

Compensatory Lengthening

Nina Topintzi

Universität Leipzig

nina.topintzi@uni-leipzig.de

0. Roadmap

§1: Brief overview of what CL is about and Hayes (1989) along with the problems it carries along

§2: CL Opacity in OT

or “How a simple analysis needs to get complicated because of theory-internal reasons”

§3: CL after loss of onset or weightless coda or “How Hayes (1989) is too restrictive”

§4: CL asymmetries or “How Hayes (1989) is not restrictive enough”

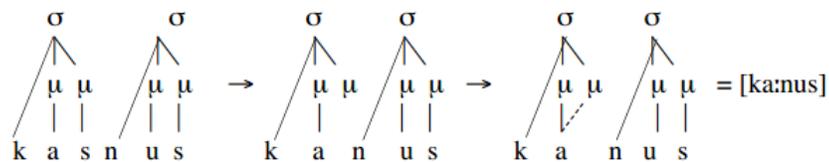
§5: Wrap-up

N.B: Due to time limitations, I will only discuss a selection of CL-approaches, with a focus on recent literature. For an overview of some of the missing accounts, see Gess (2011). For an overview of CL in GP, see Beltzung (2008) and the literature mentioned there.

1. CL: Intro & Hayes (1989)

- ⊙ CL = the lengthening of a segment in compensation for the loss or reduction of another (Gess 2011: 1513)
- ⊙ Prototypically, CVC CL as in Latin /kasnus/ → [ka:nus] ‘gray’
- ⊙ Early analyses, e.g. de Chene and Anderson (1979), Hock (1986), various papers in Wetzels & Sezer (1986)
- ⊙ Hayes (1989): single most influential paper on CL
 - CL as moraic conservation

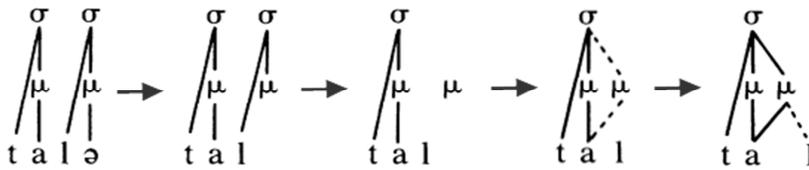
(1) Latin *s*-deletion



- Accounts for various cases of CL, either through V- or C-lengthening
- ⊙ Became the standard theory of CL, but carries along some problems, the major of which are
 - Opacity
 - Onset loss should never produce CL
 - Related to that: only moraic segments should cause CL
 - Parasitic Delinking (Step 3 below) *ad hoc* mechanism to capture CVCV CL



(2) Generating Middle English [talə] → [ta:l] in Hayes' analysis



2. CL is opaque

- ⊙ CL à la Hayes requires the order
 - RULE 1. WBYP
 - RULE 2 CL-triggering segment loss
 - OUTCOME CL
- ⊙ Counter-bleeding opacity: if R2 applied first, it'd remove the environment for R1 application
- ⊙ Problem: Standard Parallel OT (POT; Prince & Smolensky 1993/2004) permits no intermediate stage for WBYP to apply, thus under the assumption that high-ranking MAX-μ ensures mora preservation
 - /kaⁿn/ → [ka^m] with CL, but /kaⁿn/ → *[ka^m] without, instead [kaⁿ] should be preferred, because it avoids a DEP-μ violation

(3) a. CL in POT --- moraic input

	/CV ⁿ C ⁿ /	NoCODA	MAX-μ	MAX-SEG	DEP-μ
a.	CV ⁿ		*!	*	
b.	CV ^m			*	

b. CL in POT --- non-moraic input

	/CV ⁿ C/	NoCODA	MAX-μ	MAX-SEG	DEP-μ
a.	CV ⁿ			*	
b.	CV ^m			*	*!

- ⊙ Recent CL approaches in OT
 - Topintzi (2006b, 2010): CL as position correspondence in Parallel OT (POT)
 - Kiparsky (2011): CL in Stratal OT
 - Samko (2011), Torres-Tamarit (2012): CL in Harmonic Serialism
- ⊙ Other CL approaches in OT [cf. Topintzi 2006a, §5.5.2 for some more details and criticism]
 - Lee (1996): preservation of the numerical integrity of segments
 - Sprouse (1997): explicitly utilizes intermediate stages, by means of enriched inputs, i.e. inputs which exhibit syllabification and moraification
 - Goldrick (2001): proposes Projection (licensing) and Pronunciation constraints and allows for the possibility that structure, e.g. a mora, projected by one segment, may be pronounced on a different segment
 - Hermans (2001): sees CL as segment preservation (not mora preservation) and segments need to have correspondents – roughly – of the same sonority type



- Sumner (2003): CL as segmental preservation through coalescence, a by-product of which is vowel or consonant lengthening
- Shaw (2007): in OT-CC; will not be reviewed as OT-CC seems to have been superseded by HS

2.1. CL as position correspondence in POT

- ⊙ Core Idea: Preservation of a position by means of a root node (segmental) or by means of a mora (prosodic)
 - Position Correspondence (POSCORR)
- (4) POSCORR: An input segment must have an output correspondent either segmentally by means of a root node or prosodically by means of a mora
- ⊙ Other constraints utilized for CVC-CL (and lack thereof)
 - M(arkedness), e.g. NoCODA
 - UNIFORMITY, to avoid coalescence
 - P-DEP- μ : a version of DEP- μ that punishes V-lengthening (but not underlying V-length), thus $V^{\mu}C^{\mu} \rightarrow V^{\mu\mu}$ violates P-DEP- μ but not DEP- μ
- (5) *General schema*
 - (i) CL after segment deletion, e.g. /kan/ \rightarrow [ka:], /kan+i/ \rightarrow [kani]
M, UNIFORMITY, POSCORR \gg P-DEP- μ
 - (ii) Lack of CL after segment deletion
 - (a) No trace of deleted position, e.g. /kan/ \rightarrow [ka], /kan+i/ \rightarrow [kani]
M, UNIFORMITY, P-DEP- μ \gg POSCORR
 - (b) Segmental correspondent, e.g. /kan/ \rightarrow [kã], /kan+i/ \rightarrow [kani]
M, P-DEP- μ , POSCORR \gg UNIFORMITY
- ⊙ Thus: CL is not about moraic preservation, but about position preservation through a mora

- (6) *Deletion of segment followed by CL: POSCORR \gg P-DEP- μ*

non-moraic input or moraic input produce the same output, i.e. $V^{\mu\mu}$

	/C ₁ V ₂ ^μ C ₃ / or /C ₁ V ₂ ^μ C ₃ ^μ /	POSCORR	P-DEP- μ
a.	V ₂ ^{μ2}	*!	
b.	V ₂ ^{μ2μ3}		*

- (7) *Deletion of segment not followed by CL: P-DEP- μ \gg POSCORR*

non-moraic input or moraic input produce the same output, i.e. V^{μ}

	/C ₁ V ₂ ^μ C ₃ / or /C ₁ V ₂ ^μ C ₃ ^μ /	P-DEP- μ	POSCORR
a.	V ₂ ^{μ2}		*
b.	V ₂ ^{μ2μ3}	*!	

- Fully parallelistic, minimal addition of machinery (P-DEP- μ is used elsewhere, cf. Bermúdez-Otero 2001; Campos-Astorkiza 2004)



- Captures a fairly extensive range of CVC CL data, including CL after: coda loss, onset loss, non-moraic segment loss (see below)
- Underlyingly weight-bearing segments, e.g. vowels, should always lead to CL (WBYP is irrelevant here), cf. apocope in Lardil, e.g. *yalulu* > *yalul* **yalu:l*, **ya:lul*, unless something stops them from doing so → unclear what that would be in Hayes (1989). Problem is resolved here through constraint ranking (cf. (5) & fn.1)
- Kiparsky (2011: 59) claims that it does not account for CL after glide formation as in Kihehe (Odden & Odden 1999) /u-ítite/ → [wiítite] ‘you spilled’, /tu-ítite/ → [twiítite] ‘we spilled’ or prenasalisation as in Luganda /ba-ntu/ → [ba:ⁿtu] ‘people’
- But prenasalisation is accounted for (Topintzi 2010: 125) and so is glide formation (2010: 123-4) at least of the type in Kinyarwaanda /ku-gu-ir-a/ → [kug^wi:ra] ‘to fall on’ where the glide acts as secondary articulation and thus do not fulfill the RtNd criterion of POSCORR¹

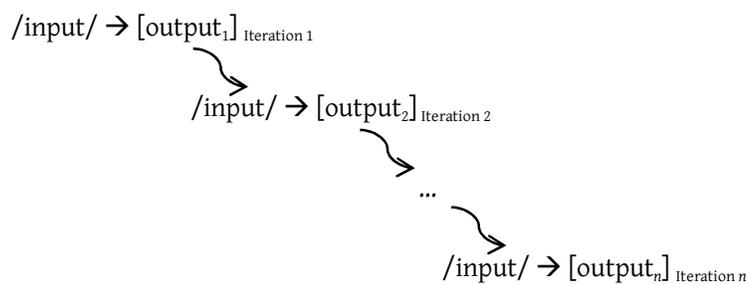
2.2. CL in some version of derivational OT

Stratal OT CL-account (Kiparsky 2011)

- ⊙ allows strata/levels, that are morphologically related, i.e. stem-level, word-level
- ⊙ in each pass, a different ranking may be in order
- ⊙ for CL, the suggestion is that the unsyllabified and moraically specified input (following RotB) goes through the first, stem level, pass and gets to be syllabified according to the language’s constraint ranking. Word-level and post-lexical phonology apply afterwards, but CL can apply exactly because there has been syllabification at the stem level

Harmonic Serialism (McCarthy 2008a,b, 2010)

- ⊙ A recent, serial version of OT, but not based on strata
- ⊙ Like POT, it involves a single constraint ranking throughout BUT not a single /input/ → [output] mapping. Instead, it comprises iterations of the type



Convergence: final output same as most recent input

- ⊙ GEN is more constrained than in POT → gradualness, thus for /lab/ [labi] or [lat] ✓, but [lati] ✗
- ⊙ Iterations/derivations are harmonically improving

¹ It seems to me that the data Kiparsky cites as problematic, i.e. Kihehe or Pali /soahaj/ → [sɔa:haj] ‘that I’, can also be accounted for even if the glide/approximant sound forms a root node of its own. Making the reasonable assumption that vowels bear a mora underlyingly, then under *HIATUS, MAX-μ >> POSCORR, P-DEP-μ, [wiitite] will be preferred over [witite] or [uitite]. Lack of lengthening after glide formation (schematically: witite), cf. Kikerewe, can also be explained under *HIATUS >> POSCORR, P-DEP-μ >> MAX-μ. In this view however, only Kinyarwaanda exhibits real CL (position preservation through a mora). Kihehe and Kikerewe satisfy POSCORR by means of a RtNd. The fact that there is also lengthening (or lack thereof) is due to MAX-μ and not POSCORR per se.



CL in HS

- ⊙ Samko (2011), Torres-Tamarit (2012)
- ⊙ Basic intuition to account for CVC CL is that the coda C must remain long enough in the derivation to acquire a mora and then delete.
- ⊙ Torres-Tamarit (after McCarthy 2008b): CVC CL = two-step process, i.e. (i) debuccalisation → (ii) root node deletion [Basic ranking for CL: WBYP >> CODACOND >> MAX; debuccalisation: CODACOND >> HAVE-PLACE, MAX-PLACE >> MAX; re-association of floating μ: MAX, *FLOAT-μ >> DEP-LINK]
 - Thus /CVC/ → [CV^uH^u] → [CV^{uM}] → [CV^{uH}] (N.B: M=floating mora)
 - Crucially: during the first iteration, a mapping of /CVC/ → [CV] is not produced by GEN, because of gradualness; otherwise [CV] would be chosen over [CVC] given the CL ranking
 - BUT: [CV] can be generated as a candidate under certain conditions in a single step, cf. a CVC input with a placeless coda, e.g. /CVh/
 - Prediction: no CL after coda loss in /CVh/. Instead the mapping [CV] should be preferred
 - Problem! CL after laryngeal coda loss in Farsi, Turkish, Choctaw, Ket, Bella Coola, etc.
- ⊙ Illustration of problem (Samko 2011: tableau (64); additional issues discussed in §5.3)

	μ t a H s i l	WBYP	hCOND	MAX[Place]	*μ/C	*SHARE	DEP-L[μ]	MAX
a.	μ t a H s i l	*!	*					
b.	μ μ μ μ t a H s i l		*!		*			
c.	μ μ μ t a s i l							*
d.	μ μ μ μ t a H s i l		*!			*	*	

- ⊙ Samko's solution (2011) in getting a moraic coda in /CVC/ is to utilize the *Fully Faithful Candidate*
 - FFCs (after OT-CC) are candidates that violate no basic faithfulness constraints, i.e. they are morafied and syllabified versions of the inputs
 - Arriving at a FFC is a separate, constraint-based, process that happens prior to the first iteration (2011: 29)
 - For an input /CVC/ and possible FFCs [CV^u] an [CV^uC^u], the latter wins, given the ranking WBYP >> MAX >> NOCODA, P-DEP-μ [N.B: Why can't CV^uC be an FFC?]
- ⊙ Derivation looks like: /CV^uC^u/ → [CV^uC^m] → [CV^{uH}] [N.B: ^m stands for μ shared between V&C]
- ⊙ Implication of HS: intermediate representations should be valid surface forms in some L
 - [CV^uH^u] in Torres-Tamarit might end up as [h] or [ʔ] for voiceless stops and fricatives, assuming these are placeless, but what about a placeless lateral or a placeless voiced stop?
 - [CV^uC^m] in Samko may emerge in some Ls (Malayalam, Levantine Arabic in Broselow et al. 1997), but also suggests that in a L where only a subset of codas undergo CL, the rest should have mora sharing, i.e. closed syllable lengthening. Questionable but testable



3. Unexpected CL

- ⊙ De Chene & Anderson (1979): a language may exhibit CL if it independently possesses a V-length contrast → counter-examples: Occitan, Friulian, Dinka (on ternary V-length contrast), Komi Ižma, etc. (Kavitskaya 2002: 23-4 and refs therein/henceforth K)
- ⊙ Modification by Hayes (1989): only weightful segments may cause CL → counter-examples
- ⊙ CL with onset loss: Samothraki Greek (Kavitskaya 2002, Topintzi 2006b, 2010), many Ls mentioned in Beltzung (2008) and below
- ⊙ CL with weightless segments (esp. non-moraic codas): Ngajan Dyirbal, Piro (Kavitskaya 2002), Squamish (Beltzung 2008), Lilloet, Nyawaygi, Hullaga Quechua, Telugu, Lhasa Tibetan, Thai (the latter two have sonorant moraic codas, but obstruent codas are non-moraic; Gordon 2006 cited in Yun 2010)

(8) *Dialects of Dyirbal (K 25 after Dixon 1990)*

	<i>Mamu</i>	<i>Ngajan</i>	
a.	waynyji-	wa:nyji-	'to go up'
b.	marbu	ma:bu	'louse'
	ɲamir	ɲami:	'hungry'
c.	gulgu	gu:gu	'brought together'
	bulal	bula:	'firefly'

(9) *Samothraki Greek (Topintzi 2010 after Katsanis 1996)*

- a. r-loss initially and CL

ríɣaɲ > í:ɣaɲ	'oregano'
rúxa > ú:xa	'clothes'
- b. r-loss in onset cluster and CL

krató > ka:tó	'I hold'
ɣráfó > ɣá:fu	'I write'
- c. onset r-loss word-medially and no CL

θaró > θaó	'I reckon'
léftirus > léftius	'free'
- d. no r-loss in coda

fanári > fanár	'lantern'
karpós > karpós	'seed'

- ⊙ No evidence for syllable weight distinctions in Ngajan or Sam.Greek (no independent V-length, QI-stress, no WdMin, etc.)
- ⊙ These languages are (partially) taken up in Kavitskaya (2002), Topintzi (2006a,b, 2010), Beltzung (2008), Yun (2010) and Kiparsky (2011)
- ⊙ Kavitskaya's (2002) CL treatment examines the phenomenon from a diachronic point of view. It's not a moraic conservation approach; instead it talks about phonologization of vowel length



(10) CVC CL (K: 9)

	a. Stage 1 (before consonant loss)	b. Stage 2 (consonant loss)	c. Phonologization
CVX	C V C []	C V []	CV:
CVY	C V C []	C V []	CV

where: X = C with longer vocalic transitions, e.g. glide; Y = stop

- ⊙ K's advantage over moraic theory à la Hayes (1989): since for her CL is not about moraic preservation, she sidesteps the onset (weightless segment) problem
- ⊙ K 99 on Sam.Greek: "I propose that the solution of the problem of CL through loss of onset *r* is connected with the acoustics of rhotics. *r* is vocalic enough to be reinterpreted as additional vowel length, or to exert lengthening effect on the following vowel and to be reinterpreted as a part of the vowel. It has the same properties when it is in the coda or in the onset, where it does not have to be moraic. Thus, *r* can be reinterpreted as additional length on preceding and following vowels"
- ⊙ Problems for the SamGreek analysis:
 - Why does the lengthening effect in Sam.Greek not appear in coda position too (the prototypical position for CL)? Compare with (9d)
 - This analysis can only work if *r* has an approximant-like nature. This is at best dubious; articulatory and acoustic study on Standard Greek *r* in *Cr* clusters highlights high variability in its realization regarding the degree of constriction (Nicolaidis & Baltazani 2011). While a different realization of *r* depending on its syllabic position might save the day, e.g. more approximant-like as an onset, more tap-like in a coda, this is not the case (Mary Baltazani, p.c.), since roughly the same variability appears irrespective of position
- ⊙ Kiparsky (2011) in a Stratal framework, claims that prevocalic /*r*/ shifted to the nucleus (due to high sonority, it could not be syllabified as an onset) and became moraic forming a diphthong with the following vowel. /*r*/ subsequently deleted producing CL
 - No motivation for the /*r*/ loss is given once it acquires diphthongal status
 - Even if this story can be justified, it fails to extend to other cases of onset CL, involving deletion of highly non-sonorous Cs
- ⊙ Beltzung (2008 and refs therein) reports: Lango, Turkana, Gyore, Anywa and Ntcham, e.g. /ku-saku/ → [kusa:u] 'field' (Ntcham); tùy-ḁ > tù:ḁ 'germination' (Anywa)
- ⊙ Kavitskaya (2002) also cannot account for these data, because
 - some involve consonants that cannot be rendered vocalic enough
 - the theory (K:68) predicts that intervocalic consonant loss should not produce lengthening, because the vocalic transitions would already be too long, so they could not be re-interpreted as V:V → Ntcham or Anywa contradict that
- ⊙ Topintzi (2006b, 2010) avoids all these issues: assuming that Sam.Greek /*r*/ cannot survive because it makes a bad onset (a claim Kiparsky actually corroborates, 2011: 60), it deletes; due to POSCORR >> P-DEP-μ, a mora is inserted to the neighbouring vowel producing lengthening. A similar story (with a different markedness constraint) would apply to the other languages



- ⊙ The difference between lengthening intervocalically (cf. Ntcham) vs. not lengthening Sam.Greek (cf. 9c), is attributed to a violable constraint, *SL-VH (*SUPER LONG VOCALIC HIATUS)
- ⊙ HS analyses of CL have (so far) not tackled CL from onset loss, but Samko (2011: 44) points to an analysis along the lines of Kiparsky (2011), since she assumes that only moraic segments can cause CL and onsets can never be such → same problem as Kiparsky above
- ⊙ CL after loss of a weightless coda (cf. Ngajan) is similarly problematic for the mora preservation analyses of CL in Stratal OT and HS, but not under POSCORR (or K for that matter)
- ⊙ A final note: Topintzi (2008, 2010) argues for the analysis of some geminates as moraic onsets. This means that onsets can also function as *targets* of CL

(11) *Pattani Malay* (Topintzi 2010: 169 and refs therein)

a. CL

buwi ~ wi 'give' sidadu ~ d:adu 'police' pimatə ~ m:atə 'jewellery'

b. Stress (slightly simplified for current purposes)

Words lacking geminates

Words with initial geminates

jàlé 'road, path'

m:átə 'jewellery'

dàlé 'in, deep'

ɟ:álɛ 'to walk'

màkèngé 'food'

4. On asymmetries and less discussed issues in CL

- ⊙ A typology of CL (adapted after Gess 2011)

Type	Trigger	Relationship	Target	Frequent?	Example
1A	C	Regressive	V	YES	blázmer > bla:mer (Old French)
1B	C	Progressive	V	SOMEWHAT	ruxa > u:xa (SamG)
1C	V	Regressive	V	YES	wizi > vi:z (Hungarian)
1D	V	Progressive	V	SOMEWHAT	kasapan > ksa:pan (Macuxi)
2A	C	Progressive	C	SOMEWHAT	wesfi > wessi (Semitic)
2B	C	Regressive	C	YES	borʃa > boʃʃa (Bengali)
2C	V	Regressive	C	VERY RARE	balite > bal:te (Bulgarian)
2D	V	Progressive	C	SOMEWHAT	likubo > kkubo (Ganda)

Diachrony vs. synchrony?

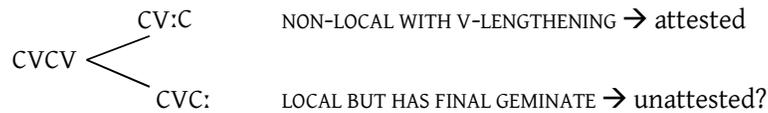
- ⊙ This table merges together cases of diachronic and synchronic CL (thanks to Jochen Trommer for pointing this out)
- ⊙ An observation made by Gess himself (but not applied in practice):
 - we may need different analyses for synchronic and diachronic CL
 - the latter may be phonetically motivated, cf. Kavitskaya's (2002: 10) analysis of CVCV CL



(12) CVCV CL (K: 10)

	a. Stage 1 (before vowel loss)	b. Stage 2 (loss of V ₂)	c. Phonologization
CVCV	C V C V □	C V C □	CV:C
CVC	C V C □	C V C □	CVC

- Indeed, 1C is virtually always a diachronic process. Notice that if this pattern is about mora/position preservation, we would also expect local CL



- CVC: should be preferable, if a language allows final geminates, cf. Hungarian which allows them, and yet: wizi > vi:z, *viz: (but perhaps no final geminates at that stage?)
 - If this pattern is about phonologization of V-length, then final gemination is correctly excluded → perhaps Kavitskaya is right
- But Gess reports also synchronic CVCV CL in Yapese (Jensen 1977) and possibly in Northern Greek dialects (listen to [ksefi:j] < [ksefji] in this commercial <http://www.youtube.com/watch?v=8mian4jTFcY>, sec.29-30; thanks to Mary Baltazani for pointing out this example)
- Additional differences between CVCV and CVC CL (K 151-169)
 - CVCV = trigger usually non-recoverable vs. CVC = trigger recoverable through alternations
 - CVCV → CV:C non-optimizing (unless there is moraic conservation on the foot level) vs. CVC → CV: optimizing by getting rid of a coda
- The ‘double flop’ instance (a variant of 1A) is almost always illustrated through Ancient Greek *odwos* > *o:dos* (Wetzels 1986: 310), a diachronic type of CL. Yun (2010: 8) also mentions Akkadian, Coptic, Trio (Western), Fula, Supyire, Kasem and Old English
- 2A is also reported to be diachronic; phonologically-speaking it is odd, since it is the onset that assimilates to the coda and not the other way round; McCarthy (2008b) excludes this possibility, unless the onset is placeless (could account for Hebrew: *jilkædenhu: > jilkædennu:) or the assimilation is one of manner or voicing → may capture some of Gess’ examples from Semitic (2011: 1517, citing Lipiński 2001: 195), but not all, cf. *wæsfı > wessi (Gurage)
- Conclusion: most analyses treat synchronic and diachronic types of CL on a par, but this may be misguided as they may be subject to different motivations → we need a CL database that addresses this point in order to reach more secure generalizations about the phenomenon

Locality

- With the exception of 1C (see above) and 1D (cf. Macuxi – possibility that it is metrically driven), CL is basically affecting neighbouring segments. Locality is usually assumed in OT analyses of CL, but not argued for explicitly



- In serial theories, this is attributed to a ban on line crossing; structures that violate it, e.g. CVCV CL only do so superficially, cf. (2) and parasitic delinking
- Use of association line crossing as a constraint: Possible in Stratal OT or HS, where syllabification has occurred at some level, but not necessarily in POT
- Some version of theory-neutral LOCAL (cf. use of constraint in Yip 2002: 84 on tone) that requires that moras should remain linked to their sponsors (either as correspondents to deleted segments, cf. POSCORR, or as correspondents to moras assigned at some level in the derivation, cf. Stratal OT or HS)
 - Gradient evaluation of LOCAL- μ^2
 - Illustration (focus on the mora of *n*)
 - [li.kan^μ.ta] → [li.kan^μ.ta] or [li.kat^μ.ta]³ No shift of the mora, so LOCAL is satisfied
 - [li.kan^μ.ta] → [li.ka^μ.ta] Shift by 1 segment, single violation of LOCAL
 - [li.kan^μ.ta] → [li^μ.ka.ta], [lik^μ.ka.ta], [li.ka.ta^μ] Shift by more segments, multiple violations of LOCAL- μ
 - Claim: CL will always be LOCAL unless overridden by a requirement that cannot be satisfied locally, e.g. if there are Ls with CVCV CL → CV:C requiring a synchronic account, it is predicted that the L in question should lack final geminates, so that locality is forced to be violated

Directionality

- Rightward CL is preferred over leftward. Another, rather, understudied area in CL
- Yun (2010) tackles it to some extent; she adopts x-slots rather than moras and argues that CL is driven by perceptual reasons. Faithfulness constraints demand the preservation of duration of sequences involving CL-triggers (2010: 40) and builds in directly into the model some of these preferences by means of fixed rankings (13) and variable position of IDENT-LENGTH

(13) Universal rankings in Yun (2010)

- a. MAX-IO[X]_{VC} >> MAX-IO[X]_{CV}
 - b. MAX-IO[X]_{V//[+son]} >> MAX-IO[X]_{V//[−son]}
 - c. MAX-IO[X]_{VC} >> MAX-IO[X]_{VCC}
- (13a) gives preference to rightward CL after C-loss (thus 1A > 1B) → no equivalent in μ -theory
 - No discussion of CL with gemination or CVCV CL (the latter might be a good thing?)
 - Main objection: if there are no moras, then a host of other things become difficult to explain, e.g. CVC's weight on a language specific basis, bimoraic templates of CVCV, CVC or CV: etc.
 - But, captures nicely a very strong preference (universal), i.e. V₁CV₂ → V₁:V₂ and not V₁V₂:
 - Only exception: Turkana /e+wɔrɔ/ → [eɔ:rrɔ] 'cloth' (Turkana; Dimmendaal 1983) → ban against long prefixal Vs

² An informal, i.e. one that has not been rigorously tested, definition of LOCAL- μ would be: Given a segment τ_x and a mora μ_x co-indexed with it, count one violation for every segment (on either side of τ_x) that μ_x migrates to.

³ The mora of the geminate in [li.kat.ta] treated either as preservation of a mora at an earlier stage in the same position or as a mora added to preserve a position, trivially satisfies LOCAL by virtue of remaining in the same (syllabic) position in the string. This might capture the intuition of some that 2B is not real CL, but standard assimilation.



(14) Deriving V:V vs. VV: asymmetry in Yun (2010: 51)

a. MAX-[X]_{VC} >> IDENT-IO [V-length] >> MAX-[X]_{CV}

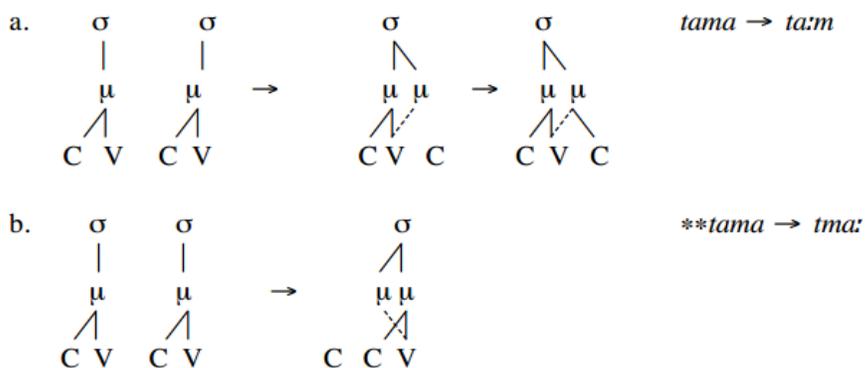
	VCV	Markedness	MAX-[X] _{VC}	ID[V-LENGTH]	MAX-[X] _{CV}	MAX-SEG
a.	VCV	*!				
b.	V:V			*	*	*
c.	VV:		*!	*		*

b. MAX-[X]_{VC} >> MAX-[X]_{CV} >> IDENT-IO [V-length]

	VCV	Markedness	MAX-[X] _{VC}	MAX-[X] _{CV}	ID[V-LENGTH]	MAX-SEG
a.	VCV	*!				
b.	V:V			*	*	*
c.	VV:		*!		*	*

- ⊙ Moraic theory only partially looks into the issue, accounting for more peripheral cases of CL and relying heavily on representations and parasitic delinking

(15) Preference for rightward rather than leftward CL in /CVCV/ (cf. Kavitskaya 2002: 33)



- ⊙ Also, given Gess' table, V-lengthening is preferred over C-lengthening. Is this a universal?
 - We'd need to check whether CL-undergoing Ls allow for both long Vs and geminates and yet prefer one or the other for CL
 - Are there Ls with CL, but with no weight contrast, e.g. Ngajan, that resolve CL through gemination rather than V-lengthening?
- ⊙ Assuming V: is better than C:, then some cases become more difficult to explain, e.g. balite > bal:te *ba:l:te (Bulgarian⁴) or likubo > k:ubo *ku:bo (Ganda)
- ⊙ Can it be that LOCAL-μ prefers CL with gemination?

⁴ Intuitively, CVC:CV is a marked structure – perhaps explaining its rarity as an output of CL – since it admits a cluster of geminate+onset C. In moraic theory (Hayes 1989 and many following him), a geminate is an *underlyingly* moraic C and as such formally equivalent to CVC^h.CV (with a singleton coda). However, the geminate here is derived and not necessarily identical to an underlying one (cf. Ringen & Vago 2011: 156, fn.3). Derived CVC:CV may thus be structurally different from CVC.CV, and may require a more complex structure, e.g. along the lines of a composite model for geminates (cf. Curtis 2003 for discussion) that incorporates Ringen & Vago's (2011) idea that geminates are segmentally bipositional.



- But this never seems to favour gemination in CVCV# CL, cf. Hungarian
 - As said, this CL is typically diachronic, so perhaps subject to different mechanisms:
 - How can this difference be formally captured?
 - There still remains a residue of synchronic CVCV# CL, e.g. Yapese
 - Maybe final gemination works differently than non-final gemination?
- ⊙ Language specific considerations need to be taken into account, e.g. Ilokano (Hayes 1989) has optional CL after glide formation by means of gemination although local V-lengthening would be possible, i.e. /luto+en/ → [lut.twén], but *[lut.wé:n] ‘cook-goal focus’
- ⊙ Observation: long Vs in Ilokano only appear in open syllables (in prefixes under reduplication or in a penultimate stressed syllable)
- ⊙ Gemination preferred because [lut.we:n] either violates *3 μ or would produce a long suffixal e at the stem level where long Vs are not allowed (Jochen Trommer & Eva Zimmermann, p.c)

5. Wrap-Up

- ⊙ Since Hayes (1989), there has been a lot of progress in CL and more in-depth discussion of its opaque character, as well as the empirical domain it applies to
- ⊙ The opacity problem
 - sidestepped in POT, cf. POSCORR in Topintzi (2006b, 2010) or string-based constraints in Yun (2010)
 - tackled more directly in Stratal OT (Kiparsky 2011) and HS (Samko 2011, Torres-Tamarit 2012) → allowing intermediate stages again, but this time in a more constrained manner, either by being morphologically-informed or limited by gradualness and harmonic improvement
- ⊙ The empirical domain of CL
 - is wider than assumed: onsets may act as triggers and targets of CL; weightless codas may trigger CL
 - is not deterministic: weightful vowels do not always induce CL (cf. Lardil apocope)
- ⊙ But there is still lots to understand
 - The boundaries of CL, although wider in some respects, seem to be more confined than usually assumed: should final V-loss and prefinal V-lengthening in /CVCV/ be considered CL, is assimilation by gemination (cf. Bengali) a type of CL and if so, how is it to be distinguished from gemination in Ilokano?
 - How big is the role of diachrony in CL and how can it be separated from that of synchronic CL in a principled way?
 - It looks that each of the mora-, segment- and string-based accounts of CL has its own virtues. Can they be reconciled in a uniform way?
 - More research into locality and directionality is needed



REFERENCES

- Beltzung, J.-M. 2008. *L'allongement compensatoire dans les représentations phonologiques: nature, contraintes et typologie*. Doctoral Dissertation, Université Paris 3 (Sorbonne-Nouvelle).
- Bermudez-Otero, R. 2001. Underlyingly nonmoraic consonants, faithfulness, and sympathy. ms, University of Manchester.
- Broselow, E., S-I Chen & M. Huffman. 1997. Syllable Weight: Convergence of Phonology and Phonetics. *Phonology* **14**: 47-82.
- Campos-Astorkiza, R. 2004. Faith in Moras: A Revised Approach to Prosodic Faithfulness. In *Proceedings of NELS 34*. ed. by K. Moulton & M. Wolf.
- Dechene, B. & S. R. Anderson. 1979. Compensatory Lengthening. *Language* **55**: 505-535.
- Dimmendaal, G. J. 1983. *The Turkana Language*. Dordrecht: Foris.
- Gess, R. 2011. Compensatory Lengthening. In *The Blackwell Companion to Phonology* (vol. 3), ed. by M. van Oostendorp, C. Ewen, E. Hume & K. Rice, pp. 1513-1536. Malden, MA: Wiley-Blackwell.
- Goldrick, M. 2001. Turbid output representations and the unity of opacity. In *Proceedings of NELS 30*. ed. by M. Hirotani, A. Coetzee, N. Hall & J.-Y. Kim.
- Gordon, M. 2006. *Syllable Weight: Phonetics, Phonology, Typology*. New York: Routledge.
- Hayes, B. 1989. Compensatory Lengthening in Moraic Phonology. *Linguistic Inquiry* **20**: 253-306.
- Hermans, B. 2001. CL in OT. *Linguistics in the Netherlands* **18**: 139-150.
- Hock, H. H. 1986. Compensatory Lengthening: In Defense of the Concept 'Mora'. *Folia Linguistica* **20**: 431-460.
- Jensen, J. T. 1977. *Yapese reference grammar*. Honolulu: University Press of Hawaii.
- Katsanis, N. [Κατσάνης, Ν.]. 1996. *The Dialect of Samothraki Greek [Το Γλωσσικό Ιδίωμα της Σαμοθράκης] Municipality of Samothraki [Δήμος Σαμοθράκης]*.
- Kavitskaya, D. 2002. *Compensatory Lengthening: Phonetics, Phonology, Diachrony*. New York: Routledge.
- Kiparsky, P. 2011. Compensatory lengthening. In *Handbook of Syllable*, ed. by C.E. Cairns & E. Raimy, pp. 33-69. Leiden & Boston: Brill.
- Lee, B.-G. 1996. The Emergence of the Faithful. ms., Available as ROA-166.
- Lipíński, E. 2001. *Semitic languages: Outline of a comparative grammar*. 2nd edn. Leuven: Peeters.
- McCarthy, J. J. 2008a. The serial interaction of stress and syncope. *NLLT* **26**: 499-546.
- . 2008b. The gradual path to cluster simplification. *Phonology* **25**: 271-319.
- . 2010. An introduction to Harmonic Serialism. *Language and Linguistics Compass* **4**: 1010-1018.
- Nicolaidis, K. & M. Baltazani. 2011. An electropalatographic and acoustic study of the Greek rhotic in /Cr/ clusters. In *Proceedings of the 17th ICPHS*, Hong Kong, China, pp. 1474-1477.
- Odden, D. & M. Odden. 1999. Kihehe syllable structure. In *The syllable: views and facts*, ed. by H. van der Hulst & N. Ritter, pp. 417-446. Berlin: Mouton de Gruyter.
- Prince, A. & P. Smolensky. 2004. *Optimality Theory: Constraint interaction in generative grammar* Malden, MA & Oxford: Blackwell.
- Ringen, C.O. & R. Vago. 2011. Geminates: Heavy or Long? In *Handbook of the Syllable*, ed. by C.E. Cairns & E. Raimy, pp. 155-169. Leiden & Boston: Brill.
- Samko, B. 2011. Compensatory Lengthening in Harmonic Serialism. Ms, UC Santa Cruz.
- Shaw, J. 2007. Compensatory lengthening via mora preservation in OT-CC: theory and predictions. In *Proceedings of NELS 38*.
- Sprouse, R. 1997. A Case for Enriched Inputs. *Handout for paper presented at TREND VII*, held 3 May 1997 at UCSC. Available as ROA-193.
- Sumner, M. 1999. Compensatory Lengthening as Coalescence: Analysis and Implications. In *Proceedings of WCCFL 18*, ed. by S. Bird, A. Carnie, J. Haugen and P. Norquest.
- Topintzi, N. 2006a. *Moraic Onsets*. Doctoral Dissertation, UCL.



- . 2006b. A (Not So) Paradoxical Instance of Compensatory Lengthening: Samothraki Greek and Theoretical Implications. *Journal of Greek Linguistics* 7: 71-119.
- . 2008. On the existence of moraic onset geminates. *NLLT* 26: 147-184.
- . 2010. *Onsets: Suprasegmental and Prosodic Behaviour*. Cambridge: Cambridge University Press.
- Torres-Tamarit, F. 2012. Compensatory Lengthening and Opaque Geminata in Harmonic Serialism. Paper presented at OCP 9, Berlin, 18-21 January 2012.
- Wetzels, W. L. 1986. Phonological timing in Ancient Greek. In *Studies in compensatory lengthening*, ed. by W. L. Wetzels & E. Sezer, pp. 279 -344. Dordrecht: Foris.
- Wetzels, W. L. & E. Sezer (eds.) 1986. *Studies in compensatory lengthening*. Dordrecht: Foris.
- Yip, M. 2002. *Tone*: Cambridge: Cambridge University Press.
- Yun, S. 2010. Implicational Universals in Compensatory Lengthening Typology: A Phonetically-based Optimality Theoretic Approach. ms, MIT.

