a.	V ₂ ^{μ2}	*!	
b.	V ₂ ^{μ3}		

CL is correctly produced independently of the input assumed, as shown in (23). The same does not hold though when we try to produce the lack of CL as (24) exemplifies.

- (24) *Deletion of segment not followed by CL: DEP- μ >> POSCORR*

(i) non-moraic input

	/C ₁ V ₂ ^{μ} C ₃ /	DEP- μ	POSCORR
a.	V ₂ ^{μ2}		*
b.	V ₂ ^{μ3}	*!	

(ii) moraic input

	/C ₁ V ₂ ^{μ} C ₃ ^{μ} /	DEP- μ	POSCORR
a.	V ₂ ^{μ2}		*!
b.	V ₂ ^{μ3}		

When the input is non-moraic (24i), lack of CL is correctly generated, but when the input is moraic, CL is instead - wrongly - produced as in (24ii.b), because no mora has been inserted and thus DEP- μ no longer distinguishes between the lengthened and the non-lengthened forms. The decision is now passed on to POSCORR which eliminates the non-lengthened form as it fails to provide a position correspondent for position 3. This is now the mirror-image of the problem Hayes' CL account faced in OT. There, generation of CL proved inconsistent; here it is the lack thereof which does.

To address this discrepancy, we only need to somehow amend DEP- μ so that it is also violated in (24ii). In other words, we need a version of DEP- μ that militates against lengthening (but of course not against underlying length). Such version is already available, as suggested in Bermudez-Otero (2001) and more fully worked out in Campos-Astorkiza (2004). The idea bears on the notion of Positional μ -licensing (Bermudez-Otero 2001).

- (25) **Positional μ -licensing:** A segment α is positionally μ -licensed by a mora iff μ is the only prosodic unit directly dominating α .

A positional μ -licenser then is a mora which is the only prosodic unit that dominates a given segment. A non-positional μ -licenser on the other hand is a μ which is not the only prosodic unit immediately dominating α . Now, if positional μ -licensing is applied to the definition of DEP- μ , then the updated P(OSITIONAL)-DEP- μ will penalise (24b) as desired.

- (26) **P-DEP- μ :** A non-positional μ -licenser mora in the output has a correspondent in the input

This constraint will only penalise inserted moras that lead to lengthening. Its evaluation and a comparison with classic DEP- μ is in (27). Only relevant structures are considered.

(27)

	Input	Output	Positional- μ -licenser	P-DEP- μ	DEP- μ	
a.	$\mu_1 \mu_2$ $\swarrow \searrow$ V	$\mu_1 \mu_2$ $\swarrow \searrow$ V	No	✓	✓	(a) long vowel
b.	μ_1 V	$\mu_1 \mu_2$ $\swarrow \searrow$ V	No	*	*	(b) lengthened vowel
c.	μ_1 V C	$\mu_1 \mu_2$ $\swarrow \searrow$ V	No	*	*	(c) non moraic C-deletion and lengthening
d.	$\mu_1 \mu_2$ V C	$\mu_1 \mu_2$ $\swarrow \searrow$ V	No	*	✓	(d) moraic C-deletion and lengthening

(27a) presents an underlyingly long vowel. Although one of the moras in the output is not a positional μ -licenser (since there is not just one μ directly dominating the segment V), P-DEP- μ is not violated, because both moras have a correspondent in the input. So (a) escapes a violation of P-DEP- μ . Naturally, (a) satisfies DEP- μ since no

mora has been added. This is not the case in (27b) where μ_2 has been inserted in the output and consequently DEP- μ is violated. P-DEP- μ is also violated, because this time the non-positionally μ -licensed μ_2 has no correspondent in the input. (c) now presents a case where in a non-moraic $/V^{\mu}C/$ sequence, the coda deletes and lengthening occurs. This is reminiscent of the situation in (b) and hence both constraints agree in their evaluation. Up to this point then, both DEP- μ constraints yield the same results. The difference arises when the $/V^{\mu}C^{\mu}/$ sequence includes a moraic C (d). By DEP- μ no violation obtains in $[V^{\mu\mu}]$, because no mora has been inserted. But for P-DEP- μ this is beside the point. Vowel lengthening again occurs and given that the non-positionally μ -licensed μ_2 linked to V has no correspondent linked to the same segment in the input, a P-DEP- μ violation obtains¹⁴.

This difference will give us the right results in (23) and (24) which are now updated as follows with P-DEP- μ (and re-numbered as (28) and (29) respectively):

(28) *Deletion of segment followed by CL: POSCORR >> P-DEP- μ*

(i) non-moraic input

	$/C_1V_2^{\mu}C_3/$	POSCORR	P-DEP- μ
a.	$V_2^{\mu 2}$	*!	
b.	$V_2^{\mu 2}\mu^3$		*

(ii) moraic input

	$/C_1V_2^{\mu}C_3^{\mu}/$	POSCORR	P-DEP- μ
a.	$V_2^{\mu 2}$	*!	
b.	$V_2^{\mu 2}\mu^3$		*

(29) *Deletion of segment not followed by CL: P-DEP- μ >> POSCORR*

(i) non-moraic input

	$/C_1V_2^{\mu}C_3/$	P-DEP- μ	POSCORR
a.	$V_2^{\mu 2}$		*
b.	$V_2^{\mu 2}\mu^3$	*!	

(ii) moraic input

	$/C_1V_2^{\mu}C_3^{\mu}/$	P-DEP- μ	POSCORR
a.	$V_2^{\mu 2}$		*
b.	$V_2^{\mu 2}\mu^3$	*!	

With this modification, our rankings of CL and lack thereof are shaped as follows¹⁵:

(30) *Deletion of a segment followed by CL --- final version:*

M, UNIFORMITY, POSCORR >> P-DEP- μ

(31) *Deletion of a segment **not** followed by CL --- final version:*

i) no trace of deleted position: M, UNIFORMITY, P-DEP- μ >> POSCORR (cf. (18))

ii) segmental correspondent: M, POSCORR, P-DEP- μ >> UNIFORMITY (cf. (19))

4.2.2 CL without total loss of a segment

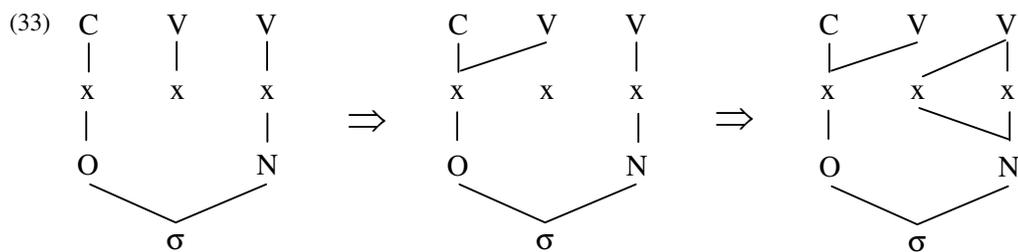
A patent question comes to mind. How about cases of CL where no segment is deleted? The obvious examples involve Bantu prenasalization and glide formation. In the former, Luganda inputs like /ba-ntu/ surface as [ba:ⁿtu] ‘people’, while in the latter /mu-iko/ becomes [m^wi:ko] ‘trowel’. No segment is lost here so why should CL occur in the first place? The answer lies in the definition of POSCORR repeated below.

- (32) POSCORR: An input segment must have an output correspondent either segmentally by means of a *root node* or prosodically by means of a mora

Work by Clements (1986), Sagey (1986) and Rosenthal (1994) suggests that the nasal or the high vowel in the sequences mentioned above has no correspondent in terms of a root node. In the output, the complex segment consists of a single root node which contains the nasal features or the ones of the high vowel. But since there is no segmental correspondent and assuming that POSCORR >> P-DEP-μ holds, some correspondent is required. This can only take the form of a prosodic correspondent by means of lengthening of the neighbouring vowel.

Kinyarwanda is a useful example for this claim (Sagey 1986). This language admits sequences of a consonant followed by a glide. Compelling evidence that this is a single complex segment with multiple articulations that consists of just one root node comes from the fact that Kinyarwanda does not allow complex onsets. More specifically, German loanwords that involve complex onsets are not preserved intact in Kinyarwanda (as predicted had the language allowed complex onsets), but are split by epenthesis so that simplex onsets are created, e.g. *Republik* > *repuburika*, *Präsident* > *perezida*, etc. Complex segments are allowed however, therefore they must be a single root node.

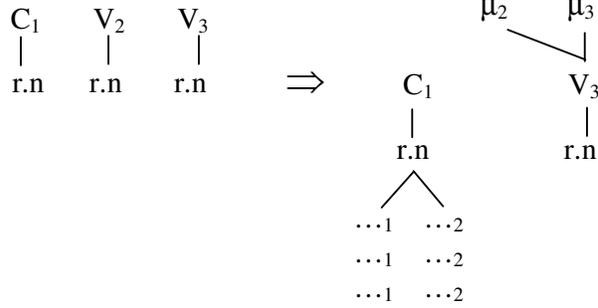
While I am not concerned with the specific features or the geometry involved in complex segments of this sort (see e.g. Sagey 1986, Lombardi 1990), the point which is of importance for us is that these segments are simply one root node. For instance, in Sagey (1986: 75), the representation of glide formation and subsequent CL is in (33) and corresponds to Kinyarwanda examples like the ones in (34):



- (34) /ku-i-Bon-a/ → [k^wi:Bona]¹⁶ ‘to see oneself’ /ku-gu-ir-a/ → [kug^wi:ra] ‘to fall on’

The first vowel creates a complex segment with the preceding consonant under a single timing slot (or root node). The first vowel’s timing slot re-associates with the second vowel leading to CL. We can translate these representations to the current model as in (35):

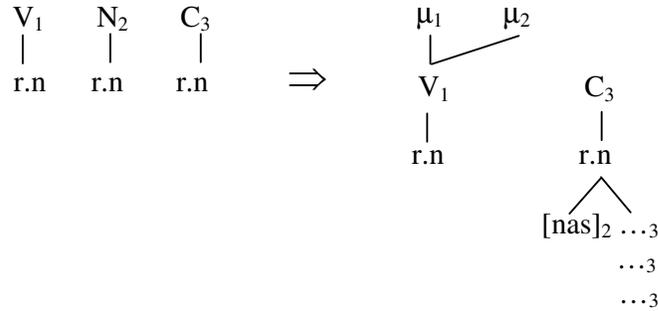
(35) *Glide formation and CL*



An input sequence such as $/C_1V_2V_3/$ then becomes $[C_1V_3^{\mu_2\mu_3}]$. The features of V_2 are included in the complex segment C_1 which consists of one root node. V_3 on the other hand lengthens so that it provides a prosodic position correspondent for V_2 , namely μ_2 . An identical process holds for examples that involve prenasalization.

- (36) $/ba\text{-}nde/ \rightarrow [ba:\text{nd}e]$ 'who?'
 $/ku\text{-}ngana/ \rightarrow [ku:\text{ngana}]$ 'to be equal'

(37) *Prenasalization and CL*



An example of these facts using an OT tableau is shown immediately below¹⁷.

(38) $/ba\text{-}nde/ \rightarrow [ba:\text{nd}e]$

	$/b_1a_2^{\mu}n_3d_4e_5^{\mu}/$	UNIFORMITY	PRENASALIZATION	POSCORR	P-DEP- μ
a.	$b_1a_2^{\mu}.n_3d_4e_5^{\mu}$	*!			
b.	$b_1a_2^{\mu}.n_3d_4e_5^{\mu}$		*!		
c.	$b_1a_2^{\mu}.n_3^{\mu}d_4e_5^{\mu}$			*!	
d.	$b_1a_2^{\mu_2\mu_3}.n_3^{\mu}d_4e_5^{\mu}$				*

Candidate (b) fails to emerge with prenasalization, so it violates whichever constraint is responsible for this violation (informally called here PRENASALIZATION). The other candidates all emerge with prenasalization. The first one presents fusion, thus it violates UNIFORMITY. Among the two remaining rivals, (c) loses since its representation implies that the nasal consonant appears with its own root node. We have argued that there is empirical evidence suggesting that this is an incorrect representation of prenasalized stops, which consist of a single root node. Thus, in (c) there is no position correspondent for n_3 , therefore such form fails POSCORR. (d)

wins since its only violation is that of P-DEP- μ due to the addition of a mora as an output prosodic correspondent for /n/.

5 Samothraki Greek compensatory lengthening analysis

With this much machinery at hand, we can now focus on SamG CL and its analysis. The core of the analysis will utilize the ranking which generates CL as in (30), namely: M, UNIFORMITY, POSCORR \gg P-DEP- μ . We will see that the markedness constraint triggering onset deletion in SamG is *ONSET/r, which is attributed to the proposed placelessness of SamG /r/ (§5.1). The absence of CL word-medially relates to the cross-linguistic tendency of avoiding super-long vocalic transitions (Kavitskaya 2002), which would have been generated had CL applied in that position (§5.5). Word-initially (§5.4) or in complex onsets followed by a V+C sequence (§5.6), no risk of that sort appears which is why CL can apply freely. CL however does not apply in two more cases: a) in complex clusters followed by V+V sequences, as it would create undesirable structures (§5.7) and b) in coda position, *ONSET/r becomes irrelevant, so no trigger for /r/ deletion exists and hence no reason to lengthen the preceding vowel (§5.3). The following sections present the detailed analysis.

5.1 On the placelessness of SamG /r/

One assumption that is going to prove crucial for the subsequent discussion is the claim that the /r/ in SamG is placeless. Such placelessness will serve to explain not only why /r/ deletes from onset position in the first place, but will also provide the basis for a specific case of vowel spreading (cf. §5.7).

The placelessness of /r/ receives support from other languages where a similar claim has also been put forward such as Yoruba (Akinlabi 1993) or English (Rice 1992). For instance, Rice observes that onset clusters *tl *dl *pw and *bw are impossible in English. Using evidence from similar cases like English, she claims that (onset) consonants cannot be identical in terms of Place structure. But then this generates a problem; /tr/ and /dr/ onset sequences are allowed. If both /r/ and /l/ are coronal, then why are r-coronal clusters accepted but l-coronal clusters are not? A possible answer is that /r/ is placeless. As a result, coronal+r sequences are admitted since they differ in terms of place of articulation.

The implication of /r/'s placelessness in Samothraki Greek is that /r/ will be able to survive in codas where placelessness is often accepted or required [cf. Selayarese ? (Rice 1992), Japanese ŋ (Yip 1991)], but not in onsets, because placeful onsets are preferred. This point is captured by the constraint *ONSET/r which is obviously dominant in SamG.

(39) *ONSET/r: /r/ is disallowed in onset position

*ONSET/r is thus the M(arkedness) constraint necessary in the schema of CL. For SamG then the CL ranking will be as in (40). A few more constraints will be added in due course to capture the full range of facts.

(40) Core ranking for SamG CL

*ONSET/r, UNIFORMITY, POSCORR >> P-DEP-μ

5.2 Summary of the patterns

Before moving on to the specifics of the analysis, let us remind ourselves of the SamG patterns along with some representative examples as summarized in the following table.

(41) Summary of SamG /r/ patterns

		Deletion	Lengthening	Examples
Coda /r/		NO	NO	fanár, karpós, arpázu (cf. (1))
Singleton onset /r/	word-initially	YES	YES	ú:xa, á:fts, é:ma (cf. (2))
	word-medially	YES	NO	kséu, luí, skáa (cf. (3))
Complex onset /r/ ¹⁸	Cr+V+C	YES	YES	xó:ma, ka:tó, dé:du (cf. (4))
	Cr+V _[i,e] +V	NO	NO	alétirja, áyírjus (cf. (5-6))

5.3 Coda /r/

Let us first start with the easiest case. In codas, /r/ does not delete, thus any lengthening is gratuitous (42b). While no ranking argument can be provided by the coda data, by looking ahead to the onset analysis, the ranking UNIFORMITY, *ONSET/r, POSCORR >> P-DEP-μ is proposed. The winner is (42a) that presents no deletion, given that /r/ escapes a *ONSET/r violation by virtue of its coda syllabification¹⁹.

(42) no CL when /r/ in coda: /karpos/ → [karpós]

UNIFORMITY, *ONSET/r, POSCORR >> P-DEP-μ

	/k ₁ a ₂ ^μ r ₃ p ₄ o ₅ ^μ s ₆ /	UNIFORMITY	*ONSET/r	POSCORR	P-DEP-μ
a.	..a ₂ ^μ r ₃ ..				
b.	..a ₂ ^{μμ3} ..				*!

5.4 Singleton /r/ word-initially

Word-initially, /r/ deletion is forced by *ONSET/r, which now becomes active. The ranking UNIFORMITY, *ONSET/r, POSCORR >> P-DEP-μ is justified as the lengthened candidate (43d) is the winner. The first three candidates are wiped out as all violate one of the top-ranked constraints by presenting coalescence, a placeless onset or by failing to show up with a correspondent for /r₁/. The final candidate manages to pass all these constraints, only at the expense of low-ranked P-DEP-μ. As a result, the initial vowel lengthens.

(43) CL of singleton /r/ word-initially: /rasu/ → [á:su]

UNIFORMITY, *ONSET/r, POSCORR >> P-DEP-μ

	/r ₁ a ₂ ^μ s ₃ u ₄ ^μ /	UNIFORMITY	*ONSET/r	POSCORR	P-DEP-μ
a.	a _{1,2} ^μ ..	*!			
b.	r ₁ a ₂ ^μ ..		*!		
c.	a ₂ ^μ ..			*!	
^{ESP} d.	a ₂ ^{μ1μ} ..				*

5.5 Singleton /r/ word-medially

In contrast to the word-initial position, word-medially no lengthening occurs. While this may seem at first glance puzzling, it proves quite a sensible thing. Imagine what would happen to an input like /l₁u₂^μr₃l₄^μ/ if lengthening took place. We would either expect *luri* [l₁u₂^{μ2μ3}l₄^μ] or *lui:* [l₁u₂^μl₄^{μ3μ4}]. In other words, a sequence V:V (or VVV) would be created. As it seems, such sequences tend to be avoided in languages as observed by Kavitskaya (2002: 43) who claims that the ban on V:V “... can be motivated by perceptual properties of VV sequences: since vowel-to-vowel **transitions** (emphasis added mine) are always very long, a two-vowel sequence is not likely to be re-interpreted as a three-vowel one”. I would like to propose that this could be encoded in the markedness constraint in (44), which of course is very highly-ranked in SamG.

(44) *SUPER-LONG VOCALIC HIATUS (*S-L VH) / *V:V

There is a reason why I have changed Kavitskaya’s term from ‘transitions’ to ‘hiatus’ in the markedness constraint. As Moira Yip (p.c.) points out, we cannot talk about ‘transitions’ in cases where the vowels flanking *r* are the same. Keeping the term ‘transitions’ would on the one hand correctly entail that when the surrounding vowels are different, no lengthening occurs after *r*-deletion, e.g. *luri* > *luí*, **luí:*, but on the other hand incorrectly predict that in a word like *θiriða* we would get *θi:íða*, instead of the actual **θiíða*, that lacks lengthening (cf. (3)). For this reason, the more neutral term ‘hiatus’ is chosen that also bans **θi:íða*.

(45) no CL in singleton /r/ word-medially: /luri/ → [luí]

UNIFORMITY, *S-L VH, *ONSET/r >> POSCORR >> P-DEP-μ

	/l ₁ u ₂ ^μ r ₃ l ₄ ^μ /	*ONSET/r	*S-L VH	POSCORR	P-DEP-μ
a.	..u ₂ ^μ r ₃ l ₄ ^μ ..	*!			
b.	..u ₂ ^{μμ3} l ₄ ^μ ..		*!		*
c.	..u ₂ ^μ l ₄ ^{μ3μ} ..		*!		*
^{ESP} d.	..u ₂ ^μ l ₄ ^μ ..			*	

*S-L VH ensures that (b) *luri* and (c) *lui:* are eliminated. (a) fails, because it does not delete the onset /r/. This time the winning candidate is actually one where POSCORR is violated, as in (d). No correspondent - segmental or prosodic - for /r₃/ exists, thus the winner shows no lengthening. As a result, our ranking needs to be slightly

modified. The relationship between (a) and (d) now offers the ranking argument *ONSET/r >> POSCORR, yielding the adapted ranking:

(46) SamG CL (to be revised): UNIFORMITY, *S-L VH, *ONS/r >> POSCORR >> P-DEP-μ

5.6 Complex clusters of the type Cr+V+C - the simple(r) case

We have now dealt with simplex /r/. It remains to see what happens with /r/ in complex clusters. I will first consider the case where the cluster is followed by a V+C sequence. As we know, the output of a cluster of this type involves /r/ deletion and lengthening as in (d). Given the grammar employed so far, the right results obtain in a manner by now familiar.

(47) CL of /r/ in Cr+V+C: /mitrun/ → [mitú:n]
 UNIFORMITY, *ONS/r >> POSCORR >> P-DEP-μ

	/m ₁ i ₂ ^μ t ₃ r ₄ u ₅ ^μ n ₆ /	UNIFORMITY	*ONSET/r	POSCORR	P-DEP-μ
a.	...t ₃ u _{4,5} ^μ n ₆	*!			
b.	...t ₃ r ₄ u ₅ ^μ n ₆		*!		
c.	...t ₃ u ₅ ^μ n ₆			*!	
d.	...t ₃ u ₅ ^{μ4μ} n ₆				*

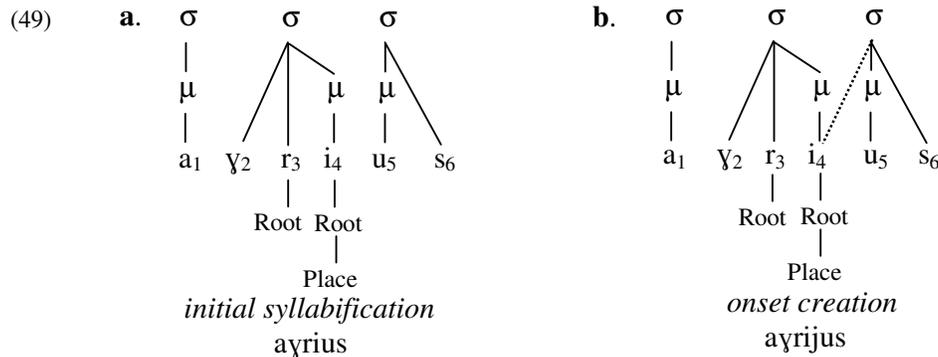
5.7 Complex clusters of the type Cr+i/e+V - the difficult case

In cases where the cluster is followed by *i/e* + *V*, no /r/ deletion occurs, but also no lengthening takes place. The question posed then is why for an input like /áγrius/ we get [áγîrjus] and not *[áγî:jus]? Katsanis considers two approaches that involve derivational epenthesis and deletion. He ends up finding both problematic (K: 57). I present some additional reasons why these are implausible.

- (48) i) áγrius → áγrijus (j-epenthesis) → áγîîjus (r-deletion) → áγîjus (coalescence) → áγîrjus (r-anaptyxis/epenthesis)
 ii) áγrius → áγrijus (j-epenthesis) → áγirijus (i-epenthesis between γr) → áγîrjus (second-i-deletion)

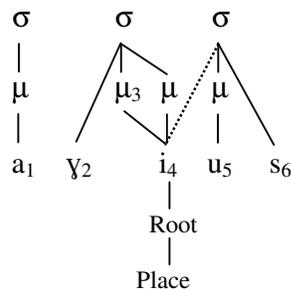
The problems (48i) faces are the following. First, /r/-epenthesis seems unlikely as it is a process generally unprecedented in Greek dialectology. But even if it was grounded, it is odd why /r/ should delete only to re-emerge later in the derivation²⁰. Moreover, why should the high central vowel coalesce, given that lengthening is not only allowed, but is in fact necessary in Cr+i/e+C? (48ii) is similarly troublesome. No good trigger for /i/-epenthesis exists, because SamG generally permits complex onset clusters e.g. *klévu* (K: 63), *kmar* (K: 64), *zmar* (K: 67)) and resolves those clusters with /r/ by deleting the latter. So why should the language choose /i/-epenthesis instead? In addition, it is extremely bizarre to argue that the second /i/ deletes, as no markedness pressure seems to be applicable here.

The alternative I offer is very different from both these approaches; I propose that what really goes on is *metathesis*, i.e. *r+i/e* becomes *i/e+r*. As a result, */r/* syllabifies in a coda and survives without getting deleted. Finally, due to */r/*'s placelessness (cf. §5.1), *i/e* can spread rightward over the */r/* and form an onset for the following syllable. The proposal should become clearer if we consider the input */ayrius/* and the possible representations its output could take stepwise. Consider first (49).



(49a) is what I call ‘initial syllabification’, i.e. the structure we should expect if nothing at all had happened. (49b) is the structure we would expect after onset creation. The next logical step would then be to get *ayijus* (50) by */r/*-deletion and subsequent lengthening, but this is not what happens. Why should this configuration be suboptimal?

(50) What *does not* happen - multiple linking in *ayijus*



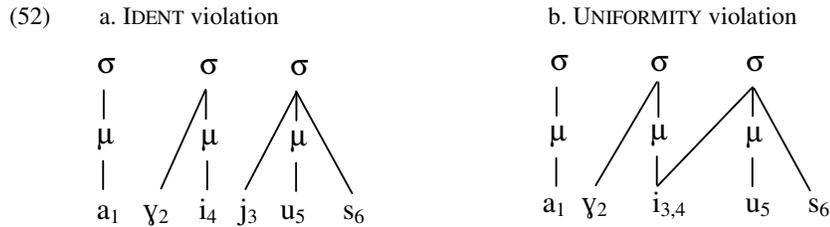
Observe that in this structure, the segment */i/* has three links to prosodic constituents (two to moras, one to a syllable). I would like to suggest that this is a configuration avoided in languages, a ban which is similar in spirit with other restrictions against ternary structures, such as the ban against trimoraic syllables (*3 μ) or against ternary feet (FTBIN). Thus, (50) can be excluded by the following proposed constraint:

(51) *3 LINKS: No ternary branching originating from a single segment

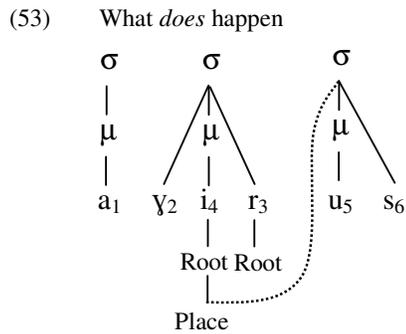
Such constraint on the surface seems to overlap with *S-L VH presented in (44) that bans V:V sequences. However, while *3 LINKS can eliminate (50), the same cannot be said for *S-L VH. The reason is that the intervening onset glide can be considered to ruin the environment upon which *S-L VH is relevant. Nonetheless, *S-L VH can be applicable if one considers other candidates like *ayius* or *ayius*, i.e. in cases

where no glide appears in an onset position²¹. Only *SL-VH but not *3 LINKS could deal with rejecting these candidates.

But there are more candidates to consider. The correspondent for /r₃/ could for instance be one that changes features of /r/ and turns it to a glide. This is not a particularly plausible candidate, but even if possible, IDENT-[F] would easily rule it out (52a). Another, slightly more plausible candidate, involves fusion, but as we have seen throughout the analysis, UNIFORMITY violations deal with this too (52b).



This leaves us with the following representation.



This configuration derives from /r/'s placelessness (cf. §5.1). /r/'s lack of a [Place] node permits /i/ to spread its place features over it and onto the onset of the next syllable without any violation of the no-crossing-line principle. At the same time, by having a metathesis between /i/ and /r/, /r/ now appears in a coda position, where it can survive (cf. §5.3) and thus no deletion occurs. Since /r/ has a segmental correspondent of itself, it satisfies POSCORR and therefore no lengthening takes place. While this form is consistent with *ONSET/r, UNIFORMITY, POSCORR, *S-L VH, *3 LINKS, IDENT-[F] (as well as DEP-SEG discussed in an endnote), it presents violations of two constraints that had not been previously discussed. One is LINEARITY which bans metathesis, as illustrated here in /...r₃i₄.../ → [...i₄r₃...]. The other is INTEGRITY which penalises splitting of a segment into two, as is the case with i₄.

To see how the right results are obtained through our tableaux, let us first consider some candidates we can easily exclude as they violate high-ranked constraints.

(54) No CL but instead metathesis: less interesting candidates

		/a ₁ ^μ γ ₂ r ₃ i ₄ ^μ u ₅ ^μ s ₆ /	*3 LINKS	*S-L VH	*ONSET/r	POS CORR	P-DEP-μ
(49)	a.	..r ₃ i ₄ ^μ u ₅ ^μ ..			*!		
	b.	..i ₄ ^{μ3μ4} u ₅ ^{μ5} ..		*!			*
	c.	..i ₄ ^{μ4} u ₅ ^{μ3μ5} ..		*!			*
(50)	d.	..i ₄ ^{μ3μ4} j ₄ u ₅ ^{μ5} ..	*!				*

The remaining candidates can be accounted for by this updated ranking:

(55) UNIFORMITY, IDENT >> INTEGRITY, LINEARITY >> P-DEP-μ

(56) No CL but instead metathesis: more interesting candidates

		/a ₁ ^μ γ ₂ r ₃ i ₄ ^μ u ₅ ^μ s ₆ /	UNIFORM- ITY	IDENT	POS CORR	INTEGR -ITY	LINEAR -ITY	P-DEP-μ
(52a)	a.	..i ₄ ^μ j ₃ u ₅ ^μ ..		*!			*	
(52b)	b.	..i _{3,4} ^μ u ₅ ^{μ5} ..	*!					
(53)	c.	..i ₄ ^μ r ₃ j ₄ u ₅ ^μ ..				*	*	

INTEGRITY and LINEARITY have to be low-ranked, because the winning candidate violates both. Moreover, we can also form an important ranking argument that LINEARITY >> P-DEP-μ as shown by re-examining simple data from CL in complex onsets.

(57) LINEARITY >> DEP-μ: /krato/ → [ka:tó] *[kartó]

		/k ₁ r ₂ a ₃ ^μ t ₄ o ₅ ^μ /	LINEARITY	P-DEP-μ
	a.	..a ₃ ^μ r ₂ ..	*!	
	b.	..a ₃ ^{μ2μ3} ..		*

This reveals that metathesis is not the preferred resolution for POSCORR satisfaction. Had it been, then we would expect metathesis rather than lengthening in cases involving Cr+V+C clusters. With all this in mind, the SamG grammar for r-deletion and CL involves:

(58) UNIFORMITY, *3 LINKS, *S-L VH, *ONSET/r, IDENT >> POSCORR, INTEGRITY, LINEARITY >> P-DEP-μ

As is evident, SamG CL presents the proposed pattern that CL needs to exhibit, namely: MARKEDNESS, UNIFORMITY >> POSCORR >> P-DEP-μ.

6 Alternatives to CL as position preservation

In this section, I will explore alternatives to CL as position preservation with special reference to SamG. Recall from the preceding discussion that the first account of SamG as sketched in Hayes (1989) fails on empirical grounds. I will thus review other accounts and will show that some fall short when they encounter the SamG data, even though they might fare reasonably well on general grounds, whereas some others fail altogether.

Two main lines of investigation will be presented. The first one includes analyses within the OT framework. Three types of accounts are examined: i) the first implicitly or explicitly accepts that an intermediate level needs to be invoked for CL treatment (Sprouse 1997, Goldrick 2000, Kiparsky 2000, McCarthy 2003, McCarthy 2005), ii) the second sees CL as segmental faithfulness (Hermans 2001) or as preservation of the numerical integrity of segments (Lee 1996) and iii) the latter is a variant of the former approach which treats CL as coalescence (Sumner 1999). I show that most of these contradict basic ideas of the OT enterprise as a whole and thus have to be discarded. One exception seems to be Sumner's (1999) approach, but this cannot account for SamG adequately.

The other line of inquiry involves a more phonetically-oriented approach as is presented in Kavitskaya (2002). In spite of some interesting aspects, the proposal exhibits certain general difficulties and some specific to SamG CL. I thus argue that the present analysis is superior as it avoids all these flaws, remains compatible with the OT tenets of parallelism and ROTB and offers a full account for the complete range of the SamG data.

6.1 OT accounts

6.1.1 Intermediate levels

The first set of analyses reviewed here makes some use of the idea that intermediate representations or levels have to be introduced to account for CL and opacity more generally. Sprouse (1997) notes that only by using an intermediate representation is it possible to express the fact that the segment that deletes and causes CL, corresponds to what would be syllabified in a coda. Hence, he explicitly admits such a level. However, in order that he restricts the proposed system so that no unlimited levels are permitted, he puts forward the idea of enriched inputs, i.e. inputs which exhibit syllabification and moraification. Not any form can be an enriched input as enriched inputs are subject to particular conditions (for more discussion, see Sprouse (1997, section 5.1)). Despite all the attempts to restrict this model, it is still quite powerful and crucially makes use of the notion of an intermediate level, which is one aspect of the model that we would like to get rid of.

Stratal Optimality Theory (Kiparsky 2000) offers a more principled way to minimise the number and nature of levels invoked by making use of morphological strata such as the stem- or word-level. The claim is that one constraint ranking may hold on one level, and a minimally different on the next level up. This could work for CL as it would be possible to say that on the first stratum the ranking **MAX-SEG**, **WBYP**, **MAX- μ** \gg **NOCODA**, **DEP- μ** maps /kan/ to [kan ^{μ}], but on the next stratum, the minimally different **NOCODA**, **WBYP**, **MAX- μ** \gg **MAX-SEG**, **DEP- μ** , forces deletion of the coda and subsequent lengthening as a response to top-ranked **MAX- μ** . Effectively, this changes /kan ^{μ} / to [kan ^{$\mu\mu$}] as desired. Once more, this account seems to imitate the operation of derivational theories, although there has been effort to subsume the strata involved to actual morphological levels. When this is possible, then Stratal OT might be an interesting contender, but when it is

not, then no obvious advantage over derivational frameworks is apparent. CL-opacity unfortunately belongs to the latter cases. It seems oblivious to morphological levels, thus its treatment in this model should employ arbitrary and unmotivated strata. This is obviously a grave defect.

McCarthy's sympathy theory (2003) faces more or less the same problems. The use of an intermediate level here is more covert by means of a sympathetic candidate, which never really emerges, but interacts with other candidates and eventually determines the correct output. Sympathy theory has faced a significant amount of criticism for a number of reasons, including the extreme complexity of the system, the introduction of intermediate forms and learnability issues as well as its failure to produce the correct results in certain occasions (cf. Bermúdez-Otero 2001). For all these reasons, I do not consider it a viable solution to the CL problem.

Goldrick's (2000) turbidity theory also refers to covert structure, but this time this is not expressed by means of a sympathetic candidate; instead, it is contained in the winner itself. This theory echoes the ideas of 'Containment' as originally proposed in Prince and Smolensky (1993), whereby deleted material was not actually deleted but still present in the output albeit unparsed. Goldrick utilizes two mechanisms: 'Projection' and 'Pronunciation'. Projection refers to the abstract structural relationship between two entities, such as the one between a vowel and its mora, whereas Pronunciation encodes the actual surface realization of that structure. Usually Projection and Pronunciation are realized on the same structure, but occasionally they do not, as in CL. In CL, a mora is projected by one segment, but is pronounced by its neighbour.

While this theory can capture CL, it too needs to refer to some form of 'intermediate' structure which in the present approach proves redundant. Moreover, a whole new set of constraints need to be introduced such as PRONOUNCE- μ , PRONOUNCE-RT, RECIPROCITY and so on whose existence seems unwarranted. Finally, under turbidity, different constraints are required to impose mora projection depending on the type of CL. In glide formation VOWEL-WEIGHT "all vowels must project their own mora" is needed, in prenasalisation NASAL-WEIGHT is needed with similar effects, and presumably in plain CVC CL - which is not discussed in Goldrick - WBYP/CODA-WEIGHT would be needed too. Apparently, in SamG CL, we would then need something like ONSET-WEIGHT. The point is that we have to employ at least four different constraints to impose weight as part of the covert structure, whereas in the CL-as-position-preservation approach, this whole issue is sidestepped because CL is simply not about mora preservation. The current model thus offers a more economical and theoretically advantageous solution.

More recently, McCarthy (2005) has presented another OT opacity-model dubbed OT-CC where CC stands for what he calls candidate chains. This framework resembles in certain aspects the harmonic serialism (HS) model of Prince and Smolensky (1993/2004). Unlike harmonic parallelism which includes just a single pass from input to output, evaluation in harmonic serialism is repeated until it converges to the correct output. However, the ranking assumed each time is the same (unlike Stratal OT, where in each stratum a different ranking is imposed). Crucially, under HS all candidates considered in each pass differ from the assumed input in just one respect (a property that inspires the use of 'candidate chains'), e.g. for an input like /kat/, candidates that are considered can only be: {kat (faithful), kati (insertion), kač (palatalization)}, but not {kači} since the latter simultaneously involves palatalization and insertion. The postulated ranking at this pass will choose one candidate as the winner. Let us assume that this is [kati], which now becomes the input for the next pass. This time {kači} is a possible candidate since it only presents palatalization. The imposed ranking in this case could be one which selects {kači} so no further evaluation takes place.

McCarthy makes use of the basic ideas of HS and builds a framework which allegedly accounts for opacity. It is outside the scope of the current work to examine this framework in detail, however a couple of notes are in place. First, note that this framework explicitly makes reference to intermediate stages, which seems to be a full circle back to serial derivational approaches. It remains to see whether such a costly move is justified. Second, McCarthy does not consider CL opacity in this work, but it seems unclear how exactly OT-CC can treat it.

To exemplify, suppose that we have an input such as $/CV_{\mu}C/$ which in the output becomes $[CV^{\mu}]$ through coda deletion and subsequent CL. According to OT-CC (see McCarthy 2005: example (8)), chains are made up of candidates which are either faithful to the input or - if less faithful - more unmarked than it. This is what is called harmonic improvement. Suppose then that in our example, coda loss occurs due to high-ranking NOCODA. This then means that the first candidate chain we need to consider is $\{CV_{\mu}C, CV_{\mu}\}$. The former candidate is totally faithful, but violates NOCODA while the latter is less faithful, but also less marked. The problem however here is that the actual winner is $CV_{\mu\mu}$. To achieve this we should perhaps first include the candidate $\{CV_{\mu}C_{\mu}\}$ in our candidate chain, so that $\{CV_{\mu}, CV_{\mu}C_{\mu}\}$. Notice though that at this stage, this candidate is less faithful - at least if the original DEP- μ evaluates it - but presents no obvious improvement in markedness (other than WBYP satisfaction, but such a constraint is contentious²²), thus it should not be a candidate included in the candidate chain. But without such a candidate it is very unclear how the desirable $CV_{\mu\mu}$ can be generated.

This hypothetical situation has been implemented by means of an analysis closer to an HS approach, which as McCarthy observes, faces difficulties with opacity. It may just as well be the case that an analysis strictly along OT-CC's lines may work. Even if it does however, it should again make reference to an intermediate stage - already problematic - plus it would require the use of $PREC(A, B)$ constraints, which stipulate that violation of constraint A must precede violation of B. Once again, this suspiciously resembles rule ordering, bringing us back closer to a serial account.

This approach then seems to explicitly deny the property that started off OT in the first place and led to its successful implementation, namely parallelism. By accepting OT-CC we run the risk of re-introducing serialism in the model and thus reverting to older theories with all the problems these faced. Moreover, if it is the case that opacity has to be handled by means of serial derivations, one wonders why such an elaborate and complex system such as the one in OT-CC has to be employed instead of the more familiar and straightforward rule-based accounts. Given that parallelism is one of the most significant tenets of OT, I contend that at least until further justification, OT-CC is not a plausible explanation for CL opacity and perhaps for opacity in general.

6.1.2 Segmental faithfulness

This type of approach keeps track of the number of segments involved (indicated by means of indices) and is in that sense similar to the coalescence approach (to be presented in the next section). Both analyses are akin to the current view, but are still different in crucial respects.

Proponents of segmental faithfulness include Lee (1996) and Hermans (2001). Lee (1996) in particular attributes CL to the preservation of the "numerical integrity of segments of a morpheme"²³. In this way, a form such as $/C_1V_2C_3C_4V_5/$ becomes $[C_1V_{2,3}C_4V_5]$. This approach fails in many respects. First, there is no

explanation why an output such as $[C_1V_{2,3}C_4V_5]$, i.e. without lengthening, should be excluded. Lee claims that the proposed analysis can handle lots of CL instances without use of moras, but it seems that only the deletion of a mora could dispose of $[C_1V_{2,3}C_4V_5]$. What is more, as the analysis progresses, moras are explicitly used to account for CL cases such as Bantu prenasalization $/muntu/ \rightarrow [mu:ntu]$. There is no point then to introduce an additional mechanism that keeps track of the segmental numerical integrity, if moras are still going to be used. Furthermore, underlying moras (cf. p. 21-26) and input syllabification (cf. p. 4, 22-24) play an important role in the generation of the patterns. Since input specification proves critical in this model, we again face the ROTB problems described earlier on.

A somewhat similar account appears in Hermans (2001), who also sees CL as segment preservation, but forms a more elaborate analysis so that he captures Hayes' claim that there is no CL after onset loss. Hermans claims that segments in an input-output correspondence relation should be identical in terms of sonority. Technically, the major segmental classes are characterised by *c*- and *v*- elements (*c*- meaning low degree of periodic energy, and *v*- meaning high degree of periodic energy). Each segment has an obligatory *c*- or *v*-head, and, depending on its type, may also have a non-head element too. This yields the following categories.

(59)

	<i>obstruents</i> c	<i>sonorants</i> c	V_2 v	V_1 v
<i>head</i>				
<i>non-head</i>		 v	 c	

V_1 refers to the first part of a long vowel, with V_2 being the second part of it. Obstruents only have a *c*-head element and V_1 's have a *v*-head only. Sonorants are consonants with some vocalic feature (non-head), while V_2 's are the reverse, by virtue of the fact that in some languages the second part of a long vowel is reduced²⁴. The idea is that three natural classes are formed: i) those with a *c*-head, i.e. obstruents and sonorants, ii) those with a *v*-head, i.e. V_1 and V_2 and iii) those whose makeup is as a whole identical, although the internal structure may differ, i.e. sonorants and V_2 .

Since segments need to belong to the same sonority class based on one of the categories above, effectively then CL cannot occur after onset loss, because the product of that type of lengthening would occupy the V_1 position. But V_1 may only be in correspondence with segments of the V_2 type, and obviously the consonantal onsets do not fall under this category. Such CL is therefore impossible. On the other hand, CL after sonorant coda loss is possible, because the product of lengthening corresponds to the V_2 position, and, as we have seen, sonorants and V_2 's may stand in correspondence.

There are two problems with this approach. The first one is that I contend that in light of the SamG data, it is wrong to conclude that CL from onset loss is impossible. Obviously, it occurs and while Hermans exceptionally seems to allow for a special case of onset CL, i.e. that of $/C_1V_2C_3V_4/ \rightarrow [C_1V_2V_3.V_4]$, both SamG patterns of word-initial and complex onset loss are missed²⁵. This is an obvious defect.

Moreover, Hermans also misses another relatively standard pattern, namely that of obstruent coda loss and lengthening, e.g. /t₁a₂p₃/ → [t₁a₂a₃]. Given that the only natural class obstruents are in is the one that also includes the sonorant consonants, I cannot see how an obstruent may delete and cause lengthening reflected on the V₂, since the latter belongs to a different class. All in all, while his proposal is an interesting one and certainly insightful in some respects, it empirically suffers, thus it is not a promising alternative.

6.1.3 CL as coalescence (Sumner 1999)

This approach also focuses on segmental preservation through the use of numerical indices. There is a certain attractive feature that this analysis makes use of which makes it stand out as more promising than the other alternatives. More specifically, it does not refer to input moraicity. In fact, although moras are included in representations, CL is not seen as the product of mora preservation. Instead, what drives CL is the following ranking:

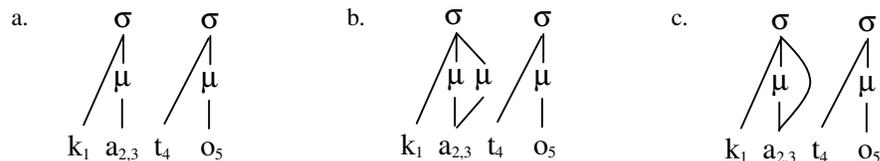
- (60) BIPOSITION >> MAX-SEG

The former constraint is defined in the following way.

- (61) BIPOSITION: An output segment representing two input segments (denoted by subscripts) must be linked to two prosodic positions [Sumner 1999: 538]

What is handy in this constraint is the reference to ‘two prosodic positions’, without necessarily imposing linking to two moras. Let us see how this approach fares when considering the Samothraki Greek cases. Things should become clearer once few possible representations are depicted. I shall first consider a rather straightforward case, such as /krato/ becoming [ka:tó]. Importantly, observe that I will consider a non-moraic input and see how the proposed analysis handles the facts.

- (62) Input: /k₁r₂a₃t₄o₅/



The first representation fails BIPOSITION since the coalesced vowel only links to one prosodic position. The third representation avoids such a problem, since the vowel has two links, one to a mora and one to a syllable; nonetheless it exhibits a structure which is at the very least controversial. Bear in mind that Rubach (1998) argues for exactly such a structure in his discussion of Slovak glides. However, even if a representation like (c) is acceptable for high vowels as glides, it seems highly unlikely for other vowels. We are thus left with (b), a candidate which satisfies BIPOSITION by virtue of the extra mora the coalesced vowel is linked to. Under this approach then, CL indeed has to do with segmental preservation, a by-product of which is vowel or consonant lengthening. Let us see these results in a tableau:

(63) *ONSET/r, BIPOSITION >> MAX-SEG

	/k ₁ r ₂ a ₃ t ₄ o ₅ /	*ONSET/r	BIPOSITION	MAX-SEG
	a. k ₁ r ₂ a ₃ t ₄ o ₅	*!		
(62a)	b. $\begin{array}{c} \mu \\ \\ k_1 a_{2,3} t_4 o_5 \end{array}$		*!	
	c. $\begin{array}{c} \mu \\ \\ k_1 a_3 t_4 o_5 \end{array}$			*!
(62b)	d. $\begin{array}{c} \mu \quad \mu \\ \diagdown \quad / \\ k_1 a_{2,3} t_4 o_5 \end{array}$			

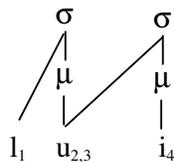
We have already discussed the candidates above with the exception of (63c). This is the candidate that presents no coalescence, so effectively /r/ has deleted altogether.

The advantage of this approach is that it needs to make no reference to underlying moraicity or particular syllabic positions. What is more, it seems much more lenient than Hayes' model in admitting CL from onset loss, which seems particularly suitable for cases like SamG. Nonetheless, Sumner acknowledges that there are cases where CL does not occur, since the segment deletes altogether and thus no lengthening applies. She is a bit cryptic at this point though in saying that there are independent constraints which decide whether segments delete or coalesce. Note that at any rate such constraints must at least dominate MAX-SEG and penalize (d) so that a candidate like (c) is preferred.

Assuming that such constraints are indeed plausible, then this approach produces good results. In particular, in Samothraki Greek, a reanalysis along Sumner's lines is possible. However, there is a single environment where her approach makes incorrect predictions. As a matter of fact, this problem also arises in one of her own examples, i.e. (23b), but she devotes no discussion on it.

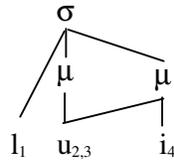
Consider our example /luri/ from (45). There we proposed a structure where no lengthening occurs, because if it did, an undesirable V:V structure would be generated. Note however that there is another structure that adheres to Sumner's proposal.

(64) /l₁u₂r₃i₄/ --- [l₁u_{2,3}.i₄] = l₁u₂w₃i₄



Although this representation captures the lack of vowel lengthening after /r/-loss word-medially, it has the implication that the high back vowel acts as an onset of the following syllable, i.e. effectively something like *lu.wi*. The same problem arises even if take [l₁u₂.i_{3,4}] as the output, i.e. *lu.ji*. These are clearly wrong since the actual output is *lui*. Perhaps, a remedy for this would be the following structure:

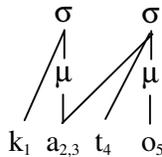
(65) /l₁u₂r₃i₄/ --- [l₁u_{2,3}i₄]



The vowels here are represented as a diphthong, but it is quite unclear what this structure entails. Furthermore, it is equally possible that the index of /r/ is linked to the vowel *i*, but as it stands we simply cannot tell. Phonetically the prediction seems to be that *u* should incorporate about two thirds of the diphthong's duration, leaving only one third for *i*.

Finally, setting aside this specific case, it is not immediately obvious why under Sumner's approach the following structure should not be generally preferred.

(66)



This is equivalent to the problematic representation in (64), but is nonetheless appealing compared to the winner (63d), since it satisfies all constraints that (d) does, but also does not incur any DEP-μ violation. It is unclear how (66) could be excluded, other than making use of phonotactic considerations, especially given that Sumner wants to keep this configuration available since she makes explicit use of it on occasion. Indeed this remains a significant problematic aspect of this model which the present approach does not share.

6.2 A phonetically-based approach - Kavitskaya (2002)

Having shown that no OT analysis constitutes a real alternative to the present approach, I will now consider a different type of account which is more phonetically-oriented. Kavitskaya (2002) proposes a model where CL is seen not as weight preservation, but as phonologisation of inherent phonetic length. The idea is that in CVC syllables, vowels are phonetically longer when followed by certain consonants whose transitions can be misheard as part of the vowel (i.e. sonorants, approximants). When these consonants delete, the "excess" length of vowels is now justified by being phonologised. Thus vowels are reinterpreted by listeners as phonemically longer. Since her model does not bear on the idea of CL as weight preservation, Kavitskaya is free to assert that onset deletion can cause CL when the consonants involved fit a certain profile.

As a matter of fact, Kavitskaya examines the SamG data too, and unlike Hayes (1989), she considers a fuller range of data, although not quite the complete set. In particular, there is no reference to the /ayrius/ → [áŷîrjus] type of cases (cf. (5)-(6)). What she claims is that the SamG *r* is vocalic enough to be reinterpreted as additional vowel length when lost. The fact that no lengthening occurs when *r* deletes in a word-medial position is attributed to a ban against VVV sequences, a position I have adopted as well. Incidentally, while such assertion also explains why no lengthening occurs in *Cr+V+V* sequences such as **ayrius*, one can only guess

why *áγîrjus is chosen instead. In fact, this pattern is not discussed at all. Moreover, the data referring to coda *r* come from Newton (1972), who claims that when *r* is in a coda position, then *i* is inserted between *r* and the following consonant. Subsequently *r* deletes as it is found intervocalically, e.g. *xartí* > *xarítí* > *xaití* ‘paper’. As we have seen earlier, in his recent study of the dialect, Katsanis (1996) claims that such forms are not accepted by modern speakers and questions the application of *i*-epenthesis generally in the dialect.

One way or another, nowadays there are words such as *karpós* or *arpázu* where the *r* is preserved intact syllable-finally, while both Kavitskaya and Katsanis accept that *r* word-finally is preserved, e.g. *fanár*. Now the question is the following: if Kavitskaya is right in the reinterpretation of *r*’s length, then why does this fail to apply when *r* is in a coda position, as in the two cases mentioned just above? This is particularly disconcerting, since coda *r*-deletion and lengthening would be the prototypical situation rather than onset *r*-deletion and lengthening. Of course, with several amendments, such as the placelessness of *r*, much of the analysis can be rescued, but this presupposes a much more prominent role for phonology than what seems to be the author’s intention.

There is one additional question; if *r* is vocalic enough and reinterpreted as extra vowel length, then why no similar result obtains with the more vocalic glides? While the status of glides in Modern Greek is notoriously controversial and confusing, there seems to be some evidence that there are some underlying glides as minimal pairs like *áðia* ‘permit, leave’ vs. *áðja* ‘empty-NEUTER-PL’ suggest (cf. Rytting 2005). It would make sense to expect a similar behaviour from glides as with the rhotic *r*, but there is no evidence for anything like that.

Finally, and on a more technical note, Kavitskaya (2002: 180-184) considers in passing a few analyses of CL within an OT framework, using the idea of moraic conservation. Importantly, she assumes that there is full syllabification and moraification in the input, a point which as we have shown is not only incompatible with core OT, but also generates wrong results. Thus, while Kavitskaya’s approach has certain appealing properties and rightly permits CL from onset loss, her analysis of SamG exhibits some defects and less worked out features.

7 Remaining issues

7.1 On the rarity of CL after onset loss

A recurring theme in this paper has been to show that “CL after onset loss” exists contra previous claims. The data from SamG have been used to corroborate this idea. However, a question that naturally comes to mind is why this pattern of CL is so exceedingly rare. To answer this question, I suggest that we also need to take into account “CL after coda loss”, i.e. the prototypical case of CL.

More specifically, consider instances of coda deletion. Although lengthening commonly follows, it is not guaranteed. What I mean by that is that not all cases of segment deletion result in CL anyway, which of course entails that numerous cases of potential CL actually never really emerge. For instance, in Diola-Fogny (Sapir 1965, Kager 1999), deletion of a coda consonant in a heterorganic cluster occurs to avoid violation of the coda condition. No lengthening of the preceding vowel arises although phonemic length exists for all vowels

(Sapir 1965: 6). For instance /let-ku-jaw/ surfaces as [lekujaw] and not *le:kujaw ‘they won’t go’ or /jaw-bu-ŋar/ becomes [jabuŋar] rather than *ja:buŋar ‘voyager’.

Another example is found in the Australian language Lardil (Kenstowicz 1994 and references cited therein). Data from Lardil stems reveal that while codas are permitted, these can only be coronals unless they are homorganic to the following onset in which case labials and velars are allowed too. Some examples are presented below.

(67) *Lardil stems and permitted codas*

i)	coronal	wu.lun	‘fruit species’
		rel.ka	‘head’
		kar.mu	‘bone’
ii)	homorganic	kuŋ.ka	‘groin’
iii)	combination	ŋam.pit	‘humpy’

Stems that show alternations between final noncoronals and \emptyset (cf. (68)) can be easily understood in the light of the generalisation above. Consonants that do not fit the profile above have to delete when in coda position (absolute form); on the other hand, when they can syllabify as onsets, they emerge intact (inflected form). Notably for our purposes, although a coda deletes, no lengthening occurs on the previous vowel, i.e. **ŋalu*: or **thurara*:, although Lardil too admits long vowels, e.g. [kudmee] ‘happy’, [maari] ‘extract’, [waaka] ‘crow’ (Hale 1981).

(68) *Stem alternations: final C-deletion*

<u>absolute</u>	<u>inflected</u>	
ŋalu	ŋaluk-in	‘story’
thurara	thurararŋ-in	‘shark’

More interestingly, Lardil also presents final vowel apocope in stems of three or more syllables. This time a vowel - which is unquestionably moraic - deletes, but again no CL applies.

(69) *Stem alternations: final V-apocope*

<u>absolute</u>	<u>inflected</u>	
yalul	yalulu-n	‘flame’
mayar	mayara-n	‘rainbow’

In fact, both V-apocope and final C-deletion may occur when the stem $\geq 3\sigma$ and the consonant in question is labial or velar. Once more, CL does not apply.

(70) *Stem alternations: V-apocope & final C-deletion*

<u>absolute</u>	<u>inflected</u>	
putu	putuka-n	‘short’
tipiti	tipitipi-n	‘rock-cod species’

Hayes (1989) acknowledges that not all instances of coda deletion lead to CL. One easy way to explain this is that in those cases, WBYP has not applied and therefore there is no mora on the coda to compensate for after its

deletion. Since the whole system ties moraicity and CL so closely, the prediction is that onsets should never trigger CL as they are non-moraic, while codas should trigger CL only when moraic. The problem now this account faces is that it ignores the effect of CL after nucleus loss. Since nuclei are *always* moraic²⁶, their deletion should always cause CL. And yet as we have seen in Lardil this is not always the case.

Numerous instances of a similar kind can be easily traced in the literature. Consulting standard textbooks is a good start. In a quick search, I have located several examples where a vowel deletes and no lengthening occurs. Some of the examples include: Chukchee vowel apocope (Kenstowicz 1994: 105), e.g. *nileq* ~ *nileqe-t* (absolute sg. ~ absolute pl.) ‘match’; Tangale vowel elision (Kenstowicz 1994: 96); Klamath first stem (short) vowel syncope between a CV-prefix and a following CV syllable, e.g. *mač^hat-ka* ‘listens’ ~ *sna-mč^hat-i:la* ‘causes to hear’ (Odden 2005: 121), and Icelandic *i*-deletion and vowel syncope (Odden 2005: 189-190). Of course, in most cases things are more complicated, so the interested reader is invited to check these works and the references cited therein for details.

In the light of the discussion above, it should be evident that the connection between moraicity and CL is thus not as robust as previously claimed. Segment deletion, even if the segment is uncontroversially moraic, i.e. a vowel, is not guaranteed to be counterbalanced by CL. This fact makes the explanation that the absence of onset CL as the product of the lack of onset weight less convincing. In other words, the absence of CL after onset loss may simply relate to the lack of CL occurring frequently in languages and not to the absence of onset weight per se. This then provides the first explanation of why no expectation for CL after onset loss is necessary. But this does not tell us much about the fact that onset CL specifically is not only fairly uncommon, but also indeed increasingly rare. I would like to suggest that there are some additional reasons at play which pertain specifically to onsets.

The most prominent is the fact that it is cross-linguistically the case that coda loss is overwhelmingly more common than onset loss. Nonetheless, some languages do present onset loss. For example, onset loss in the beginning of the word is repeatedly observed in Australian languages, but in virtually none has CL occurred. An important exception however seems to be Proto-Austronesian (PAN) initial consonant loss, which according to Zewen (1977: 9-10, 36), has led to CL, relics of which can still be seen in Marshallese (MAR) today (71). The fact is still that since onset loss is on its own quite rare, the chance for CL to a large extent obviously decreases.

(71) Proto-Austronesian initial consonant loss and CL in Marshallese²⁷

PAN	MAR	Gloss
bitun	i:ju	‘star’
binih	i:ne	‘seed’
ɣaput’	ä:ut	‘to wrap around’
pukət	o:k	‘drag net’
puna	u:n	‘origin’
hat’ap	a:t	‘smoke’

The other factor which seems to be at play is that deletion of an onset and subsequent lengthening can often lead to super-long vocalic hiatus, which as we have seen before, is universally dispreferred (Kavitskaya 2002). This type of hiatus would arise in virtually all cases where there is a singleton intevocalic onset, i.e. VCV, that gets deleted. By ranking the constraint *S-L VH highly, we can ensure that lengthening of this type cannot occur. As

a matter of fact, Samothraki Greek is an excellent example of this type. /r/ deletes from an onset position, and causes lengthening whenever this would not violate *S-L VH, namely word-initially and in a complex onset cluster. Word-medially, its deletion would result in the prohibited V:V configuration, thus no lengthening takes place. This makes the strong prediction that there will be no language with compensatory lengthening after onset loss word-medially, without equivalent lengthening word-initially. In fact, this prediction is reflected in the Marshallese example above PAN *yaput'* > MAR *ä:ut*, where the MAR form presents deletion of both the initial as well as the medial onset. Lengthening however occurs only after the former's loss. More generally then, since *S-L VH seems to be cross-linguistically highly-ranked, it falls out that it frequently destroys the environment for CL after onset loss.

If we put all these properties together, it should no longer be a surprise that "CL from onset loss" is very rare.

7.2 Onsets as a trigger and as a target of CL

So far we have looked at data where onsets function as a trigger for CL, that is, the deletion of a segment in an onset position causes lengthening of the following vowel. SamG is obviously a clear example of that sort, but it is not the only one. The Marshallese data in (71) fall under the same category and so do data from Onondaga (Iroquoian) which resembles the patterns of SamG to a large extent.

Onondaga (Michelson 1988) lost the phoneme /r/ of the Proto-Lake-Iroquoian (PLI) in all environments. More specifically, Intervocally, *r* became *w* after round vowels and *y* after *i*. In all other occasions *r* was lost. The latter case resulted in a sequence of vowels, but no lengthening appeared. More interesting is the pre- and post-consonantal *r* loss. In both cases, *r* deleted and caused lengthening of the vowel before it (pre-consonantal *r* e.g. PLI: **katq̄ryeʔs* 'I'm breathing, Onond: *katq̄ryeʔs*) or after it (post-consonantal). The latter is the focus of our attention.

(72) *Post-consonantal /r/-deletion and lengthening in Onondaga*

PLI * <i>yotshíʔkreʔ</i>	'clouds'	Onond <i>otshíʔke:ʔ</i>
PLI * <i>óʔkraʔ</i>	'snow flake'	Onond <i>óʔkæ:ʔ</i>

Michelson states that the derived vowels are truly long and tautosyllabic since they are counted as one vowel for stress purposes. Citing Woodbury (1981), Michelson proposes to explain the Onondaga pattern based on two facts. First, in dictionaries of the seventeenth and mid-eighteenth centuries, **Cr* is frequently represented as *Cer*. Second, Mohawk regularly inserts an epenthetic *e* to break up **Cr* clusters. One could then assume that Onondaga at some point inserted an epenthetic *e* too, which after *r*-deletion would coalesce with the following vowel leading to a fused long vowel, i.e. **CrV* > **CerV* > **CeV* > **CVV* > **CV:*.

While this is a possible explanation of the facts, it is not the only one. Onondaga's *r*-deletion patterns seem to be consistent with what happens in SamG, i.e. lengthening of a vowel after cluster simplification, but not when the *r* deletes word-medially²⁸. With regard to the information the dictionaries offer, numerous questions can be raised, but the most pertinent one is why only some of the **Cr* forms are represented as **Cer*. Wouldn't one expect all the forms to show up with this pattern had it been so pervasive? More importantly,

other than the dictionaries, there does not seem to be any conclusive empirical evidence that indeed there was *e*-epenthesis in the **CrV* context. The fact that it occurs in Mohawk does not imply that it happens in Onondaga. What is more, Mohawk has extensive *e*-epenthesis, which is in direct contrast to the distribution of *e*-insertion in Onondaga. As a matter of fact, Michelson herself (1988: 146-7) describes just a single case of *e*-epenthesis, namely between an extrasyllabic non-laryngeal consonant and a following consonant.

All these facts show that this issue has not been resolved and should at the very least make us sceptic towards the **Cer* approach. A compensatory lengthening approach is available, especially given the resemblance of the facts to the SamG data. This matter then remains open for further investigation.

Apart from the onset as a trigger for CL, we can also find data where presumably the onset is the target of CL. In other words, the loss of a position causes lengthening of an onset consonant which is now rendered a geminate. Certain dialects of Trique (Oto-Manguan) provide a compelling example of this type, whereby a vowel is lost or shortened and the preceding consonant lengthens as a response to that.

Trique is a tonal language, which also shows stress, consistently on the final syllable. This syllable is special in being able to license numerous tonal, consonantal and vocalic contrasts (see Hollenbach 1977 for details). One of the contrasts in the San Juan Copala dialect of Trique (henceforth SJC), involves the presence of fortis stops and fortis sibilants in the onset of the final syllable and nowhere else. The fortis consonants are voiceless and unaspirated (Hollenbach 1977: 36-37). Phrasal stress is marked both by the intensity on the ultima of the word receiving it and by lengthening of open long vowels or shortening of short vowels. “The latter is accompanied by *compensatory lengthening of an immediately preceding fortis stop, fortis sibilant or resonant* (Hollenbach 1977: 49; emphasis added mine)”.

If this is accurate, then it seems to be the case that reduction of the short vowels is compensated for by lengthening of the preceding **onset** consonant. Given that it is the stressed syllable which gets affected and that Copala Trique lacks codas (with the exception of word-final coda, which has to be a laryngeal (Hollenbach 1977: 36)), it seems reasonable that the newly created geminate syllabifies wholly in the onset and is rendered a moraic onset (a position compatible with Topintzi 2006).

A similar instance of CL appears in the San Andrés Chicahuaxtla (SAC) dialect of Trique. SAC differs from SJC in having both fortis resonants /m: n: l: y: w:/ which only appear in monosyllables and lenis sonorants /m n l y w/. The characteristic of fortis resonants is that they are quite long (Hollenbach 1977: 50). Interestingly, SAC fortis resonants developed mainly by a lengthening of simple resonants to compensate for the loss of a penult (Longacre 1957: 18)²⁹.

(73) *SAC Trique fortis resonants* (Hollenbach 1977: 58; a dot underneath a vowel indicates it is short)

SJC	yumɛ	SAC	mj	‘sweet potato’
SJC	yana	SAC	n:a	‘loft’
SJC	yuwe	SAC	w:e	‘century plant’

If initial moraic geminates are best described as weightful onsets as Topintzi (2006) argues, then this is another case of CL that *leads* to a moraic onset.

7.3 CL and consonant loss rather than syllable constituent (coda, onset) loss

The data from Trique above seem to support that onsets can be moraic. This possibility opens up a huge issue that merits detailed discussion and as a result cannot be tackled in the present paper (but see Topintzi 2006). Although this conclusion seems practically inescapable in Trique, it is not of direct importance to SamG. Indeed, the proposed analysis for SamG CL does not hinge on that issue. What is more, an analysis that ignores input moraicity has been devised which side-steps this issue. This was partly done for theoretical reasons, so as to conform to some of the major tenets of OT, such as avoidance of serialism (as serialism contradicts parallelism) and liberation from input restrictions (as input restrictions contradict Richness of the Base). As explained before, had the onset been allowed to carry a mora, then we would run against ROTB. Had we allowed an analysis along Hayes' (1989) lines, then we would need to impose serialism.

What is more, it should by now be evident that if CL works pretty much in the way described in this paper, then reference to terms such as "CL after coda loss" or "CL after onset loss" does not make much sense, as there is no point in the 'derivation' during which these consonants syllabify. Although the terms above are descriptively convenient, the lost segments cannot be identified as codas or onsets, since in the input no syllabification applies, while in the output, the segments in question no longer surface. Consequently, Hayes' claim that CL occurs only after coda loss is significantly weakened. It thus makes much more sense to simply talk about CL after consonant deletion.

8 Conclusion

In this paper, I have proposed an analysis that treats compensatory lengthening not as mora conservation, but instead as position preservation through a mora. A segment requires an output correspondent, either segmentally as a root node, or prosodically as a mora. This is dubbed POSCORR. However, when the segment deletes, only the prosodic correspondent solution remains. In the languages where POSCORR is of high priority, lengthening through a mora occurs as a response to its satisfaction. The net effect is CL. In the languages where such lengthening is banned, POSCORR is either sacrificed so that no output correspondent for the segment exists, or it is satisfied by having a segmental correspondent by means of fusion. Moraicity is then seen as the remedy, rather than the trigger of CL.

Unlike previous approaches, this analysis does not require reference to underlying moras (Lee 1996, Lin 1997) or intermediate representations (Hayes 1989) to account for CL. Given a certain ranking, CL can occur in a familiar input-output manner, thus this solution is obviously advantageous because it does not compromise OT's orientation to the output. It also does not conflict with other important tenets, such as the Richness of the Base, which although not entirely uncontroversial, has nonetheless remained influential.

Apart from being theoretically advantageous, the analysis has empirical merits too. Without any additional assumptions, it can account for the existence of CL in languages which present deletion of an onset, a situation that is explicitly ruled out as an option in virtually all accounts of CL. If CL is about position preservation through moras, then it is actually expected that we should get such cases since syllabification is

beside the point. What matters are positions and whether there is a prosodic correspondent to the position emptied by a deleted segment.

In this respect, “CL as position preservation” treats all segments as equal and thus as equally able to cause CL³⁰. The fact that CL from onset is not as common, probably relates to the rarity of onset loss in the first place as well as the avoidance of super long vocalic sequences that would be caused word-medially had onset loss led to lengthening.

The current analysis of course does not claim to be a full theory of CL, since this would require exhaustive investigation of the reported CL cases. This is by far beyond the scope of the present work. Presumably, certain additions and modifications would be required to account for the full inventory of CL cases. One thing is sure though. Cases that can be descriptively dubbed as “CL after onset loss” unambiguously exist and must be accounted for satisfactorily and directly. This paper hopes to have done exactly that.

Notes

* I am most grateful to Moira Yip for extensive discussion on this material and invaluable suggestions that helped me improve it. I also wish to thank two anonymous reviewers of the Journal of Greek Linguistics, the audiences at the 17ISTAL, ICGL7 and at the University of Patras Colloquium Series for useful comments as well as Marianna Ronga and Maria Tsolaki for help with the data. All errors are of course my own.

¹ /r/-deletion but no lengthening is also reported in the dialect of Taifiri of Kallipolis and occasionally in the dialect of Kapi in Lesbos (Papadopoulos 1927). Papadopoulos also reports /l/-deletion in Samothraki Greek, a non-productive and practically no longer attested process also acknowledged in Katsanis (1996) and Afroudakis (p.c.). Unlike them however, Papadopoulos claims that sometimes /l/-deletion leads to CL. Thus, we get interesting examples where both /r/- and /l/-deletion apply with subsequent CL, e.g. *lutrá* > *urtá*: ‘baths’.

² Andriotis (1926-1928) suggests that /r/-deletion and lengthening occurs in a coda position too, e.g. *várka* > *vá:ka* ‘boat’. Other works of the same period, e.g. Papadopoulos (1927) agree with more recent accounts such as Krekoulas (1964), Afroudakis (1985) and Katsanis (1996), who report that coda /r/ survives.

³ Raising in [mía] is unexpected, since the mid-vowel is stressed. For this idiosyncratic process, the reader should consult Katsanis (1996: 37-38). It is also unclear why [θííða] receives stress on the first vowel instead of [θííða].

⁴ I maintain Katsanis’ representation of [î] and [ê]. However according to Ronga (p.c.), [î] is most likely IPA [i] and [ê] is [ɐ]. She also maintains that [i] is the unstressed realization of [ê].

⁵ According to Maria Tsolaki (p.c.), some of the words above are instead pronounced as: *axérjastus*, *kirjás*’ (where the final *s* is palatalized) and *gémnus*.

⁶ A reviewer notes that Turkish has the word *harita* ‘map’ that seems to be from the Greek [xarta]. The presence of the nonharmonic -i- in the Turkish root could possibly suggest a borrowing with -i- already in place. This could be then used as indirect evidence for a form *hariti* for the word ‘paper’ in SamG. Baris Kabak (p.c.) suggests that this form was indeed borrowed into Turkish from Greek, but via Arabic (a suggestion also corroborated by dictionary entries) with the vowel /i/ already present. Islam Youssef (p.c.) further reports that the word [xaarita] exists for some Arabic speakers with the meaning ‘map’, although the actual dictionary meaning is ‘restive animal’. However, most Arabic speakers use the form [xariita] with a long underlying /i/ vowel which has both the meaning of ‘map’ and another genuine Arabic meaning ‘leather pot’. Two points seem pertinent at this stage. First, while the /i/ seems to be present underlyingly, this cannot be certain, since Arabic has various epenthetic processes to break up clusters (although restrictions are loosened across syllable edges). Perhaps then, /i/ is epenthetic rather than copied from the original Greek word. Second, since vowel epenthesis is virtually unattested in SamG, it is unlikely that Arabic - even with the /i/ already present - borrowed ‘map’ from this dialect of Greek.

⁷ Constraints used: NOCODA: codas are banned; MAX-μ: the total number of input moras is the same with the total number of output moras, i.e. no mora deletion; MAX-SEG: no deletion of input segments; DEP-μ: the total number of output moras is the same with the total number of input moras, i.e. no mora insertion.

⁸ Here and throughout, “moraic specification” refers to consonantal moras. I assume that vowels are always moraic, hence the term - as used in the paper - contrasts /CV^HC/ (no moraic specification) with /CV^HC^H/ (moraic specification).

⁹ However, this should not be taken to mean that moraic onsets are generally impossible. Topintzi (2006) argues extensively that moraic onsets exist in some languages providing evidence from Bella Coola, Pattani Malay, Pirahã, Marshallese and others.

¹⁰ Which is why the present analysis improves significantly on the one proposed for SamG in Topintzi (2005a). There, underlying moraicity of /r/ was an important assumption for the analysis. This approach also predicted that word-medial /r/-deletion would lead to a slight phonetic lengthening which would not really be perceived. On the contrary, now we can straightforwardly express the fact that no lengthening occurs in this environment at all. For comparison, the reader is referred to that work.

¹¹ An additional alternative somehow analogous to (14a) would have mora fusion (although μ_3 does not necessarily have a correspondent in the input): [k₁a₂^{μ_{2,3}}], but this would apparently violate some sort of UNIFORMITY with respect to moras.

¹² This raises the question of what the winning candidate in a language without CL would look like. This is tackled next.

¹³ Tranel (1987: 49-50) also discusses lengthening in French. There are certain conditions under which this may occur. As far as nasal vowels are concerned, lengthening applies when they are in a stressed and *closed* syllable. In the cases mentioned here, the nasal vowel is not in a closed syllable, therefore no lengthening may show up.

¹⁴ We could also consider other structures, which nonetheless have no direct bearing to the present discussion. For completeness though, consider non-moraic /CV^μC/ leading to either [CV^μC] or to moraic [CV^μC^μ]. Both outputs are positionally μ -licensed since only one mora directly dominates each of the mora-bearing segments, thus these forms vacuously satisfy P-DEP- μ which evaluates non-positional μ -licensors. We could also consider input geminates and geminated consonants. If a geminate is a /C^μ/ that in the output also acquires a link to the onset of the next syllable, then it would not be positionally μ -licensed, as both a mora and a syllable immediately dominate it in the output. However, P-DEP- μ is satisfied given that the geminate's mora has a correspondent in the input. Geminated consonants are a different story though. These are presumably /C/ in the input and receive the same output representation with original geminates. The difference now lies in the fact that their mora has no correspondent in the input and hence they violate P-DEP- μ . This is presumably a desirable result, at least in the cases where we would like to penalise surface gemination that has no source underlyingly, but admit one that stems from the input.

¹⁵ A reviewer correctly observes that sequences such as VCC sometimes end up as V:C (i.e. with CL) and sometimes simply undergo cluster reduction, i.e. VC. Although it is an interesting question to ask why this variation occurs, it is beyond the scope of the current paper, thus I leave it open for future research.

¹⁶ I choose here a transcription like [k^{wi}:Bona] instead of Sagey's [kwiiBona] so that I simultaneously stress the fact that we are talking about complex segments with multiple articulations as well as long vowels (rather than a sequence of identical vowels).

¹⁷ The language bans codas, so I do not consider candidates that have codas.

¹⁸ In the case of complex onsets, centralization occurs when the C=velar and the V=front. I abstract away from this process in the present paper.

¹⁹ We could also consider the candidate: [..a₂^μr₃^μ..] that syllabifies the /r/ as a moraic coda. Although /r/ would escape a P-DEP- μ violation due to the fact that its mora is a positional μ -licensor, it would be presumably ruled out by something like *MORAIC CODA which bans weightful codas.

²⁰ A reviewer points out that this argument is less strong given that the introduction of innovative processes in a language is one of the factors that produce differentiation among dialects. Moreover, since one can find other cases where seemingly opposite processes occur in complementary environments in one and the same language, e.g. English (see below), then processes like SamG r-deletion and subsequent r-epenthesis can be considered plausible. While I will agree with the first point, I believe that the second one is less convincing. More concretely, in the non-rhotic dialects of English, /ɹ/ deletes in a coda position, but re-appears by means of the so-called "intrusive-r" in cases like *the idea* [ɹ] *is*... Note that the "intrusive-r" is due to epenthesis and as such contrasts with the "linking-r" that is present underlyingly and can emerge in the onset in cases like: *ca*[ɹ] *alarm*, c.f. Bermudez-Otero (2007). In this view, the scenario in (48i) would be parallel to the English case, as outlined here, since it too involves r-deletion and r-epenthesis. Importantly however, in the SamG case, r-deletion would occur in an onset position and r-epenthesis in a coda position, that is, the inverse of what happens in English, where r-deletion avoids or simplifies codas and r-epenthesis offers an onset to an otherwise onsetless syllable. Both processes seem well-motivated, but this is not the case in SamG; even if r-deletion in an onset position is reasonable (since in the case of clusters it simplifies complex onsets), the same cannot be said for the inserted /r/ in codas, which only complicates syllable structure. In sum, the r-deletion-and-r-epenthesis processes are not sufficiently motivated in SamG.

²¹ We could alternatively provide an onset by inserting a glide instead of spreading. This would avoid the *S-L VH violation, but would violate DEP-SEG, the constraint against segment insertion, which I claim is highly ranked in the language.

²² I mean ‘contentious’ here in a very specific way. Although WBYP, or its equivalent MORaic CODA, are widely used in the literature (Morén 1999/2001; McGarrity 2003; Crosswhite 2001; Broselow, Chen and Huffman 1997 among many others), this constraint is basically a convenient stipulation to assign moraicity onto coda consonants, which would otherwise surface as non-moraic. Moreover, in many cases, WBYP is used alongside its opposite *MORaic CODA (for striking examples, it suffices just to look at Crosswhite 2001, or McGarrity’s 2003 tableaux (24) and (25)). Both these constraints are markedness constraints, but it is unclear what the unmarked option for a language really is: should codas be moraic or not? I will not explore this issue in this work. It should nonetheless be evident that OT-CC could only be viable if it assumed that moraic codas are the unmarked option for a language.

²³ As far as I can see, there is no definition of the notion ‘numerical integrity’ anywhere, but we are being told how this can be achieved, i.e. by the general ranking MARKEDNESS >> MAX >> FAITH.

²⁴ I am not convinced that this renders the ‘v-head, c-non-head’ as V₂’s, but for the sake of clarity, I am following Hermans’ suggestion.

²⁵ Plus the only exceptional pattern that Hermans admits is exactly the one that SamG bans (i.e. lengthening after word-medial deletion).

²⁶ This is overwhelmingly the case. However, there are instances where some vowels have been claimed to contribute no mora. See for instance Shaw (2004) who argues that schwa in Mohawk is non-moraic.

²⁷ Cluster simplification did not lead to CL, as in PAN *mpənəd > MAR bəŋ ‘obstructed’. Also note that /b/ stands for a rounded bilabial stop and /t/ for an unreleased dental stop.

²⁸ Onondaga differs from SamG in also presenting deletion of /r/ from the coda with subsequent lengthening. With respect to Onondaga initial /r/, Michelson (1988) reports that stem-initial /r/ alternates with y/∅, but at least in the examples provided, this does not coincide with the word-initial position.

²⁹ Note that the Trique CL facts probably suggest some modification to the CL theory presented in this paper with regard to assumptions about root node preservation in reduced vowels (SJC) and the addition of only one mora even when two segments delete (apparently due to limitations of the number of moras per syllable) as in SAC.

³⁰ At a closer inspection, some refinement may be required, since it is an oversimplification to say that all segments are equal. For instance, epenthetic segments in some languages are invisible for certain processes, e.g. the Marshallese epenthetic yV- prefix in reduplication (Hendricks 1999).

References

- Afroudakis, Angelos [Αφρουδάκης, Άγγελος] (1985) *Collection of language material from the island of Samothraki* [Συλλογή γλωσσικού υλικού από τη νήσο Σαμοθράκη]. Manuscript No. 1293, Research Centre of Modern Greek Dialects and Idioms, Academy of Athens (Κέντρον Ερεύνης των Νεοελληνικών Διαλέκτων και Ιδιωμάτων - Ι.Α.Ν.Ε Ακαδημίας Αθηνών)
- Akinlabi, Akinbiyi (1993) Underspecification and the Phonology of Yoruba /r/. *Linguistic Inquiry* 24(1): 139-160
- Andriotis, Nikolaos [Ανδριώτης, Νικόλαος] (1926-1928) *Collection of language material from Samothraki* [Συλλογή γλωσσικού υλικού εκ Σαμοθράκης]. Manuscript No: 449, Research Centre of Modern Greek Dialects and Idioms, Academy of Athens (Κέντρον Ερεύνης των Νεοελληνικών Διαλέκτων και Ιδιωμάτων - Ι.Α.Ν.Ε Ακαδημίας Αθηνών)
- Bermúdez-Otero, Ricardo (2001) Underlyingly nonmoraic coda consonants, faithfulness and sympathy. ms., University of Manchester

- Bermúdez-Otero, Ricardo (2007) Word-final prevocalic consonants in English: representation vs. derivation. Paper presented at OCP4, Rhodes, Greece, 18-21 January 2007
- Broselow, Ellen, Su-I Chen and Marie Huffman (1997) Syllable weight: convergence of phonology and phonetics. *Phonology* **14**: 47-82
- Campos-Astorkiza, Rebeka (2004) Faith in Moras: A Revised Approach to Prosodic Faithfulness. In *Proceedings of North East Linguistics Society* **34**, K. Moulton & M. Wolf (eds.), Amherst, MA: GLSA
- Clements, George N. (1986) Compensatory Lengthening and Consonant Gemination in LuGanda. In *Studies in Compensatory Lengthening*, Leo Wetzels and Engin Sezer (eds.), pp. 37-77, Dordrecht: Foris
- Crosswhite, Catherine (2001) Predicting syllabicity and moraicity in Dihovo Macedonian. In *University of Rochester Working Papers in the Language Sciences* **2(1)**, J. S. Magnuson and K. M. Crosswhite (eds.), pp. 1-22
- Goldrick, Matthew (2000) Turbid Output Representations and the Unity of Opacity. In *Proceedings of the North East Linguistic Society*, Vol. 1, Masako Hirotoni, Andries Coetzee, Nancy Hall and Ji-yung Kim (eds.) pp. 231-245, Amherst, MA: GLSA
- Hale, Ken (1981) Part I: Remarks on Lardil phonology and morphology. In *A Preliminary dictionary of Lardil*, Ken Hale et al., Document 0028 of the Aboriginal Studies Electronic Data Archive (ASEDA), Canberra: AIATSIS (Australian Institute of Aboriginal and Torres Strait Islander Studies)
- Hayes, Bruce (1989) Compensatory Lengthening in Moraic Phonology. *Linguistic Inquiry* **20**: 253-306
- Heisenberg, August (1921) Die Liquida P im Dialekt von Samothrake. In *For G.N Chatzidakis [Αφιέρωμα εις Γ.Ν. Χατζιδάκιν]*, pp. 89-99, Athens: P.D. Sakellarios [Ἐν Αθήναις: Π.Δ. Σακελλάριος]
- Hendricks, Sean (1999) *Reduplication without template constraints: A study in bare-consonant reduplication*. Doctoral Dissertation, University of Arizona
- Hermans, Ben (2001) CL in OT. *Linguistics in the Netherlands* **18**: 139-150
- Hollenbach, Barbara (1977) Phonetic vs. phonemic correspondence in two Trique dialects. In *Studies in Otomanguean phonology*, William R. Merrifield (ed.), pp. 35-67, Dallas: Summer Institute of Linguistics and University of Texas at Arlington
- Kager, René (1999) *Optimality Theory*. Cambridge: Cambridge University Press
- Katsanis, Nikolaos [Κατσάνης, Νικόλαος] (1996) *The dialect of Samothraki Greek [Το γλωσσικό ιδίωμα της Σαμοθράκης]*. Municipality of Samothraki [Δήμος Σαμοθράκης]
- Kavitskaya, Darya (2002) *Compensatory Lengthening: Phonetics, Phonology, Diachrony*. New York and London: Routledge
- Kenstowicz, Michael (1994) *Phonology in Generative Grammar*. Oxford: Blackwell
- Kiparsky, Paul (2000) Opacity and Cyclicity. *The Linguistic Review* **17**: 351-367
- Krekoukias, Dimitrios A. [Κρεκούκιας, Δημήτριος Α.] (1964) *Collection of language material from Samothraki [Συλλογή γλωσσικού υλικού εκ Σαμοθράκης]*. Manuscript No 839, Research Centre of Modern Greek Dialects and Idioms, Academy of Athens (Κέντρον Ερεύνης των Νεοελληνικών Διαλέκτων και Ιδιωμάτων - Ι.Α.Ν.Ε Ακαδημίας Αθηνών)
- Lee, Byung-Gun (1996) The Emergence of the Faithful. Available as ROA-166
- Lin, Yen-Hwei (1997) Syllabic and moraic structures in Piro. *Phonology* **14**: 403-436

- Lombardi, Linda (1990) The nonlinear organization of the affricate. *Natural Language and Linguistic Theory* **8**: 375-425
- Longacre, Robert E. (1957) Proto-Mixtecan. Publication 5 of the Indiana University Research Center in Anthropology, Folklore, and Linguistics, published as Part 3 of Vol. 23, No. 4 of the *International Journal of American Linguistics*
- Mateus, Maria Helena and Ernesto d'Andrade (2002) *The Phonology of Portuguese*. Oxford: Oxford University Press
- McCarthy, John (2003) Sympathy, Cumulativity, and the Duke-of-York Gambit. In *The Optimal Syllable*, Caroline Féry and Ruben van de Vijver (eds.), pp. 23-76, Cambridge: Cambridge University Press
- McCarthy, John (2005) Candidate chains. Paper presented at OCP2, Tromsø, Norway
- McCarthy, John and Alan Prince (1995) Faithfulness and reduplicative identity. In *University of Massachusetts Occasional Papers in Linguistics* **18**, Jill Beckman, Laura Walsh Dickey and Suzanne Urbanczyk (eds.), pp. 249-384, Amherst, MA: GLSA
- McGarrity, Laura W. (2003) *Constraints on patterns of primary and secondary stress*. Doctoral Dissertation, Indiana University
- Michelson, Karin (1988) *A Comparative Study of Lake-Iroquoian Accent*. Dordrecht: Kluwer
- Morén, Bruce (2001) *Distinctiveness, Coercion and Sonority: A Unified Theory of Weight*. Doctoral Dissertation 1999, University of Maryland at College Park. Published 2001 by Routledge
- Newton, Brian (1972) Loss of /t/ in a Modern Greek dialect. *Language* **48**: 566-572
- Odden, David (2005) *Introducing Phonology*. Cambridge: Cambridge University Press
- Papadopoulos, Anthimos [Παπαδόπουλος, Άνθιμος] (1927) *Grammar of Modern Greek Northern Idioms [Γραμματική των Βορείων Ιδιωμάτων της Νέας Ελληνικής]*. Athens: P.D. Sakellarios [Εν Αθήναις: Π.Δ. Σακελλάριος]
- Prince, Alan and Paul Smolensky (1993) *Optimality Theory: Constraint Interaction in Generative Grammar*. Published 2004, Malden, MA; Oxford, UK; Victoria, Australia: Blackwell
- Rice, Keren (1992) On Deriving Sonority: A Structural Account of Sonority Relationships. *Phonology* **9**: 61-99
- Rosenthal, Sam (1994) *Vowel/Glide Alternation in a Theory of Constraint Interaction*. Doctoral Dissertation, University of Massachusetts, Amherst. Published 1997, New York and London: Garland Publishing
- Rubach, Jerzy (1998) A Slovak argument for the onset-rhyme distinction. *Linguistic Inquiry* **29**: 168-179
- Rytting, Anton C. (2005) An iota of difference: Attitudes to yod in lexical and social contexts. *Journal of Greek Linguistics* **6**: 151-185
- Sagey, Elizabeth (1986) *The Representation of Features and Relations in Nonlinear Phonology*. Doctoral Dissertation, MIT
- Sapir, Edward (1965) *A grammar of Diola-Fogny*. Cambridge University Press
- Shaw, Patricia A. (2004) Inside Access: The Prosodic Role of Internal Morphological Constituency. In *The Nature of Word: essays in honor of Paul Kiparsky*, K. Hanson and S. Inkelas (eds.), Cambridge, MA: MIT Press
- Sprouse, Ronald (1997) A case for enriched inputs. Handout for paper presented at TREND VII, held May 3, 1997 at UCSC. Available as ROA-193

- Sumner, Meghan (1999) Compensatory lengthening as coalescence: analysis and implications. In *Proceedings of WCCFL 18*, S. Bird, A. Carnie, J. Haugen and P. Norquest (eds.), pp. 532-544, Somerville, MA: Cascadilla Press
- Topintzi, Nina (2005a) Solving the Samothraki Greek Compensatory Lengthening Puzzle. In the *Proceedings of the 17th International Symposium of Theoretical & Applied Linguistics (17ISTAL)*, Thessaloniki, Greece, 15-17 April 2005
- Topintzi, Nina (2005b) Compensatory lengthening in Samothraki Greek: some problems and theoretical consequences. In *The Proceedings of the 7th International Conference on Greek Linguistics (ICGL7)*, York, UK, 8-10 September 2005
- Topintzi, Nina (2006) *Moraic onsets*. Doctoral Dissertation, UCL
- Tranel, Bernard (1987) *The sounds of French: an introduction*. Cambridge: Cambridge University Press
- Woodbury, Hanni (1981) The loss of a phoneme. *International Journal of American Linguistics* **47**: 103-120
- Yip, Moira (1991) Coronals, Consonant Clusters and the Coda Condition. In *The Special Status of Coronals. Internal and External Evidence*, Caroline Paradis and Jean-François Prunet (eds.), pp. 61-78, San Diego: Academic Press
- Zewen, François-Xavier Nicolas (1977) *The Marshallese language: A study of its phonology, morphology, and syntax*. Berlin: Verlag von Dietrich Reimer

Περίληψη

Η εργασία αυτή μελετά την αποβολή του /t/ από θέση έμβασης και την επακόλουθη αναπληρωματική έκταση (ΑΕ) που πραγματώνεται με την επιμήκυνση του ακόλουθου φωνήεντος. Το φαινόμενο αυτό θέτει δύο σημαντικά προβλήματα για την τρέχουσα φωνολογική θεωρία. Πρώτον, σύμφωνα με τη μοραϊκή θεωρία οι εμβάσεις δε φέρουν ποτέ συλλαβικό βάρος, επομένως η αποβολή τους δε θα πρέπει ποτέ να οδηγεί και σε ΑΕ. Δεύτερον, η ΑΕ συνιστά μία κλασική περίπτωση φωνολογικής αδιαφάνειας (opacity), η οποία καθίσταται το σημαντικότερο 'αγκάθι' για τη Θεωρία του Βέλτιστου (Optimality Theory). Το παρόν άρθρο αντιμετωπίζει και τα δύο αυτά προβλήματα και προτείνει ότι η ΑΕ πρέπει να αναλυθεί ως διατήρηση **θέσης** μέσω της παρουσίας μορών και όχι ως διατήρηση των μορών αυτών καθεαυτών, όπως ισχυρίζεται ο Hayes (1989). Βασικό πλεονέκτημα αυτής της ανάλυσης είναι ότι καμία αναφορά στον αριθμό των μορών του εισαγομένου δε χρειάζεται να γίνει και συνεπώς οι εμβάσεις, όπως και οι έξοδοι, μπορούν να οδηγήσουν σε ΑΕ. Επιπλέον, το προτεινόμενο μοντέλο μεριμνά ώστε να είναι σύμφωνο με βασικές αρχές της Θεωρίας του Βέλτιστου, όπως η αντιστοίχιση εισαγομένου-εξαγομένου χωρίς ενδιάμεσα στάδια και ο Πλούτος της Βάσης (Richness of the Base). Καταφέρνει έτσι να εξηγήσει πολυάριθμες περιπτώσεις αναπληρωματικής έκτασης, ενώ εφαρμόζεται με επιτυχία στη Σαμοθρακίτικη ΑΕ, διεξοδική ανάλυση της οποίας ακολουθεί με σκοπό να μελετηθεί το φαινόμενο στην πλήρη του διάσταση.