Kevin M. Ryan
Author's accepted manuscript (June 2026)
To appear in *Phonology*

Geminate licensing in Tamil and elsewhere: ganging cumulativity, not constraint conjunction

Abstract

Two strategies have been proposed to model gang effects, namely, constraint conjunction and cumulativity via weighting. While arguments against conjunction have focused on its overgeneration, I demonstrate that it also undergenerates. For example, within the Tamil syllable, a long vowel may precede an obstruent geminate but not a sonorant geminate. Ruling out a purely weight-based analysis, I conclude that (C)V:(C)N: syllables are excluded because they are doubly marked, violating both *SONGEM and *3μ. The joint effect of these constraints can be modeled by cumulativity in Harmonic Grammar. It cannot, however, be modeled using local conjunction in Optimality Theory, as no domain of conjunction captures the correct generalization. Thus, Optimality Theory is not a viable theory of (even categorical) phonotactics. Further, cumulativity can address at least some apparent moraic mismatches. Finally, the article offers a new argument against the MPARSE approach to phonotactics.

In constraint-based grammar, a GANG EFFECT refers to the situation in which two (or more) constraints overpower a higher-weighted constraint (Jäger & Rosenbach 2006:939f). An oft-cited example since Pater (2009, 2016) is geminate devoicing in Japanese loanword phonology (Nishimura 2003, 2006; Kawahara 2006, 2011). In brief, voiced geminates are generally permitted (e.g. [webbu] 'web'); thus, IDENT(voice) > *VOIGEM. Moreover, multiple voiced obstruents are generally permitted (e.g. [bagu] 'bug'); thus, IDENT(voi) > OCP(voi). However, *VOIGEM and OCP(voi) cannot both be violated by the same word (e.g. *[doggu] 'dog'). In such forms, the geminate devoices ([dokku]).[1]

In terms of weighted constraints in Harmonic Grammar (HG; Legendre et al. 1990; Smolensky & Legendre 2006), while the two markedness constraints *VOIGEM and OCP(voi) must individually have lower weights than IDENT(voi), their summed weight must exceed that of IDENT(voi). In this way, the two markedness constraints 'gang up' against faithfulness. A weighting scheme that achieves this effect is illustrated by tableaux (1–3).

(1)

	/webbu/	$\mathscr{H}$	IDENT(voi) 3	*VOIGEM 2	OCP(voi) 2
☞	a. webbu	-2		-1	
	b. weppu	-3	-1		

[1] In fact, the repair is optional (*op. cit.*). The optionality is readily implemented in HG without altering the analysis of the gang effect (see Pater 2009:1025 for details).

(2)

	/bagu/	$\mathscr{H}$	IDENT(voi) 3	*VOIGEM 2	OCP(voi) 2
☞	a. bagu	-2			-1
	b. baku	-3	-1		
	c. pagu	-3	-1		

(3)

	/doggu/	$\mathscr{H}$	IDENT(voi) 3	*VOIGEM 2	OCP(voi) 2
	a. doggu	-4		-1	-1
☞	b. dokku	-3	-1		
	c. toggu	-5	-1	-1	

If ranking were strict, as in Optimality Theory (OT; Prince & Smolensky 1993/2004), no ranking of these constraints would generate the desired optima. Nevertheless, as Nishimura (2003, 2006) observes, the case can be analyzed in OT using local constraint conjunction (Smolensky 1993, 2006). I refer to OT with local conjunction as OT-LC, in contrast with HG (without conjunction).[2] In this case, the ganging constraints can be combined into a single, conjoined constraint, namely, *VOIGEM $\&_{PWd}$ OCP(voi), which is violated iff both constraints are violated within the specified domain (prosodic word). The OT-LC analysis is presented in (4–6).

(4)

	/webbu/	*VOIGEM $\&_{PWd}$ OCP(voi)	IDENT(voi)	*VOIGEM	OCP(voi)
☞	a. webbu			$*$	
	b. weppu		$*!$		

(5)

	/bagu/	*VOIGEM $\&_{PWd}$ OCP(voi)	IDENT(voi)	*VOIGEM	OCP(voi)
☞	a. bagu				$*$
	b. baku		$*!$		
	c. pagu		$*!$		

(6)

	/doggu/	*VOIGEM $\&_{PWd}$ OCP(voi)	IDENT(voi)	*VOIGEM	OCP(voi)
	a. doggu	$*!$		$*$	$*$
☞	b. dokku		$*$		
	c. toggu		$*$	$*!$	

Given this case, one might conjecture that *any* gang effect modelable via weighting in HG could be modeled via conjunction in OT. That is, schematically, any three-constraint system with

[2] Conjunction is also compatible with HG, and indeed arguably motivated there by certain types of superlinearity (Green & Davis 2014; Shih 2017; Milenković 2026). That said, for present purposes, I invoke only HG without conjunction. As I will argue, OT-LC is insufficient; HG is needed. Whether LC is needed in HG is an orthogonal question.

the HG weighting conditions in (7) could be analyzed (one might suppose) by the OT grammar in (8). For instance, Pater (2009:1014) notes, 'conjunction of this type can probably account for many, if not all, observed patterns of cumulative constraint interaction'.

(7) $w(\mathrm{X}) > w(\mathrm{Y})$
$w(\mathrm{X}) > w(\mathrm{Z})$
$w(\mathrm{X}) < w(\mathrm{Y}) + w(\mathrm{Z})$

(8) $\mathrm{Y\&Z} \gg \mathrm{X} \gg \mathrm{Y, Z}$

As the present article argues, Pater's hedges are warranted: There are attested gang effects that can be implemented with cumulativity in HG but not with conjunction in OT. The reason that (7) and (8) are not always interchangeable is that the latter must specify a domain of conjunction (McCarthy 2002:18). In some cases, I demonstrate, there is no domain that captures the generalization implemented by (7). With HG ganging, no domain is specified; rather, gang effects are determined by co-relevance (Pater 2016:16–8; see §9.3 below for further discussion). While locality in OT-LC depends on a (stipulated or derived) domain of conjunction, for ganging in HG, locality is emergent, determined by the nature of the constraints involved.[3]

The article proceeds as follows. In order to highlight the theoretical contribution, it is organized backwards, such that the core analysis (and dilemma for OT-LC) is presented first, in the next section (§1), and then the language data are elaborated in the subsequent sections. In particular, §2 overviews Tamil phonotactics and weight. §3 establishes the key phonotactic restriction, showing that not only is V:(C)N: absent from the lexicon, but it is also productively avoided. An alternative, purely weight-based analysis (e.g. *3μ acting alone) is rejected by §4. The empirical facts are revisited by §5, arguing that geminates cannot be reanalyzed (e.g. as fortis/lenis singletons). §6 offers a general analysis of Tamil moraic phonology. Moving beyond Tamil, cases from other languages are considered in §7. Alternative analyses are rejected in §8. The discussion, in §9, focuses on three issues, namely, the novelty of this undergeneration argument vis-à-vis the previous literature (§9.1), the use of gang effects to address cases of moraic inconsistency (§9.2), and the reasons for not using MPARSE as a model of phonotactics here (§9.3). §10 concludes.

1 Analysis of a Tamil phonotactic restriction

Literary Tamil has both sonorant and obstruent geminates, but only the latter can follow a long vowel within the syllable. For instance, while V:T: and VN: are permitted, *V:N: is not.[4] I argue that this restriction reflects the cumulative markedness of two well supported constraints, namely, *SONGEM (no sonorant geminates; Kawahara 2007) and *3μ (no trimoraic syllables; Prince &

[3] Under the original version of conjunction, the domain is stipulated. Łubowicz (2005:4) proposes that markedness constraints can be conjoined only if they have the same loci of violation. I will argue that even the former, anything-goes version of conjunction undergenerates. *A fortiori*, any more restricted version likewise undergenerates. Note further that ganging in HG is not constrained in the same way by co-relevance when gaps are analyzed using MPARSE (Breiss & Albright 2022:7, 25; see §9.3 below). I do not adopt MPARSE here.

[4] Throughout, C refers to a consonant, V a vowel, T an obstruent, N a sonorant (not just nasal), and J a rhotic or glide (i.e. non-lateral non-syllabic approximant). The onset is omitted, being irrelevant. In IPA transcriptions, long vowels are indicated by :, but geminate consonants are doubled.

Smolensky 1993/2004:229). In what follows, assume that all geminates shown in tableaux are moraic (see §2.3, §4, and §5 for further discussion). MAX-μ penalizes deleting a mora and hence shortening an underlyingly long segment.

First, sonorant geminates are generally tolerated, as with [enn̲a] 'what' in (9). Second, superheavy (trimoraic) syllables are generally tolerated, as with [muːkku] 'nose' in (10). However, superheavies are not tolerated if they are induced by geminate sonorants, as with hypothetical (richness-of-the-base) input /eːnn̲a/ in (11). (As will be clarified presently, the awkward phrasing 'are induced by' rather than 'contain' is deliberate.) Such an input must be repaired in some manner, for example, by simplifying the geminate.[5]

(9)

/enn̲a/		$\mathscr{H}$	MAX-μ 3	*SONGEM 2	*3μ 2
☞	a. enn̲a	-2		-1	
	b. en̲a	-3	-1		

(10)

/muːkku/		$\mathscr{H}$	MAX-μ 3	*SONGEM 2	*3μ 2
☞	a. muːkku	-2			-1
	b. muːku	-3	-1		

(11)

/eːnn̲a/		$\mathscr{H}$	MAX-μ 3	*SONGEM 2	*3μ 2
	a. eːnn̲a	-4		-1	-1
☞	b. eːn̲a	-3	-1		
	c. enn̲a	-5	-1	-1	

If the violations of *SONGEM and *3μ come from unrelated loci, the fully faithful candidate wins, as with [inn̲aːʈʈu] 'this position' in (12).[6] Although the violations of *SONGEM and *3μ come from the same syllable in (12)—the nasal is (in part) the onset of that syllable—shortening the geminate does not address the superheaviness problem. Therefore no shortening occurs.

[5] In fact, this repair is attested, both in sandhi (across words) and in historical change (§3). Synchronically within words, the forbidden sequence apparently does not arise (§3), hence my use of a hypothetical input here. The analysis predicts that if such a sequence were to arise within a word, it would be repaired the same way that it is repaired across words.

[6] All examples are from Kamban's corpus (§2). Some additional examples of this pattern are [elloːrkkum] 'for all', [immaːtt̲am] 'this change', [ajɲɲuːtt̲u] 'five hundred', [kaɳɳoːʈʈam] 'impaired vision', and [immuːppiʈajjeː] 'in this old age', among others.

(12)

	/i͟n͟na:ʈʈu/	$\mathscr{H}$	MAX-μ 3	*SONGEM 2	*3μ 2
☞	a. i͟n͟na:ʈʈu	−4		−1	−1
	b. i͟na:ʈʈu	−5	−1		−1
	c. i͟n͟na:ʈu	−5	−1	−1	
	d. i͟na:ʈu	−6	−2		

Turning to the OT-LC analysis of these facts, one could posit *SONGEM &$_\sigma$ *3μ. At first blush, this achieves the desired optima, as in (13–15).

(13)

	/e͟n͟na/	*SONGEM &$_\sigma$ *3μ	MAX-μ	*SONGEM	*3μ
☞	a. e͟n͟na			*	
	b. e͟na		*!		

(14)

	/mu:kku/	*SONGEM &$_\sigma$ *3μ	MAX-μ	*SONGEM	*3μ
☞	a. mu:kku				*
	b. mu:ku		*!		

(15)

	/e:͟n͟na/	*SONGEM &$_\sigma$ *3μ	MAX-μ	*SONGEM	*3μ
	a. e:͟n͟na	*!		*	*
☞	b. e:͟na		*		
	c. e͟n͟na		*	*!	

But this OT-LC analysis erroneously predicts that a sonorant geminate will also simplify as the *onset* of a superheavy, as in (16), in which ☜ indicates the desired winner, while 💣 is the winner generated by the tableau. In Tamil, a sonorant geminate is free to be (in part) the onset of a superheavy; it just cannot be the *cause* of superheaviness.

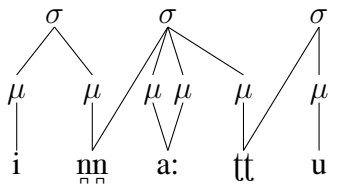
(16)

	/i͟n͟na:ʈʈu/	*SONGEM &$_\sigma$ *3μ	MAX-μ	*SONGEM	*3μ
☜	a. i͟n͟na:ʈʈu	*!		*	*
💣	b. i͟na:ʈʈu		*		*
	c. i͟n͟na:ʈu		*	*!	
	d. i͟na:ʈu		**!		

Given standard assumptions, [i͟n͟na:ʈʈu] has the prosodic structure in (17). Because geminates are heterosyllabic, [n͟n͟] is part of the following syllable, but does not contribute to its weight. One cannot change the definition of locality to count only segments that are *fully* contained within the domain, because that would render the conjoined constraint vacuous in Tamil; all geminates are heterosyllabic. Moreover, no other domain of conjunction would work; the foot is too large, the

mora too small. The rime is not available as the domain, as there is no such constituent in moraic theory (McCarthy & Prince 1986/1996; Hayes 1989; §8). Nor can *SONGEM be replaced by hypothetical *N_μ, penalizing only the sonorant mora (§8). Additional logical possibilities, including invoking non-constituents as domains or adopting a different, non-moraic theory of prosody, are likewise problematic, as discussed in §8, which addresses alternative proposals.

(17)



2 Background: Tamil syllable structure and weight

2.1 The language and inventory

With roughly 80 million native speakers, Tamil (southern India, Sri Lanka, Singapore) is spoken by about 1% of the world's population, the second most widely spoken Dravidian language, after Telugu. Here, I focus on the standard, literary variety, as represented by Kamban's 12th-century (Middle Tamil) epic, the Tamil version of the *Rāmāyaṇa*, one of the two most famous Tamil literary works. Besides its canonicality, this corpus serves as a good point of reference for several reasons: While not modern, it is still somewhat accessible to speakers and conforms to traditional grammars; at the same time, it predates certain complicating sound changes now found in spoken varieties (e.g. once transparent singleton vs. geminate [c, cː] is nowadays usually pronounced [s, tʃ]). Second, it is a large corpus, at 42,132 lines and 251,384 word tokens. Third, it is composed in quantitative meters, which furnishes independent diagnostics of weight. Finally, more sandhi is shown than in modern texts. Standard editions include a 'surface', metricized version, with sandhi, alongside a word-by-word version, without sandhi.[7]

Thus, in what follows, 'Tamil' should be understood to refer to Kamban's Tamil. Some minor points vary for other dialects. At any rate, the key generalization about sonorant vs. obstruent geminates holds for all varieties of literary Tamil, Old (c. 100 CE) to Modern.

Middle Tamil has the consonant and vowel inventories in (18) and (19), respectively.[8] Among consonants, voicing is not contrastive, but length is. Every consonant may geminate except for the rhotics, a well-attested exclusion cross-linguistically (Podesva 2002, Kawahara 2007:30, Ryan 2019:125, Nikolaev 2021). Palatal and velar nasals are uncommon as geminates, but possible.[9] Singleton plosives are usually voiced (and sometimes otherwise lenited) between voiced segments, but as is usual, I will not indicate this in transcriptions; length is the primary distinction (see also §5

[7] I use the Annamalai University edition (Kamban 1956; Hart & Heifetz 1988; Wentworth 2022), as available online at Project Madurai (`www.projectmadurai.org`).

[8] See Christdas (1988), Annamalai & Steever (1998), and Keane (2004) on Modern Tamil and Lehmann (1998:75–99) and Wilden (2018:21) on Old Tamil.

[9] Additionally, one marginal consonant is omitted from (18), namely, the fricative called *āytam* [x]. This phone is rare (0.04% of word tokens) and occurs only between a short vowel and singleton plosive. Because it cannot be an onset, it cannot be a geminate (except due to onomatopoetic overlengthening; Rajam 1992:242).

on the phonetics). Where alveolars contrast with dentals, I transcribe the former with underlines: [t̲, n̲]. This is the standard practice in transliterating Tamil, and also consistent with the IPA, since ̲ is a diacritic for retraction. Phonemic /t̲/ is realized as a trill [r̲] as a singleton, at least intervocalically for Kamban. This allophony does not affect the main argument: What matters is that [tt̲], as an obstruent geminate, is free to follow a long vowel; the realization of /t̲/ as a singleton is irrelevant.[10]

(18)

	labial	dental	alveolar	retroflex	palatal	velar
plosive	p(ː)	t̪(ː)	r̲ t̲ː	ʈ(ː)	c(ː)	k(ː)
nasal	m(ː)	n̪(ː)	n̲(ː)	ɳ(ː)	ɲ(ː)	ŋ(ː)
lateral			l(ː)	ɭ(ː)		
rhotic			ɾ	ɻ		
glide	ʋ(ː)				j(ː)	

Similarly, all phonemic vowels come in short and long pairs. As for diphthongs, Tamil is traditionally said to have two (*ai* and *au*), each with its own letter, presumably influenced by Sanskrit. Nevertheless, these can be analyzed as vowel-glide sequences, joining a larger series. Stressed *ai* represents [aj], filling the gap in [_, aːj, ej, eːj, oj, oːj, uj, uːj].[11] *Au*, for its part, is marginal, found in 0.03% of word tokens, mostly loanwords from Sanskrit. Therefore, in what follows, I take the Tamil vowels to comprise only monophthongs.

(19)

	front	central	back
high	i(ː)		u(ː)
mid	e(ː)		o(ː)
low		a(ː)	

Short /u/ has a reduced allophone known traditionally as *kur̲riyalukaram* 'overshort *u*', typically in word-final position (Rajam 1992:44–6, Niklas 1993:29–33, Wilden 2018:23).[12] For example, [muːkku] 'nose' might be more narrowly transcribed as [muːkkɯ] or [muːkkɨ]. This article transcribes reduced [u] as 'u' throughout, as its allophony is irrelevant here. The present transcriptions are consistent with Tamil orthography, which does not encode the distinction. Moreover, the exact distribution of the reduced allophone in Kamban is unclear; for one thing, medieval grammarians differ as to whether it occurs after sonorants (Niklas 1993:31). Nor is it clear which IPA symbol to use for the reduced allophone in Kamban's time. Reduction has no effect on moraicity; reduced vowels (when not elided) scan like other short vowels in Kamban.[13]

[10] Standard transliteration employs *r̲r̲* for what I transcribe as [tt̲]. The latter is the historical source, and was likely still Kamban's pronunciation of the geminate. Compare Christdas' (1988:140) dialect of Tamil, which to this day retains [tt̲]. The usual Modern Tamil pronunciation of *r̲r̲* is [tr̲], retaining the historical stopness but adding rhotacism.

[11] This analysis is also adopted by Christdas (1988:163) and Krishnamurti (2003:90). It is further reinforced by poetic rhyme, in which the first post-vocalic consonant must match, but vowel quality is ignored. *Stressed* /aj/ is singled out because /aj/ is monophthongized (to [e] or similar) when unstressed.

[12] Short /i/ likewise has a reduced allophone in premodern Tamil, though it is less common than overshort *u*.

[13] In this respect, Middle Tamil differs from Old Tamil, where the distinction between full and reduced vowels is more important for metrics (Rajam 1992:117, Chevillard 2014:55).

As a distinct issue, some analysts take overshort *u* to be epenthetic. For example, surface [muːkkɯ] might be analyzed as underlying /muːkk/. This article treats word-final /u/ as underlying (assuming that it elides before another vowel, as in §2.3), but nothing critical to the present analysis hinges on that decision.

2.2 Syllable structure and phonotactics

Turning to syllable structure, the onset is straightforward: Complex onsets are forbidden across the board. Empty onsets are permitted, but only word-initially. Lexically (that is, ignoring sandhi), only ten word-initial onsets are possible, namely, [Ø, p, t̪, c, k, m, n̪, (ɲ), ʋ, j] ([ɲ] being uncommon). Medially, any consonant may be an onset (except *āytam*; footnote 9), but the onset must be filled.

The nucleus, as mentioned, is a monophthong, short or long.

As for the coda, lexically, a *word-final* coda may be only certain sonorants, namely, [j, ɾ, ɻ, l, ɭ, m, n, ɳ]. Word-medially, there are more possibilities; indeed, any consonant can be a coda, if nothing else as part of a geminate. Limited two-consonant codas are possible, even after long vowels. Specifically, one of [j, ɾ, ɻ] may precede a geminate or nasal-plosive cluster (e.g. the interludes [ɾt̪t̪], [jt̪t̪], [ɾn̪t̪], and [jmp] in (26) and (33) below).

A complete accounting of the phonotactics of medial sequences is not critical for present purposes (see Christdas 1988), but a few generalizations nearly summarize the system: Homorganic nasal-plosive sequences, like geminates, are ubiquitous (e.g. (20a)). Certain heterorganic nasal-plosive sequences are also found, but less frequently (one-to-ten), and typically across morphemes (e.g. (20b)). Obstruent-obstruent clusters have only the form {t̪, ʈ}{p, c, k} and are always derived (e.g. (20c)). Non-geminate sonorant-sonorant clusters are somewhat infrequent, but several combinations are possible (e.g. (20d)). Finally, obstruent-sonorant clusters are never permitted.

(20)
a.	an̪t̪a	‘that’
b.	in-pam	‘delight’
c.	/n̪aɭ-pu/ → [n̪aʈ-pu]	‘friendship’
d.	aːɳ-maj	‘manliness’

Summarizing, the Tamil syllable has the basic schema in (21), where C2 ∈ [j, ɾ, ɻ]. If C2 is present, C3 cannot be an approximant. Sonority sequencing is never violated. The restriction of C2 to [j, ɾ, ɻ] (notated ‘J’ below) entails that a geminate may be preceded by a glide or rhotic, but not by a nasal or obstruent.

(21) (C1)V(ː)(C2)(C3)

2.3 Syllable weight and moraicity

The present analysis takes all Tamil codas (including geminates) to be moraic. That is, Tamil exhibits ‘weight by position’ (WbyP) and is compatible with the moraic theory of geminates (Hayes 1989; Davis 2011). This section briefly surveys several weight-sensitive phenomena in Middle Tamil, all of which support analyzing codas as moraic.

First, all Tamil meters are quantitative. For example, one of Kamban's meters is a catalectic trochaic octameter, as in quatrain (22).[14] In this meter, every non-final foot is heavy-light. Closed syllables count as heavy, regardless of their position in the word (e.g. /aɽuṯṯa/, /ṯuruṯṯi/).

(22) a. ṯaːṯu / koɳɽa /ciːṯa' / meːʋu / caɳṯu, / caɳṯa / meɳ̲ r̲a/l̪ir, Ø
b. poːṯu, / koɳɽ' a/ɽuṯṯa/poːṯu, / poŋku / ṯiː ma/ruɳṯi/n̲aːl Ø
c. ʋeːṯu / koɳɽa/t' eɳ̲n̲a, / meːn̲i / ʋeɳṯu / ʋeɳṯu, / ʋimmu / ṯiː Ø
d. uːṯu / ʋan̲ r̲u/ruṯṯi/poːl, u/jiṟṯ' u/jiṟṯ', u/jaŋki/n̲aːn̲. Ø

Second, codas count for weight for prosodic word minimality. A monosyllable must contain a long vowel or coda (or diphthong; §2.2). However, because an isolated word (of any length) cannot end with an obstruent (§2.2), the minimality of isolated words does not confirm the moraicity of obstruents specifically.

That said, minimality also plays a role in sandhi, and in that context, obstruent codas (both singleton or geminate) are confirmed to be moraic. In Kamban, word-final /u/ normally elides before another vowel. In (22), for instance, apostrophes (except the first) indicate this type of elision. Elision is suppressed, however, if it would render the word subminimal. Thus, elision applies to words of type (C)VCCu (e.g. /koɳɽu/ 'taking'), (C)VCːu (e.g. /paṯṯu/ 'ten'), and (C)VːCu (e.g. /poːṯu/ 'when'), as well as longer words, but not words of type (C)VCu (e.g. /kaɽu/ 'bitter'). Elision is blocked in the last case because it would render the moraic trochee degenerate (monomoraic), as *[(ka)(ɽ' V...)], given that consonants resyllabify across words.[15]

Third, consider stress. In descriptions characterizing Tamil stress as quantity-sensitive, codas count for weight. Per Christdas (1996) (as analyzed by Gordon 2004), stress is initial, but shifts to the peninitial iff the initial is light (i.e. (C)V) and the peninitial has a long vowel. This rule implies that at least initial-syllable codas (as always, including geminates) are moraic. While Gordon's (2004:footnote 2) analysis of the rule assumes that non-initial codas are non-moraic, Ryan (2019:62) presents an alternative analysis of the same rule that preserves coda moraicity in all positions.[16] At any rate, stress is not strong evidence vis-à-vis Kamban's Tamil, as it is difficult to say exactly how stress worked at that time; even for present-day Tamil, descriptions vary.

Fourth, morphophonology arguably supports the moraic status of consonants. For example, verbs may be causativized by geminating the final obstruent, as in (23) (Rajam 1992:58f, 524–33, Wilden 2018:78), which can be analyzed as mora affixation. In (a–b), the initial syllables are already heavy in the bases, but in (c), causativization changes the weight category of a syllable.

[14] In Tamil terms, this meter is a variety of *el̪ucīr canta viruttam* (Tirumurukan̲ 1987:57). This example is Kamban's verse 3254, translation: 'When sandal paste to cool him mingled with soft shoots and buds were spread on him, sweet-smelling powders and lovely then as if the skin were burning up from a caustic salve, his body burned, burned as he suffered. He sucked his breath and sucked his breath like a powerful bellows blowing and blowing up a raging fire' (Hart & Heifetz 1988:178).

[15] This analysis takes /u/ to be underlying. Alternatively, one might treat [u] as epenthetic (when *not* prevocalic), or perhaps relate elision to the distribution of the reduced allophone of /u/ (§2.1). Even with these alternatives, however, the moraic trochee would still be a conditioning factor for epenthesis or reduction, supporting moraic codas. Further, pronouns such as /aṯu/ 'that' and /iṯu/ 'this' sometimes elide, ostensibly flouting the stated generalization. As function words, they are free to cliticize, in which case they are not subject to minimality.

[16] Ryan (2019:62) employs (and argues independently for; cf. Ryan 2020) a constraint StressVV that allows long vowels to be more stress-attracting than VC rimes even when both are bimoraic. Paramore (2025:17) invokes the same constraint, calling it S$\rightarrow$ $[\mu_V \mu_V]_\sigma$.

(23)	a.	aːɽu	‘move’	aːɽɽu	‘shake’
	b.	oːɽu	‘run’	oːɽɽu	‘drive’
	c.	peruku	‘increase’	perukku	‘(cause to) increase’

Fifth, in Old Tamil (Middle Tamil will be addressed presently), when an adjective or numeral of the form (C)VCV is prevocalic, the word-final vowel deletes and a preceding segment lengthens. If the final consonant is geminable, it geminates, as in (24a–b) (Niklas 1993:21, Rajam 1992:59f, 86–8). If the consonant is a rhotic, and thus non-geminable, the preceding vowel is lengthened, as in (c–d). Ryan (2019:129) suggests that this process is compensatory lengthening, taking the mora to be shifted to an earlier segment. Alternatively, it could be regarded as another type of minimality-driven lengthening, since the modifier would be subminimal without lengthening. Either way, coda consonants are confirmed to be moraic. At any rate, by Kamban’s time, this kind of gemination appears to be losing productivity. While the cited words (among others) persist in Middle Tamil, Kamban’s usual strategy for prevocalic (C)VTV is to suppress elision, as discussed above.

(24)	a.	/kut̲u eɻut̻̲tu/	→	[kut̲t̲-eɻut̻̲tu]	‘short vowel/letter’
	b.	/n̻̲eɽu eɻut̻̲tu/	→	[n̻̲eɽɽ-eɻut̻̲tu]	‘long vowel/letter’
	c.	/iru aɽi/	→	[iːr-aɽi]	‘two feet’
	d.	/peru eɻil/	→	[peːr eɻil]	‘great handsome’

Summarizing so far, the available Middle Tamil evidence broadly supports the moraicity of codas. The survey is not exhaustive,[17] but the point is to situate the analysis of the phonotactic treated by this article, not to analyze all aspects of Tamil weight. Indeed, as I will argue, the phenomenon under focus in this article adds to the evidence that codas are uniformly moraic, as that assumption enables a straightforward analysis that is otherwise not available (see §4 for arguments against a counteranalysis in which obstruent geminates are treated as non-moraic).

Before closing this section, two additional points about weight are in order.

First, I leave open the possibility that the rhotic codas /ɾ/ and /ɻ/ are non-moraic in Middle Tamil, as argued by Ryan (2019:122–36) on the grounds that VR is subminimal as a prosodic word, VR scans as light in meter, and R cannot geminate. Ryan (2019) proposes the constraint *μ/R, and the same could be adopted here (highly weighted) without affecting the rest of the analysis.

Second, because the present analysis takes Tamil codas (except perhaps R) to be uniformly moraic, it follows that (C)VːC$_1$ syllables are trimoraic. Indeed, this assumption is critical to the proposed analysis, which invokes a gang effect involving *3μ. Although some scholars reject trimoraic syllables (starting with Trubetzkoy 1939), I assume, with Hayes (1989) and many others, that they are possible. I first briefly survey some cross-linguistic evidence for trimoraic syllables, followed by some Tamil-specific considerations.

Cross-linguistically, several stress systems treat VːC syllables as heavier than Vː syllables (e.g. Hindi, Kashmiri, Mankiyali, Nanti, Pulaar; Ryan 2019:35–43, 59). Second, quantitative meters can make the same distinction, gradiently or categorically (e.g. Finnish, Greek, Persian, Sanskrit, Tamil; *ibid.*:141, 148–51). Third, Hayes (1989:291f) cites compensatory lengthening as support, as it can apply to syllables that would already be heavy without it. Fourth, diphthongs can be

[17] On further gemination phenomena in present-day varieties of Tamil, see Christdas (1988), Sampath Kumar (2016:99f), and Venkatakrishnan & Kar (2025). Gemination is also richly involved with Middle Tamil morphophonology in ways that are beyond the scope.

trimoraic, as in Vedic Sanskrit, where [aːi, aːu] are superheavy, contrasting with heavy [ai, au] (Ryan 2021) (see also Blevins & Harrison 1999:217 on Gilbertese). Fifth, trimoraic monophthongs are firmly established (Remijsen et al. 2019; see also Becker et al. 2021, Baković 1996:5). Finally, two theory-internal points. Prince & Smolensky's (1993/2004:229) constraint *3μ, widely adopted in subsequent work, presupposes the possibility of trimoraic syllables, as otherwise no constraint would be needed. Moreover, the moraic theory of geminates (Hayes 1989; Davis 2011) predicts the trimoraicity of VːCː syllables.[18]

Turning to Tamil, the phenomenon analyzed by this article is itself sufficient evidence to posit trimoraic syllables, but some additional independent support—though not critical—can also be cited. First, phonetically, rime durations are in fact trimodal. As I show in §5, the duration ratio is nearly 1:2:3 among the syllables analyzed as mono-, bi-, and trimoraic. Second, mora affixation (as in (23)) implies trimoraicity, as it augments already bimoraic syllables. One cannot reanalyze mora affixation as a fortis/lenis alternation, since when the vowel of the base is short, the geminated form becomes categorically heavy. Similarly, boundary gemination applies even after a long vowel (§3). Finally, Ryan (2011:436–9) demonstrates that VːC is treated as gradiently but significantly heavier than Vː by Tamil meter.

3 A length co-occurrence restriction

As stated (§1), a long vowel may precede a geminate within the same syllable, but only if the geminate is an obstruent, as in (25). This generalization holds not just of Middle Tamil, but of all varieties of literary Tamil. Christdas (1988:259), for instance, observes that it holds of her dialect (Kan̪n̪iyākumari), and Sampath Kumar (2016:99) also notes it for spoken Tamil.

(25) a. jaːppu 'poetry'
b. ʋaːt̪̪tu 'duck'
c. puːʈʈu 'locking'
d. n̪̪aːkku 'tongue'
e. *jaːmma, *ʋaːlla, *puːɳɳa, *n̪̪aːjja, etc.

The generalization (that a long vowel may precede a geminate obstruent but not geminate sonorant) ignores an intervening glide or rhotic. That is, a glide or rhotic may intercede between a long vowel and obstruent geminate, as in (26a–b) (on their moraic structure, see §6). But a geminate sonorant remains illicit after a long vowel even if such a consonant intervenes, as in (26c).

(26) a. aːɾt̪̪tu 'feed'
b. maːjt̪̪tal 'to kill'
c. *aːɾn̪̪n̪̪a, *maːjn̪̪n̪̪al, etc.

Geminate sonorants are freely permitted after short vowels (including with an intervening glide or rhotic). As mentioned in §2, however, the rhotics cannot be geminate in any context.

[18] One might reply that such a geminate shares its mora with the preceding long vowel, but that would be an additional stipulation, not a prediction generated by the moraic theory of geminates.

(27)

a.	cemmaj	‘redness’
b.	mu̱ṉṉam	‘intention’
c.	illaj	‘no’
d.	u̱ḻam	‘heart’
e.	ṉojja	‘weak’

The gap is not accidental. Counts for Kamban’s text are in (28), and any other Tamil text would show the same gap. J represents [j, ɾ, ɻ], which, as mentioned, may intercede between a long vowel and geminate. The two tokens of Vː(J)Nː are both [ṯeːm-maɾu] ‘sweet fragrance’, which could just as well be written as two words.

(28)

	Tː	Nː
V(J)_	27,106	9,692
Vː(J)_	7,712	2

Moreover, the restriction is productive. First, Christdas (1988:260f) observes that while Tamil verbs may be causativized by geminating the final consonant, as in (29a–d) (see also (23) above), the strategy is unavailable for VːN verbs, such as (e) (cf. also Rajam 1992:58f, 524–33, Wilden 2018:78).[19] The same holds of Kamban’s Tamil. However, the argument is weakened by the fact that causative gemination does not target VN verbs, even though VNː would be licit. For Kamban, V(ː)N verbs causativize either by overwriting the sonorant with a geminate plosive (e.g. [muːl̪] ‘emerge’ vs. [muːt̪t̪i] ‘initiate’) or by periphrasis. At any rate, the fact remains that there are numerous processes of gemination in Tamil (more to be described presently), and all conspire to avoid creating Vː(J)Nː.[20]

(29)

a.	aːt̪u	‘move’		aːt̪t̪u	‘shake’
b.	maːɾu	‘(undergo) change’		maːttu	‘(cause) change’
c.	eːɾu	‘climb’		eːttu	‘cause to climb’
d.	t̪aːŋku	‘bear’		t̪aːkku	‘lift’
e.	kaːl(u)	‘flow’		*kaːllu	‘(e.g. pour)’

Second, as also argued by Christdas (1988:264), literary Tamil [ṉɾ̱] becomes [ɳɳ] in spoken dialects (e.g. [oṉɾ̱u] > [oɳɳɯ] ‘one’). However, original [ṉɾ̱] yields a singleton after a long vowel (e.g. [muːṉɾ̱u] > [muːɳɯ] ‘three’).[21]

Another argument, not raised by Christdas (1988), concerns sandhi. In Kamban’s corpus, when two identical sonorants flank a word boundary, the first optionally deletes (unless doing so would violate prosodic minimality). To avoid interference by minimality, I consider only cases in which the first word is at least disyllabic. Sonorants are counted as identical if either they are underlyingly identical (given the lexical phonotactics in §2, this situation arises only for /m m/) or would have

[19] For (b–c), recall that singleton /ṯ/ surfaces as [ɾ̱] when intervocalic. In (d), the nasal deletes because nasal-geminate sequences are illicit (§2.2).

[20] Christdas (1988:262f) further observes that the sets of stem-final sonorants and suffix-initial sonorants are complementary, such that Vː(J)Nː- does not arise via morphology.

[21] Venkatakrishnan & Kar (2025:116) offer an independent case: In spoken dialects, (C)VN words undergo paragoge with gemination (e.g. [kal] > [kɐllɯ] ‘stone’). In (C)VːN, by contrast, paragoge occurs without gemination (e.g. [paːl] > [paːlɯ] ‘milk’).

to totally assimilate if both surfaced (e.g. /ɳ n̪/ → [ɳ ɳ]). If the first sonorant follows a short vowel, it surfaces 13% of the time, vs. 2% after a long vowel, a significant difference (Fisher's exact test $p < 0.001$). Moreover, the exceptional 2% (3 tokens) all straddle a sentence boundary. Thus, within an intonation phrase, V:(J)N_x#N_x (where the two sonorants surface identically) never occurs (excepting the two tokens of [t̪e:m-maru] mentioned above). V:(J)T_x#T_x, by contrast, is extremely common, and even actively formed: In certain (frequent) syntactic contexts, plosives geminate across word boundaries: V(:)(J) T → V(:)(J)T T (Wilden 2018:29f, cf. also Mohanan 1982:36f on closely related Malayalam).

In conclusion, the fact that an obstruent but not sonorant geminate may follow a long vowel within the syllable is established not just by the lexicon, but also by morphophonology, diachrony, and sandhi. It is clearly a productive restriction, and has evidently been stable throughout Tamil's literary history of two millennia.

4 The geminate restriction is not purely weight-based

Cross-linguistically, restrictions on geminates following long vowels are usually attributed to weight. In the simplest case, they fold in with a general prohibition on trimoraic syllables (*3μ). For example, Koasati has long vowels, geminates, and other codas (Kimball 1991). But a long vowel cannot immediately precede a coda (geminate or otherwise). That is, syllables of type V:C are forbidden. Assuming that codas are moraic (WbyP; weight by position), *3μ can be top-ranked.[22]

In Koya, a Dravidian language related to Tamil, it is reported that a long vowel cannot immediately precede a geminate, but may precede other codas (Tyler 1969). Davis (2011:120f) analyzes Koya by positing that geminates, but not other codas, are moraic, as could be achieved by MAX-μ ≫ *C_μ ≫ WbyP. This entails that V:C: alone is (or would be) trimoraic. Given this assumption, *3μ can once again be top-ranked to complete the analysis.[23]

The question, then, is whether Tamil could submit to a similar analysis, one based purely upon weight (e.g. unconjoined *3μ). To be sure, my analysis of Tamil invokes weight in some manner, in that *3μ is part of the gang effect. The question here is whether *3μ could act alone, given the right assumptions about moraification. I will argue that no such analysis is tenable.

Specifically, let us consider whether an analysis is possible in which a V:T: syllable (ignoring the optional intervening glide or rhotic) is bimoraic, whereas a V:N: syllable (ditto) is trimoraic. If that were the case, *3μ alone could achieve the needed restriction. Tableaux (31–30) illustrate the logic of the counteranalysis. As indicated by the (lack of) subscript μ, the geminate is taken to be moraic in (31a) but (here) non-moraic in (30a). MAX (or IDENT(long) if one prefers) is used here to penalize degemination regardless of the geminate's moraicity.

[22] The Koasati example is complicated by the fact that weight for stress ignores the coda. See Ryan (2019:72f) for an analysis of Koasati that preserves the moraicity of codas. At any rate, closed-syllable shortening is widely attested, as in Icelandic (Kager 1999:267f), Kashaya (Buckley 1991), Wolof (Faust & Ulfsbjorninn 2025:30), Yawelmani (Newman 1944:25), and other languages.

[23] The empirical facts just described for Koya have been called into question (Curtis 2003:167–70, Faust & Ulfsbjorninn 2025:59). Nothing here hinges on whether the case goes through.

(30)

	/muːmma/	*3μ	MAX
	a. muːm$_\mu$ma	∗!	
☞	b. muːma		∗

(31)

	/muːkku/	*3μ	MAX
☞	a. muːkku		
	b. muːku		∗!

For this analysis to work, one would first need to allow non-moraic true geminates. Geminates like [kk] (see (25) for further examples) are phonetically long, as transcribed, and cannot be reanalyzed as singletons (see §5). Some linguists do not admit non-moraic true geminates, but let us assume, for the sake of argument, that they are possible, and move forward with the counteranalysis.

The second problem with this analysis is that it cannot merely be stipulated that obstruent geminates following long vowels are non-moraic; it needs to fall out from some ranking. To be sure, it is not unnatural in general for obstruent geminates to be lighter than sonorant geminates (e.g. Topintzi & Zimmermann 2025:345, A8). Cross-linguistically, weight for syllable structure may be contingent on sonority. One example is Osthoff's Law, whereby a sonorant coda cannot co-occur with a long vowel, as in Lithuanian (Zec 1995:101) and other Indo-European languages (Sayeed 2017). As two further examples, in Kwak'wala, a sonorant cannot follow a full vowel, which has been analyzed as a monomoraic maximum (Bach et al. 2005). In Huehuetla Tepehua, loss of a sonorant but not obstruent coda triggers compensatory lengthening (Kung 2007:139f).

Thus, one might consider invoking a constraint such as *T$_\mu$ to render moraicity contingent on sonority. The problem with this approach for Tamil, however, is that there is strong evidence that obstruent geminates are moraic in the general case (§2.3). The evidence is especially clear after short vowels. The defender of the counteranalysis might therefore admit that Tː is moraic after a short vowel (such that the whole syllable is two moras), but not after a long vowel (lest the whole syllable be three moras). In other words, the moraicity of Tː, one might argue, is contingent on whether a long vowel precedes it, as summarized by (32).

(32) Desired outcomes under a possible counteranalysis

a.	VTː$_\mu$	Tː is moraic after a short vowel; satisfies *3μ, thus allowed
b.	VNː$_\mu$	Nː is moraic after a short vowel; satisfies *3μ, thus allowed
c.	VːTː	Tː is non-moraic after a long vowel; satisfies *3μ, thus allowed
d.	VːNː$_\mu$	Nː is moraic after a long vowel; violates *3μ, thus prohibited

On this analysis, we may begin by ranking *3μ on top, as the whole idea is that *3μ is never violated, thus ruling out (d). Next, we need to rank the constraint implementing sonority-sensitivity (e.g. T$_\mu$) against WbyP to determine which codas are moraic. The problem is, neither possible ranking works: If T$_\mu$ ≫ WbyP, (a) is misgenerated; if the reverse, (c) is misgenerated.[24]

[24] The issue remains if one employs a constraint like WbyP-N ('coda sonorants are moraic') to implement sonority-sensitivity. Since WbyP is still part of CON, a constraint favoring non-moraicity is needed to permit some codas to be non-moraic, such as *C$_\mu$. In that case, neither ranking of WbyP and *C$_\mu$ works.

Another strategy that one might consider to confine the effect of *T_μ to post-Vː position would be constraint conjunction. For example, *T_μ &$_\sigma$ *3μ penalizes moraic obstruents in trimoraic syllables. Ranked above WbyP, it would permit Tː to be moraic after a short vowel but not after a long vowel, as desired. But this analysis also fails: Syllables of the type Vː(J)N are licit, as exemplified by (33). The ranking just specified would render the sonorant codas moraic in (33), and thus the syllables as wholes trimoraic, defeating the *3μ analysis. The generalization is that a sonorant *geminate* cannot follow a long vowel, not that a sonorant *coda* cannot. Indeed, the coda can even contain two sonorants after a long vowel, as in (d–e).[25]

(33) a. paːmpu ‘snake’
b. aːn̼ra ‘excellent’
c. iːɳʈu ‘here’
d. ceːr̼n̼taːr ‘dependents’
e. kaːjmpan̼aj ‘fruit palm’

To conclude, this section considers an alternative analysis whereby the illicitness of *Vː(J)Nː is motivated by *3μ, while Vː(J)Tː is allowed because Tː is non-moraic, and thus *3μ does not apply. This counteranalysis has three problems. First, it requires non-moraic true geminates. While some scholars argue for this possibility (see Topintzi & Zimmermann 2025 for a recent review), others view it as highly marked or impossible (cf. e.g. Hayes 1989, Davis 1999, Ryan 2019:11, Paramore 2025). Second, even if one accepts non-moraic geminates, a ranking is needed to render them non-moraic in the right context. In this case, all geminates must be rendered moraic except Tː following a long vowel. As I argued, no ranking of the standard constraints—even allowing for conjunctions—captures the needed state of affairs. A third point against this counteranalysis is that there is independent evidence that Tː remains moraic even after a long vowel, though it is admittedly modest (last paragraph of §2.3). For instance, Tː is not substantially compressed after a long vowel (§5), as one might expect if it were non-moraic specifically in that context.

5 Against a reanalysis of the empirical facts

The foregoing discussion assumes that geminates are correctly transcribed as such. If Tː in the context Vː(J)Tː were not actually a geminate (say, if T vs. Tː were actually a voicing or fortis/lenis distinction), an analysis in terms of unconjoined *3μ would become available. This section confirms that Tː is indeed a geminate, even after a long vowel. Therefore, no such reanalysis is tenable.

Figure 1 shows consonant durations based on audio tapes accompanying Hart (1999). Each bar represents 20 tokens (except for the glide/rhotic categories, which are more sparsely populated), gathered in order from passages read aloud by the book’s author, starting at chapter 20. Only stressed, word-initial syllables of words containing at least two syllables were measured, and rhotics (being non-geminable flaps) were not counted as N. All consonants were intervocalic unless otherwise indicated. In the aggregate, sonorant geminates are 2.3 times as long as sonorant singletons, and obstruent geminates 2.9 times as long as obstruent singletons. More to point, after

[25] The hypothetical analysis fails for another, independent reason. It erroneously predicts that the initial syllable of VTːVːN is light. Tː is a coda, but also part of the following, superheavy syllable, so its mora would be suppressed by the conjunction.

a long vowel, T: is 2.7 times as long as singleton T. While there is naturally some compression of T: after a long vowel (mean 140 ms) vs. short vowel (mean 170 ms), in either case, T: is roughly three times the duration of a singleton.

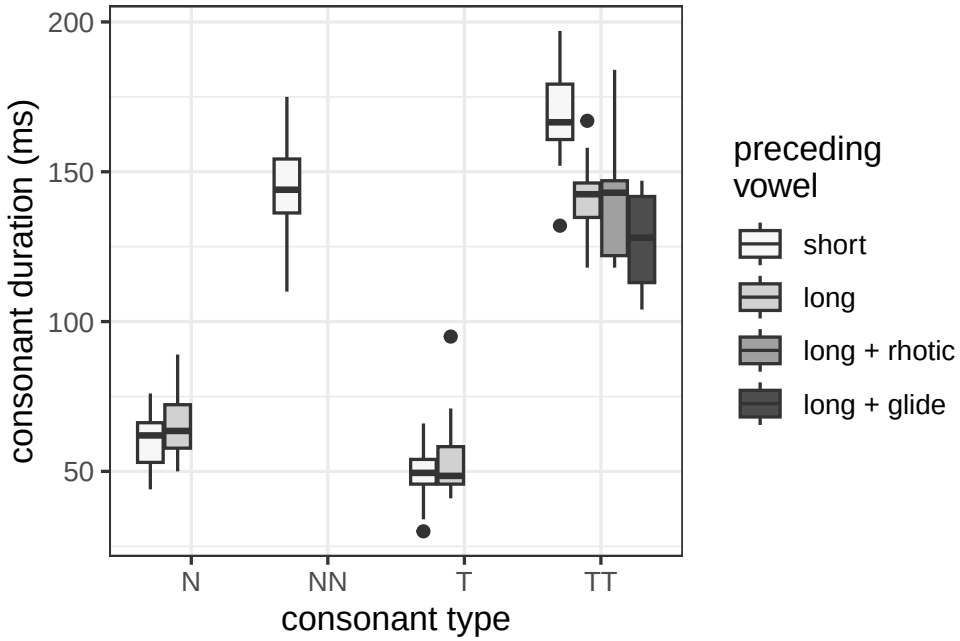


Figure 1: Duration of different consonant types (four groups) according to whether the preceding vowel is long or short (shading).

Thus, the transcriptions are correct. While the recordings do not represent Kamban's period, given that gemination is a conservative feature of Dravidian, the persistence of such contrasts in Modern Tamil all but guarantees them at earlier stages. While this phonetic study is small in scale, it need not be belabored, as I am merely corroborating what is already the consensus. It is a sanity check, so to speak. The fact that T: is a geminate (including after a long vowel) is further reflected by the script as well as by traditional grammars, which explicitly treat geminates as moraic (i.e. contributing to *māttirai* count), like other codas (Zvelebil 1989:7f, Rajam 1992:46f, Niklas 1993:36f). (§2.3 presented other support for geminacy.)

Figure 2 is based on the same data as Figure 1, but shows overall rime durations. Only geminate codas are considered, counting half of their duration towards the rime. The rime durations comprise three groups, corresponding to what I analyze as monomoraic, bimoraic, and trimoraic, with mean durations of 78, 149, and 230 ms, respectively. That is, bimoraic rimes are 1.9 times as long as monomoraic rimes, and trimoraic rimes are 2.9 times as long as monomoraic rimes, for a ratio of roughly 1:2:3. Interestingly, the intervening glide/rhotic 'J' in V:(J)T: does not add to the overall duration of the rime, as I analyze via mora sharing in §6.

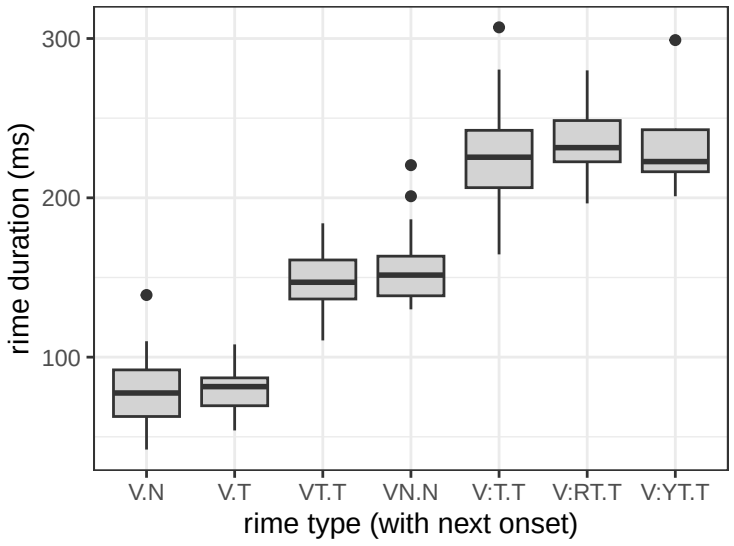


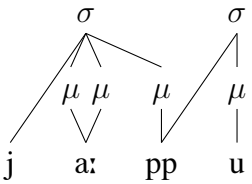
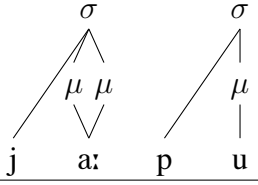
Figure 2: Rime durations comprise three groups (monomoraic, bimoraic, and trimoraic). The following onset is shown but not counted as part of rime duration.

6 A general analysis of moraicity in Tamil

This section presents a general analysis of the moraicity of Tamil syllables that is compatible with the analysis of the gang effect in §1. It posits the following. First, a short vowel contributes one mora, and a long vowel two. Second, a geminate always contributes a mora to the preceding syllable. Third, coda consonants are moraic unless the syllable would exceed three moras.

The three-constraint HG grammar in §1 nearly already encodes this state of affairs; the only addition here is WbyP (penalizing any coda consonant that does not project its own mora) at 2.5. This addition does not alter the winners in §1. As illustrated by tableau (34) for [ja:ppu] 'poetry', because MAX-μ is highly weighted, a trimoraic syllable cannot be avoided by shortening a long vowel or geminate.

(34)

/ja:ppu/	$\mathscr{H}$	MAX-μ 3	WbyP 2.5	*SONGEM 2	*3μ 2
☞ a. 	−2				−1
b. 	−3	−1			

If J (a rhotic or glide) intervenes between the long vowel and geminate, as with [kaːjt͡tu] ‘ignite’ in (35), J does not contribute its own mora. Rather, it shares a mora with the long vowel. No candidate is shown in (35) in which both [j] and the geminate are moraic, as the result would be a quadrimoraic syllable. I assume that such syllables are either ruled out by GEN (i.e. are impossible in general) or else by highly-weighted *4μ, not shown. Furthermore, I assume that J affiliates with a vocalic rather than consonantal mora, as is supported by the phonetics (§5): VːJ (mean 167 ms) does not significantly differ from Vː (mean 169 ms). The latter could be ruled out by a constraint penalizing a mora dominating multiple consonants. The direction of sharing, however, has no impact on the core proposal here, so I will not pursue the formulation of such a constraint. I take (a) to violate WbyP because [j] does not possess its own mora; rather, it shares a mora with the vowel. Hayes (1989:258) is explicit that mora-sharing VC violates WbyP.[26]

(35)

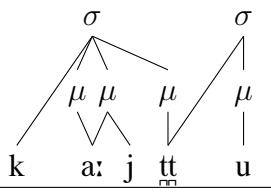
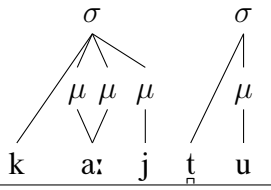
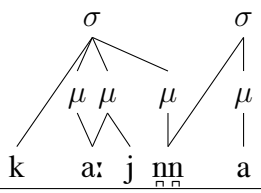
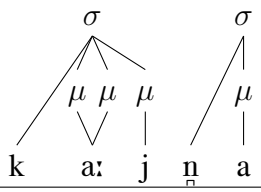
/kaːjt͡tu/	$\mathscr{H}$	MAX-μ 3	WbyP 2.5	*SONGEM 2	*3μ 2
☞ a. 	−4.5		−1		−1
b. 	−5	−1			−1

Tableau (36) reveals that a hypothetical illicit form such as *[kaːjn͡na] is properly rejected. The glide intervening between the long vowel and geminate does not affect the analysis.

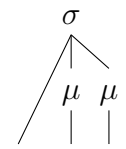
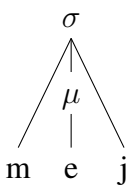
[26] An anonymous reviewer points out that this analysis, as it stands, does not distinguish between the licit sequence JTː and the illicit sequence *NTː. The distinction is not weight-related: Illicit *VNTː is not heavier than licit VːTː. One possible explanation is that *NTː would be a kind of ‘supergeminate’, since the whole sequence would share a place node, unlike JTː. At any rate, this issue is orthogonal to the proposal developed by this article.

(36)

/kaːjn̪n̪a/	$\mathscr{H}$	MAX-μ 3	WbyP 2.5	*SONGEM 2	*3μ 2
a. 	-6.5		-1	-1	-1
☞ b. 	-5	-1			-1

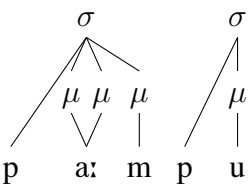
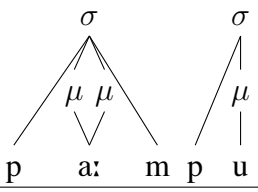
In (37), coda [j] is parsed as moraic in [mej] ‘truth’.

(37)

/mej/	$\mathscr{H}$	MAX-μ 3	WbyP 2.5	*SONGEM 2	*3μ 2
☞ a. 	-0				
b. 	-2.5		-1		

Finally, my analysis allows trimoraic V:N where N is a singleton, as with [paːmpu] ‘snake’ in (38).

(38)

/paːmpu/	$\mathscr{H}$	MAX-μ 3	WbyP 2.5	*SONGEM 2	*3μ 2
☞ a. 	-2				-1
b. 	-2.5		-1		

7 Potential cases from other languages

Generalizing from the Tamil case, this article's argument would likewise be supported by (at least) any language with the following properties: (1) VːC is a possible (non-final) syllable, (2) geminates occur, (3) VːC is not permitted if C is a geminate (or a subset of geminates), and (4) codas are moraic across the board. This section describes a few potential additional cases of this profile, though all require further investigation. As a reminder, if a language forbids geminates after long vowels but does not allow *any* coda after a long vowel (e.g. Koasati, Wolof), it is not relevant, as the absence of VːCː can then be attributed to *3μ alone (§4).

First, Hindi permits both VːC syllables and geminates, but not their combination as VːCː (Ohala 1983:6). As Curtis (2003:164–6) notes, codas are moraic in Hindi—indeed, there is evidence that they are moraic even after a long vowel—and therefore the prohibition on VːCː does not reflect a bi-moraic maximum. The analysis of Tamil cross-applies, merely substituting *GEM for *SONGEM, that is, MAX-μ $>$ *3μ, *GEM, weighted to gang. The HG analysis, unlike the OT-LC analysis, once again captures that a geminate may *precede* a long vowel in Hindi.

Second, Hungarian features long vowels, geminates, and VːC syllables. Codas are clearly moraic after V and have been proposed to be moraic after Vː as well (e.g. Morén 1999:241–6, Cohn 2003:79f), though Curtis (2003:263) notes that the moraic status of (non-geminate) C in VːC is unclear. While VːCː is possible in derived environments (due to affixation and assimilation), it is highly marked within morphemes. Curtis (2003:263f) and Törkenczy (2004:119) observe that with rare exception, Hungarian forbids VːCː within morphemes (see the latter for discussion of the exceptions). Moreover, (at least some of) the exceptions are optionally repaired (references in Curtis 2003:263). The analysis just suggested for Hindi would work.

A third possible case is Fula, specifically, the Pulaar dialect of Mauritania and Senegal. Per Paradis (1988:79f), Pulaar permits VːC syllables unless the coda is a geminate. It is clear from Paradis (1988) and other sources that codas are moraic at least in VC. Niang (1993, 1997) reports at length that VːC outweighs Vː for stress in Pulaar, which supports the moraicity of C in VːC. However, per Niang's (1993; 1997) description, geminates *are* permitted after long vowels. Thus, while Paradis (1988) and Niang (1993, 1997) ostensibly analyze the same language, there might

be dialectal differences. Further study is needed. If the geminate restriction described by Paradis (1988) coexists with the stress system described by Niang (1993, 1997) (note that Paradis 1988 does not discuss stress), it would furnish another case like Tamil.

A fourth possible case is Thurgovian Swiss German. Sonorant and obstruent geminates both occur, but only the latter may follow a long vowel (Kraehenmann 2003:§5). In this respect, the case resembles Tamil. Nevertheless, there are several complications. First, unlike Tamil, V:N does not occur at all: A sonorant coda of any kind (geminate or not) may only follow a short vowel. Second, gemination is contrastive *word-initially* in plosives, but not in sonorants (or fricatives). Third, sonorant but not obstruent codas contribute to weight for stress.[27] That is, stress treats VN as heavy but VT: as light. In sum, the case is an intriguing parallel, but I can do no better here than to cite Kraehenmann's (2003) book-length treatment.

8 Alternative analyses

I have already argued in §4 against an alternative account whereby Tamil's geminate restriction is driven exclusively by weight, that is, by *3μ acting alone, on the assumption that T: is non-moraic following a long vowel. This section addresses other conceivable counter-proposals.

8.1 Brute force

First, there is brute force, positing all-in-one *V:(C)N: (cf. Christdas 1988:259). This approach misses the generalization that the restriction reflects the interaction of two simple, natural components, namely, *3μ and *SONGEM. Because it is completely stipulative, it opens the door to equally stipulative but unnatural constraints such as *V:(C)T: (switching the values of sonorant and obstruent) or *VC: (penalizing a geminate after a short vowel), among countless others.[28] The present analysis not only avoids a stipulative constraint, but avoids the addition of *any* constraint: All parts of the analysis are independently needed; I merely show how they work together.

8.2 Self-conjunction or counting cumulativity

My proposal involves *3μ ganging up with *SONGEM. Note that there is some resemblance between these two constraints, in that both penalize prototypically heavy elements. (Indeed, the HG analysis capitalizes on this co-relevance, which is a prerequisite to ganging.) In particular, a long vowel and a sonorant geminate are both long $[+sonorant]$ segments. Therefore, consider a hypothetical self-conjunction, *LONGSONORANT$^2_\sigma$, which penalizes their co-occurrence within a syllable. This constraint correctly rules out V:(J)N:, but erroneously also rules out N:V:. (In this respect, it has the same problem that the conjunction *SONGEM &$_\sigma$ *3μ had in §1.) A second problem is that *LONGSONORANT is dubious as a constraint. Given the independent necessity of both *LONGVOWEL and *SONGEM, it is at least unnecessary. Moreover, while both segment types are

[27] For this reason, Kraehenmann (2003:172) treats word-internal syllable-final sonorants as being part of the nucleus.

[28] Similarly, relatively 'omnivorous' statistical learners (cf. Albright & Hayes 2003, Hayes & Wilson 2008, or modern LLMs) could, with the right settings, learn constraints like *V:(C)N:, but they could equally learn any other constraint of the same complexity, regardless of naturalness.

descriptively 'long', in moraic formalism, their length manifests in different ways—as two moras for the vowel, but one mora for the consonant. Thus, such a constraint penalizes a disjunction of formal objects.

An HG analysis based on counting cumulativity is likewise unavailable. Imagine that *LONGSONORANT (if feasible) were weighted 1.5 and MAX-μ were weighted 2. Even though two violations of the former (for, say, V:N:) would incur a greater penalty than than one violation of the latter (for, say, shortening N:), shortening of N: would only alleviate one violation of *LONGSONORANT, not the needed two. In other words, MAX-μ and *LONGSONORANT trade off symmetrically, but counting cumulativity requires an asymmetric trade-off (Pater 2009). Ganging cumulativity is critical for this case.

8.3 Alternative conjunctions

I argued in §1 that the conjunction of *SONGEM $\&_\sigma$ *3μ does not work as needed. Other conceivable conjunctions fare no better. Consider, say, *LONGVOWEL $\&_\sigma$ *N$_\mu$, penalizing the co-occurrence of a long vowel and moraic sonorant within a syllable. This conjunction has two problems. First, *N$_\mu$ is not a viable constraint; if anything, sonorants *prefer* to be moraic. *N$_\mu$ therefore generates markedness reversals, in which obstruent but not sonorant codas count as heavy. Second, N:V: would still violate this constraint. Even though the mora of N: is part of the preceding syllable, the *segment* N: would still incur a violation from *N$_\mu$, and the *segment* N: is still a part of the following syllable. A third but perhaps weaker objection is that this counteranalysis requires singleton N to be non-moraic in V:N, *contra* my analysis, which takes such sonorants to be moraic (§6). There is modest evidence that C in general is moraic in V:C (§2.3, last paragraph).

As another possibility, replacing *3μ with *LONGVOWEL—such that the whole conjunction is *SONGEM $\&_\sigma$ *LONGVOWEL—would not work. N:V: would remain erroneously penalized under this formulation. A gang effect of *SONGEM and *LONGVOWEL would also not work in HG, since the needed asymmetric trade-off would no longer arise, such that incurring one violation of MAX-μ addresses both markedness constraints.

8.4 Redefining *SONGEM (locality hacking)

One might further attempt to rescue the conjunction *SONGEM $\&_\sigma$ *3μ (as refuted in §1) by redefining *SONGEM to penalize only the mora of the sonorant geminate, not the whole segment. Under that definition, N:V:T: no longer erroneously incurs a violation of the conjunction, since the mora of the sonorant is not part of the trimoraic syllable. This counterproposal has two problems. First, geminacy can be ascertained only by evaluating the whole segment (i.e. checking whether the root node is multiply-linked). Thus, the domain of *SONGEM is properly the segment. The marked element of N: is not its mora alone (recall the argument against hypothetical *N$_\mu$ in §8.3). Second, this kind of locality hacking, so to speak, is *ad hoc*, and as such not cost-free. It adds a stipulation to the definition of *SONGEM—namely, that it penalizes specifically the mora—that is otherwise unmotivated, and opens up the possibility of redefining the localities of all constraints in *ad hoc* ways. Walker (2017:12), for instance, argues against locality hacking vis-à-vis a constraint on vowel harmony. To be restrictive and non-*ad hoc*, the domain of a constraint should be emergent (its structural description), not an arbitrary extra parameter.

8.5 Alternative domains of conjunction

As noted in §1, the conjunction *SONGEM $\&_{\sigma}$ *3μ would work if it were possible to make its domain the rime (i.e. hypothetical *SONGEM $\&_{rime}$ *3μ), but the rime is not a constituent in moraic theory. But consider a version of conjunction that allows domains to be non-constituents. Indeed, it is not unusual to find non-constituents (per moraic theory) such as 'onset', 'coda', or even 'cluster' invoked in the literature, though the propriety of this practice (using English words rather than constituents) is questionable. Łubowicz (2005) proposes that conjoined markedness constraints must have the same locus of violation, which entails that domains need not be constituents (e.g. constraints can penalize biphones). However, *3μ (locus = syllable) and *SONGEM (locus = segment) could not be conjoined under that proposal. Thus, for the sake of argument, imagine that any non-constituent may be a domain. *SONGEM & *3μ would work if its domain were the rime (and *3μ were redefined to evaluate the rime rather than the syllable). That said, not only is the rime not a constituent of moraic theory, it is not clear that it can even be stated as a regular expression over elements, at least in the general case. One might begin by referring to the span from the first moraic segment to the end of the syllable, $[_{\sigma}X^{*}(X_{\mu}X_{(\mu)}*)]_{\sigma}$. This is already awkward, but it gets worse. Many moraic theorists accept at least one of the following: The rime may contain no moraic segment (e.g. as with certain reduced vowels; Kager 1989), or the onset may contain a moraic segment (e.g. word-initial geminate; Topintzi & Zimmermann 2025). Another problem with allowing non-constituents as domains is that unless done in highly constrained way (e.g. Łubowicz 2005), learning is infeasible (cf. Pater 2016:15 on the unlearnability of conjunction even without this freedom). The learnability of the HG proposal here, by contrast, is straightforward. That is, there is no need for the learner to construct conjoined constraints, and algorithms exist to optimally weight given constraints (e.g. Goldwater & Johnson 2003).

Finally, given the unavailability of the rime in moraic theory, one might argue for a different theory of prosody. After all, some theories, such as X-slot theory (Levin 1985), do formalize the rime. This article is hardly the place to adjudicate between prosodic theories.[29] One could imagine a version of moraic theory with a rime node, but that would overhaul the theory; McCarthy & Prince (1986/1996:56–61) and Hayes (1989:298f), among others, reject such a move. Indeed, the rejection of the rime is a point of agreement between moraic phonologists and CV phonologists (Clements & Keyser 1983; Scheer & Szigetvári 2005). Another alternative to moraic theory is interval theory, which replaces the syllable by the span between successive vowel beginnings (i.e. rime plus following onset, if any; Steriade 2019). The interval would work as a domain of conjunction for Tamil, as in *LONGVOWEL $\&_{interval}$ *SONGEM (replacing *3μ with *LONGVOWEL because interval theory does not recognize the mora). This constraint correctly penalizes V:N: but not N:V:, as N: in the latter is fully excluded from the interval. It follows that the present case might be construed as an argument for interval theory. Nevertheless, OT-LC with intervals would not work for other cases, such as English **spip* in §10, as the word-initial onset is not part of any interval. A fuller consideration of moraic vs. interval theory is beyond the scope.

[29] For overviews of the argumentation that led to the widespread adoption of moraic theory starting in the 1980s, see Hayes (1989), Broselow (1995), Davis (2011), and Odden (2011).

8.6 On covert conjunction

An anonymous reviewer brings up the issue that it is not always formally clear whether a given constraint is simple vs. a conjunction. One might argue, for instance, that *SONGEM—critical to my proposal—is really the conjunction *SON $\&_{seg}$ *GEM. One response is that *SON (whether defined as $*[+sonorant]$ or $*[-syllabic, +sonorant]$) is dubious as an independent constraint. Sonorancy seems not to be marked in a context-free sense.[30] Another response is that even if *SONGEM is decomposable, it is incidentally so; admitting some such substantively motivated constraints is still more restrictive than admitting local conjunction as a general mechanism.

The same might be said of $*3\mu$: Perhaps this constraint is covertly a self-conjunction, namely, $*\mu^3_\sigma$. The same responses apply. While it is clear that trimoraic syllables are marked ($*3\mu$ is widely employed, not just by Prince & Smolensky 1993/2004, but before and since), it is unclear that moras are marked in general (i.e. unconjoined $*\mu$, as distinct from constraints on segment types being moraic such as $*C_\mu$ and $*T_\mu$). Indeed, it is not even clear that bimoraicity *per se* is marked (hypothetical $*2\mu$), though of course long vowels and codas are marked. Thus, once again, $*3\mu$ is not a cumulative 'worst of the worst' situation (i.e. independently marked elements becoming more marked in the combination), but rather markedness emerging only in the combination. Pater (2016:15) likewise argues that while *CODAVOICE and *[x] could be regarded as being conjunctions, there are good reasons not to do so.

9 Discussion

The proposed analysis of Tamil geminate phonotactics posits no new constraints or machinery. It relies on two markedness constraints, both with broad independent support, namely, *SONGEM and $*3\mu$. It employs weighted constraints (HG), which are also widely supported (e.g. Hayes & Wilson 2008; Zuraw & Hayes 2017; Smith & Pater 2020; Flemming 2021; Breiss & Albright 2022). It makes the least controversial assumptions possible about prosodic structure. These instruments in concert permit a simple analysis of the prohibition on V:(C)N:, which is that *SONGEM and $*3\mu$ gang up against faithfulness (MAX-μ).

The interest is therefore not in the analysis *per se*, but in three ramifications of the analysis. First, it demonstrates that attested gang effects which can be analyzed using cumulativity in HG cannot always be analyzed using conjunction in OT. Second, it shows that co-occurrence restrictions on geminates and long vowels cannot always be attributed (exclusively) to weight (cf. Davis 2011). Third, as elaborated in §9.3, it contributes to the debate on whether phonotactic ill-formedness should be analyzed using MPARSE. MPARSE fails in at least two ways here, both misgenerating Tamil and failing to connect the lexical gap to repairs in sandhi and diachrony.

9.1 Local conjunction both overgenerates and undergenerates

Most previous criticism of OT-LC focuses on overgeneration. To give one example, the grammar in (39–40) devoices obstruents iff there is a coda anywhere in the word (Pater 2016:16f). Analysts

[30] Central Rotokas is sometimes cited as a language without sonorant consonants, assuming a particular phonemic analysis (Robinson 2006:207). But sonorant consonants (not to mention vowels, which are also $[+son]$) regularly surface in that language; for example, /d/ is 'typically realized as a flap' (*ibid.*).

such as Łubowicz (2005) and Wolf (2007:19) have therefore proposed restricting the locality of conjunction in order to rein it in. But OT-LC's overgeneration is not confined to locality problems. For instance, even with a shared domain of the segment, it can generate a language with contrastive voicing in codas but not onsets, via IDENT(voi) $\&_{seg}$ NOCODA (Ito & Mester 1998; see Hayes & White 2015:283f for additional examples and Łubowicz 2005:6 for a response). Conjunction has also been critized on learnability grounds, including the explosion of parameters, the need for conjunctions whose parts are unmotivated, and the risk of learners overexploiting conjunction (Jesney 2005:38f). It is also criticized on conceptual grounds, as it undermines OT's core principle of strict domination (Kager 1999:400, McCarthy 2002:18).[31]

(39)

/balata/		NOCODA $\&_{PWd}$ *D	IDENT(voi)	*D	NOCODA
☞	a. balata			*	
	b. palata		*!		

(40)

/balatak/		NOCODA $\&_{PWd}$ *D	IDENT(voi)	*D	NOCODA
	a. balatak	*!		*	*
☞	b. palatak		*		*

Less has been said about OT-LC's undergeneration compared to HG (with the exception of Walker 2017, on which, see below). This article is one of the first treatments of a case in which the HG grammar X $>$ Y, Z (where Y and Z are weighted to gang up against X) cannot be analyzed by the OT-LC grammar Y&Z $\gg$ X $\gg$ Y, Z. That said, the possibility of this kind of undergeneration on the part of OT-LC has been recognized. For instance, Pater (2016:18) writes, 'As McCarthy (2002) points out, HG does not require constraints in a gang effect to share a domain, and it is possible to construct patterns that only HG generates (for example, on the basis of cumulative effects of MAX and markedness constraints, which Moreton & Smolensky (2002) claim are impossible in OT-LC)'.

Neither Pater (2016) nor McCarthy (2002:e.g. 18, 177) offer a concrete example (real or hypothetical), but I construct one in (41–43). Say that AGREE requires immediately adjacent vowels to agree in nasality. In this grammar,[32] intervocalic /k/ deletes unless doing so introduces a violation of AGREE; that is, as a rule, k $\rightarrow$ Ø / $V_{\alpha\text{nas}}$__$V_{\alpha\text{nas}}$. HG generates this scenario, as shown. In OT, by contrast, no ranking of these three constraints generates this scenario, even though it can be expressed as a single rule (and thus is not classically opaque). Nevertheless, this kind of interaction (do X unless it creates a new violation of Y) is in general natural, and McCarthy (2003a) proposes Comparative Markedness (CM) to implement it. With OT-CM, $\text{AGREE}_{\text{N}} \gg$ *VkV $\gg$ MAX implements the case. Thus, it is not clear that HG (without CM or LC) makes an unwanted prediction in generating such cases; indeed, OT needed to be patched precisely to generate them.[33]

[31] Conjunction has further been criticized for its use in analyzing derived environment effects and counterfeeding opacity (Wolf 2007:24, with references), but generally not in a way that pits OT against HG. Jesney (2009:297) treats gang effects in positional licensing, but allows that OT-LC could cover them.

[32] Assume that all non-shown faithfulness constraints are highly weighted.

[33] As another example, Moreton & Smolensky (2002) claim that MAX and DEP cannot be conjoined, having intrinsically different domains. McCarthy (2002:177) presents a hypothetical case of opacity solved by MAX &

(41)

	/eko/	$\mathscr{H}$	*VkV 3	MAX 2.5	AGREE 2
	a. eko	-3	-1		
☞	b. eo	-2.5		-1	

(42)

	/ẽko/	$\mathscr{H}$	*VkV 3	MAX 2.5	AGREE 2
☞	a. ẽko	-3	-1		
	b. ẽo	-4.5		-1	-1

(43)

	/ẽo/	$\mathscr{H}$	*VkV 3	MAX 2.5	AGREE 2
☞	a. ẽo	-2			-1
	b. ẽ	-2.5		-1	

It therefore remains unclear whether HG overgenerates by allowing MAX to gang up with markedness or with DEP, or by not imposing domains on gang effects. To be sure, HG overgenerates in other ways. For instance, with gradient constraints, it yields what Pater (2016:30–4) terms 'unbounded trade-offs' (or what Ryan 2017:335 calls 'cutoff-point pathologies'). Likewise, with scalar constraints, it overgenerates what Pater (2016:26) terms 'sufficient reward thresholds' (e.g. the relative weight pathology described by Ryan 2019:243).[34] These pathologies are addressed in part by redefining constraints (e.g. to be categorical; McCarthy 2003b) or by couching HG in serialism. At the same time, counting cumulativity is often used as an argument *for* HG (e.g. Ryan 2014; McPherson & Hayes 2016; Breiss 2020; Kawahara & Breiss 2021; Kim 2022). This article is thus far from the first to argue that OT undergenerates relative to HG, but is among the first to argue more specifically that LC undergenerates relative to cumulativity.

Walker (2017) is the only other publication of which I am aware that uses real language data to make the latter point—that HG can exploit a gang effect to generate a case that OT-LC with the same constraints (plus conjunctions) cannot analyze. In Yakut, rounding spreads progressively between vowels unless the first is high and the second is non-high. This spreading is local in the sense that a round non-high vowel does not spread its rounding to another non-high vowel across an intervening round high vowel. Walker (2017) analyzes the case as a gang effect in HG, such that GESTURALUNIFORMITY[round] (penalizing spreading round across vowels that differ in height) gangs up with *ROLO (penalizing a round non-high vowel) against SPREAD[round]if[-high] (self-explanatory). This analysis correctly accounts for locality. The OT-LC reanalysis (conjoining GESTURALUNIFORMITY[round] and *ROLO) works for adjacent vowels but incorrectly generates long-distance spreading (*ibid.*:11). Walker (2017:12f) discusses previously proposed non-HG analyses, arguing that they involve less well-motivated constraints, among other problems.

DEP. However, with HG and no conjunction, no weighting of the constraints employed by McCarthy (2002) can generate the opaque scenario that he describes. Thus, even if the conjunction of MAX & DEP overgenerates (as Moreton & Smolensky 2002 suggest), it is not clear that HG with MAX and DEP overgenerates in the same way.

[34] 'Stress shifts from syllable x to y only when y exceeds x by at least z units' (*ibid.*).

9.2 On moraic inconsistencies: the distribution of geminates is not always weight-based

As discussed in §4, constraints on the co-occurrence of geminates and long vowels are usually attributed to weight (e.g. Davis 2011:119–26). I have argued that weight-related constraints (e.g. *3μ) can gang up with other, non-weight-related constraints (e.g. *GEM or *SONGEM).[35] This approach has an additional benefit not yet mentioned: It resolves certain MORAIC INCONSISTENCY problems, whereby different phenomena within the same language appear to diagnose moraicity differently (Gordon 2006). For example, in Koya, codas appear to be moraic for stress, but not for syllable structure. Topintzi & Zimmermann (2025:328) imply that different systems in Koya might simply assess moras differently. Ganging, however, allows moraicity to be consistent: Codas can be moraic across the board, while V:C: is ruled out because *3μ and *GEM gang up against MAX-μ. Indeed, Tamil itself might be regarded as a case of moraic inconsistency, since the existence of V:(C)T: might be taken to imply that T: is non-moraic (although that is not my analysis), while other evidence supports the moraicity of T:. While presumably not a panacea, it remains to be explored how far gang effects go towards resolving apparent moraic inconsistencies cross-linguistically.

9.3 On MPARSE as a model of phonotactic illformedness

Another approach to phonotactics, not employed here, is to treat lexical gaps as competitions in which the null parse (•) wins (e.g. Breiss & Albright 2022; Cabrera 2025). For the present case, the analysis of the gang effect could remain intact (same weights), but MAX-μ in §1 could be replaced by MPARSE, which penalizes the null parse, as in (44–46). Instead of showing candidates with shortening/degemination, these tableaux show the null parse.

(44)

	/en̲n̲a/	$\mathscr{H}$	MPARSE 3	*SONGEM 2	*3μ 2
☞	a. en̲n̲a	−2		−1	
	b. •	−3	−1		

(45)

	/mu:kku/	$\mathscr{H}$	MPARSE 3	*SONGEM 2	*3μ 2
☞	a. mu:kku	−2			−1
	b. •	−3	−1		

(46)

	/e:n̲n̲a/	$\mathscr{H}$	MPARSE 3	*SONGEM 2	*3μ 2
	a. e:n̲n̲a	−4		−1	−1
☞	b. •	−3	−1		

[35] Compare also non-weight-driven geminate co-occurrence restrictions (Markoulaki 2018).

The biggest problem with MPARSE here is that it breaks the co-relevance constraint on ganging otherwise intrinsic to HG (Pater 2016:17), and as a consequence fails to generate Tamil correctly. In particular, with the MPARSE analysis, even if the superheavy and sonorant geminate occur in unrelated parts of the word, the word will be (erroneously) rendered illformed, as with [inna:ʈʈu] 'this position' in (47). Compare the (desired) symmetric trade-off in (12) above to the (undesired) asymmetric trade-off in (47) here. Under my HG analysis without MPARSE, co-relevance emerges in the sense that degemination is triggered only when it simultaneously obviates violations of both *SONGEM and *3μ, which is only possible when those violations occur in the same syllable. In just that case, a single violation of MAX-μ (shortening a single segment) solves both problems. With HG with MPARSE, by contrast, • wins as long as *SONGEM and *3μ are both violated, regardless of where the violations occur.

(47)

/inna:ʈʈu/	$\mathscr{H}$	MPARSE 3	*SONGEM 2	*3μ 2
☜ a. inna:ʈʈu	−4		−1	−1
💣 b. •	−3	−1		

Note that because (47a) is the fully faithful candidate, it is tied with • on all faithfulness constraints. Therefore, the weights of faithfulness constraints (not shown) are moot. Moreover, increasing the weight of MPARSE (to more than 4) to fix (47) would break (46), where • is the needed winner on this counteranalysis.

In conclusion, an MPARSE-based analysis in HG, like OT-LC (with or without MPARSE), fails to generate the Tamil case. It arguably fails in a second way as well. If null candidates are available for single-word inputs, they should also be available for multi-word inputs. However, in sandhi, V:(C)N: is repaired by shortening. Similarly, in historical change, newly arising V:(C)N: violations are repaired by shortening (§3). Thus, speakers do repair violations, they just lack the opportunity to do so in existing words. For the non-MPARSE model pursued here, this is not a problem. The existing lexicon lacks V:(C)N:, so it is moot whether it would be repaired; there is no harm in predicting that it *would* be. But in any context in which the marked configuration would arise (due to sandhi, diachrony, or, hypothetically, loanword adoption), it is handled properly. The model without MPARSE might seem to miss a behavioral fact, namely, that speakers presumably recognize that a word containing V:(C)N: is illicit. But in general, recognizing that a word is illicit is not incompatible with repairing the word if it does occur.[36] For other pros and cons of MPARSE, including learnability problems (NB. • is not observed by the learner), see Breiss & Albright (2022:7, 24f) and Cabrera (2025:8f). The larger role of MPARSE in phonological theory (e.g. in accounting for paradigm gaps) is beyond the scope.

[36] For example, if one asks an American English speaker, 'Would you normally pronounce l-a-t-e-r as [leɪtɻ]?', most would reply in the negative, recognizing that [leɪtɻ] is (largely) illicit. Of course, when /leɪtɻ/ is submitted to their grammar, it yields [leɪɾɻ], not •.

10 Conclusion

In most cases, if a gang effect can be implemented using cumulativity in HG, it can also be implemented via local conjunction in OT (OT-LC); one need only conjoin the ganging constraints. One might therefore be tempted to conjecture that this implication always holds (at least for non-gradient phenomena). However, as McCarthy (2002:18) and Pater (2016:18) note (without giving any specific examples), there are in principle gang effects analyzable by HG but not by OT-LC. After all, OT-LC imposes restrictions on domains (however defined) and on the types of constraints that can be conjoined (§9.1). HG ganging, by contrast, is unconstrained in these ways; gang effects are restricted only by the architecture itself: The constraints involved must be co-relevant (Pater's 2009 asymmetric trade-off).[37]

This article presents an attested gang effect that can be modeled with cumulativity in HG but not with conjunction in OT. Most previous criticism of OT-LC focuses on its overgeneration. While the present article is not the first to argue that OT-LC also undergenerates (§9.1), it is one of the first to demonstrate that OT-LC undergenerates gang effects (not just in principle, but real cases) that HG has no trouble generating, another potential case being Walker (2017:11).

In Tamil, both obstruent and sonorant geminates occur, but the latter cannot follow a long vowel within the syllable. That is, descriptively, *Vː(C)Nː. As I argue (§4), this prohibition cannot be analyzed as being driven exclusively by *3μ. Rather, it reflects the cumulative markedness of *SONGEM and *3μ. The components of this analysis—the constraints, prosodic structures, and use of weighted constraints—are all relatively uncontroversial. What is novel is the demonstration that (1) these simple parts, working together, generate gang effects that OT-LC cannot generate and (2) such cases are attested (in Tamil, and likely other languages as well; §7). The problem for conjunction is that no domain (e.g. syllable) captures the correct generalization (§8).

OT is therefore not a viable model of phonotactics, even within the realm of categorical, non-scalar phenomena. HG or a related regression/neural approach is needed.

Gang effects are well established in phonotactics, but are often not diagnostic between HG and OT-LC. For example, Albright (2012, with references; cf. also Pizzo 2015) finds that syllables of the type $sK_{\alpha}VK_{\alpha}$, where the consonants are, say, identical non-coronals (e.g. [spɪp, skʌk, smæm]), are productively underrepresented in English, more so than the modest markedness of the independent parts (sK and $K_{\alpha}VK_{\alpha}$) would predict, plausibly a gang effect of *sK and OCP(place). This case, unlike the Tamil case, is amenable to both HG (e.g. *sK and OCP(place) ganging up against MAX) and OT-LC (e.g. *sK $\&_{\sigma}$ OCP(place)). Similarly, Jesney (2009) argues that there are gang effects in Tamil phonotactics unrelated to the geminate phenomena treated by this article, and they too could be analyzed using either HG or OT-LC (though she presents other arguments for the superiority of HG for licensing, independent of the arguments in this article).

Some questions remain open. First, it remains to be seen whether the Tamil-like cases enumerated in §7 (or others like them) require the same analysis, and more generally whether OT-LC undergeneration relative HG arises via other constraints. Second, I argued in §9.2 that cumulativity has untapped potential to address moraic inconsistency problems, but it is unclear how much headway it might yield. Third, this article's results are not incompatible with conjunction in HG (Green & Davis 2014; Shih 2017; Milenković 2026). They suggest only that OT with LC is insufficient; the intrinsic cumulativity provided by HG is needed. Whether LC is also needed in HG is

[37] As discussed in §9.3, co-relevance is weakened if MPARSE is used.

an orthogonal question and area of ongoing research.

References

Albright, Adam. 2012. Additive markedness interactions in phonology. Handout, UCLA Linguistics Colloquium, January 13.

Albright, Adam & Bruce Hayes. 2003. Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition* 90. 119–161. https://doi.org/10.1016/s0010-0277(03)00146-x.

Annamalai, Elayaperumal & Sanford B. Steever. 1998. Modern Tamil. In Sanford B. Steever (ed.), *The Dravidian languages*, 100–128. London and New York: Routledge. https://doi.org/10.4324/9781315722580-4.

Bach, Emmon, Darin Howe & Patricia Shaw. 2005. On epenthesis and moraicity in Northern Wakashan. Paper presented at the Society for the Study of the Indigenous Languages of the Americas.

Baković, Eric J. 1996. Foot harmony and quantitative adjustments. MS., Rutgers Optimality Archive.

Becker, Michael, Seoyoung Kim & Filomena Sandalo. 2021. A prosódia do verbo kadiwéu. MS, University of Massachusetts, Amherst.

Blevins, Juliette & Sheldon P. Harrison. 1999. Trimoraic feet in Gilbertese. *Oceanic Linguistics* 38(2). 203–230. https://doi.org/10.1353/ol.1999.0012.

Breiss, Canaan. 2020. Constraint cumulativity in phonotactics: Evidence from artificial grammar learning studies. *Phonology* 37(4). 551–576. https://doi.org/10.1017/S0952675720000275.

Breiss, Canaan & Adam Albright. 2022. Cumulative markedness effects and (non-)linearity in phonotactics. *Glossa* 7(1). 1–32. https://doi.org/10.16995/glossa.5713.

Broselow, Ellen. 1995. Skeletal positions and moras. In John A. Goldsmith (ed.), *Handbook of phonological theory*, 175–205. Oxford: Blackwell.

Buckley, Eugene. 1991. Kashaya closed-syllable shortening and prosodic syllabification. In Dawn Bates (ed.), *The proceedings of the Tenth West Coast Conference on Formal Linguistics*, 65–74. Stanford: CSLI Publications.

Cabrera, Marisabel. 2025. One-shot vs. competitions phonotactics in modeling constraint cumulativity. *Proceeding of the Annual Meeting on Phonology (AMP)* 2023/2024. 1–10.

Chevillard, Jean-Luc. 2014. Metres in Tamil Bhakti literature and the problem of their (occasional) description in treatises (Studies on Tamil metrics–2). In Valérie Gillet (ed.), *Mapping the chronology of Bhakti: Milestones, stepping stones, and stumbling stones: Proceedings of a workshop held in honour of Pandit R. Varadadesikan. Varadadesikan*, vol. 124 Collection Indologie, 39–96. Pondicherry: Institut Français de Pondichéry and École française d'Extrême-Orient.

Christdas, Prathima. 1988. *The phonology and morphology of Tamil*: Cornell University dissertation.

Christdas, Prathima. 1996. Syllable prominence and stress in Tamil. MS., Cornell University.

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

Cohn, Abigail C. 2003. Phonological structure and phonetic duration: The role of the mora. *Working Papers of the Cornell Phonetics Laboratory* 15. 69–100.
Curtis, Emily. 2003. *Geminate weight: Case studies and formal models*: University of Washington dissertation.
Davis, Stuart. 1999. On the representation of initial geminates. *Phonology* 16. 93–104. https://doi.org/10.1017/s0952675799003711.
Davis, Stuart. 2011. Quantity. In John Goldsmith, Jason Riggle & Alan C. L. Yu (eds.), *The handbook of phonological theory, 2nd edition*, 103–140. New York: Blackwell Publishing. https://doi.org/doi.org/10.1002/9781444343069.ch4.
Faust, Noam & Shanti Ulfsbjorninn. 2025. The typology of weight in strict CV metrics: A challenge to moraic theory. *Linguistic Inquiry* 56. 1–60. https://doi.org/10.1162/ling.a.58.
Flemming, Edward. 2021. Comparing maxent and noisy harmonic grammar. *Glossa* 6(1). 1–42. https://doi.org/10.16995/glossa.5775.
Goldwater, Sharon & Mark Johnson. 2003. Learning OT constraint rankings using a Maximum Entropy model. In Jennifer Spenader, Anders Eriksson & Osten Dahl (eds.), *Proceedings of the Stockholm Workshop on Variation within Optimality Theory*, 111–120.
Gordon, Matthew. 2004. Positional weight constraints in Optimality Theory. *Linguistic Inquiry* 35(4). 692–703. https://doi.org/10.1162/ling.2004.35.4.692.
Gordon, Matthew. 2006. *Syllable weight: phonetics, phonology, typology*. New York, NY: Routledge Press. https://doi.org/10.4324/9780203944028.
Green, C. & Stuart Davis. 2014. Superadditivity and limitations on syllable complexity in Bambara words. In Ashley W. Farris-Trimble & Jessica A. Barlow (eds.), *Perspectives on phonological theory and development: In honor of Daniel A. Dinnsen*, 223–247. Amsterdam: John Benjamins. https://doi.org/10.1075/lald.56.17gre.
Hart, George & Hank Heifetz. 1988. *The Forest Book of the Rāmāyaṇa of Kampaṉ*. Berkeley: University of California Press.
Hart, Kausalya. 1999. *Tamil for beginners*. University of California, Berkeley: International and Areal Studies.
Hayes, Bruce. 1989. Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20. 253–306.
Hayes, Bruce & James White. 2015. Saltation and the P-map. *Phonology* 32(2). 267–302. https://doi.org/10.1017/s0952675715000159.
Hayes, Bruce & Colin Wilson. 2008. A Maximum Entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39. 379–440. https://doi.org/10.1162/ling.2008.39.3.379.
Ito, Junko & Armin Mester. 1998. Markedness and word structure: OCP effects in Japanese. MS., University of California, Santa Cruz.
Jäger, Gerhard & Anette Rosenbach. 2006. The winner takes it all – almost: Cumulativity in grammatical variation. *Linguistics* 44. 937–971. https://doi.org/10.1515/ling.2006.031.
Jesney, Karen. 2005. *Chain shift in phonological acquisition*: University of Calgary MA thesis.
Jesney, Karen. 2009. Licensing in multiple contexts: An argument for Harmonic Grammar. In M. Ryan Bochnak & al. (eds.), *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, vol. 45 1, 287–301. Chicago: Chicago Linguistic Society.
Kager, René. 1989. *A metrical theory of stress and destressing in English and Dutch*. Dordrecht: Foris.

Kager, René. 1999. *Optimality Theory*. Cambridge, U.K.: Cambridge University Press. https://doi.org/10.1017/cbo9780511812408.
Kamban. 1956. *Kampar Iyaṟṟiya Irāmāyaṇam*. Annamalai: Annamalai University Press. Critical edition of the *Kamparāmāyaṇam* (c. 1200 CE).
Kawahara, Shigeto. 2006. A faithfulness ranking projected from a perceptibility scale: The case of [+voice] in Japanese. *Language* 82(3). 536–574. https://doi.org/10.1353/lan.2006.0146.
Kawahara, Shigeto. 2007. Sonorancy and geminacy. *University of Massachusetts Occasional Papers in Linguistics* 32. 145–186.
Kawahara, Shigeto. 2011. Japanese loanword devoicing revisited: A rating study. *Natural Language and Linguistic Theory* 29. 705–723. https://doi.org/10.1007/s11049-011-9131-7.
Kawahara, Shigeto & Canaan Breiss. 2021. Exploring the nature of cumulativity in sound symbolism: Experimental studies of Pokémonistics with English speakers. *Laboratory Phonology* 12(1). 1–29. https://doi.org/10.5334/labphon.280.
Keane, Elinor. 2004. Illustrations of the IPA: Tamil. *Journal of the International Phonetic Association* 34(1). 111–116. https://doi.org/10.1017/s0025100304001549.
Kim, Seoyoung. 2022. A MaxEnt learner for super-additive counting cumulativity. *Glossa* 7(1). https://doi.org/10.16995/glossa.5856.
Kimball, Geoffrey. 1991. *Koasati grammar*. Lincoln: University of Nebraska Press. https://doi.org/10.2307/415814.
Kraehenmann, Astrid. 2003. *Quantity and prosodic asymmetries in Alemannic: Synchronic and diachronic perspectives*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110197228.
Krishnamurti, Bhadriraju. 2003. *The Dravidian languages*. Cambridge, U.K.: Cambridge University Press. https://doi.org/10.1017/cbo9780511486876.
Kung, Susan Smythe. 2007. *A descriptive grammar of Huehuetla Tepehua*: University of Texas, Austin dissertation.
Legendre, Géraldine, Yoshiro Miyata & Paul Smolensky. 1990. Harmonic grammar: A formal multi-level connectionist theory of linguistic well-formedness: Theoretical foundations. Technical Report, University of Colorado, Boulder, Department of Computer Science.
Lehmann, Thomas. 1998. Old Tamil. In Sanford B. Steever (ed.), *The Dravidian languages*, 75–99. London and New York: Routledge. https://doi.org/10.4324/9781315722580-3.
Levin, Juliette. 1985. *A metrical theory of syllabicity*: Massachusetts Institute of Technology dissertation.
Łubowicz, Anna. 2005. Locality of conjunction. In John Alderete, Han Chung-Hye & Alexei Kochetov (eds.), *Proceedings of the West Coast Conference on Formal Linguistics*, vol. 24, 254–262. Cambridge, MA: Cascadilla Press.
Markoulaki, Katerina. 2018. *Re-approaching Schneider's Law: Geminates and syllable margins*: Aristotle University of Thessaloniki MA thesis.
McCarthy, John. 2003a. Comparative markedness. *Theoretical Linguistics* 29. 1–51. https://doi.org/10.1515/thli.29.1-2.1.
McCarthy, John J. 2002. *A thematic guide to Optimality Theory*. Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9780511613333.
McCarthy, John J. 2003b. OT constraints are categorical. *Phonology* 20. 75–138. https://doi.org/10.1017/s0952675703004470.
McCarthy, John J. & Alan Prince. 1986/1996. Prosodic morphology 1986. *Linguistics Department Faculty Publication Series* 13. 1–99.

McPherson, Laura & Bruce Hayes. 2016. Relating application frequency to morphological structure: The case of Tommo So vowel harmony. *Phonology* 33(1). 125–167. https://doi.org/10.1017/s0952675716000051.

Milenković, Aljoša. 2026. Superadditive cumulativity in categorical prosodic patterns: Prosodic minimality in Bosnian/Croatian/Montenegrin/Serbian. *Phonology* 43(e4). 1–33. https://doi.org/10.1017/S0952675726100311.

Mohanan, Karuvannur P. 1982. *Lexical phonology*: Massachusetts Institute of Technology dissertation.

Morén, Bruce T. 1999. *Distinctiveness, coercion, and sonority: A unified theory of weight*: University of Maryland, College Park dissertation. https://doi.org/10.4324/9781315023779.

Moreton, Elliott & Paul Smolensky. 2002. Typological consequences of local constraint conjunction. In Line Mikkelsen & Chris Potts (eds.), *Proceedings of the 21st West Coast Conference on Formal Linguistics*, 306–319. Somerville, MA: Cascadilla Proceedings Project.

Newman, Stanley. 1944. *Yokuts language of California*. New York: Viking Fund.

Niang, Mamadou Ousmane. 1993. *Syllable "sonority" hierarchy and the stress patterns in Pulaar: A metrical approach*: University of Illinois, Urbana-Champaign dissertation.

Niang, Mamadou Ousmane. 1997. *Constraints on Pulaar phonology*. Lanham, Maryland: University Press of America. https://doi.org/10.2307/417503.

Niklas, Ulrike. 1993. *Amitacākarar iyaṟṟiya yāpparuṅkalakkārikai, Kuṇacākarar iyaṟṟiya uraiyuṭaṉ (The verses on the precious jewel prosody)*. Pondichéry: Institut Français de Pondichéry.

Nikolaev, Alexander. 2021. Rhotic degemination in Sanskrit and the etymology of Vedic ūrú- 'thigh,' Hittite UZU(u)walla- 'id.'. *Indo-European Linguistics* 9(1). 171–202. https://doi.org/10.1163/22125892-bja10011.

Nishimura, Kohei. 2003. *Lyman's Law in loanwords*: Nagoya University MA thesis.

Nishimura, Kohei. 2006. Lyman's Law in loanwords. *Phonological Studies* 9. 83–90.

Odden, David. 2011. The representation of vowel length. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, 1–26. New York: John Wiley & Sons, Ltd. https://doi.org/10.1002/9781444335262.wbctp0020.

Ohala, Manjari. 1983. *Aspects of Hindi phonology*. Delhi: Motilal Banarsidass.

Paradis, Carole. 1988. On constraints and repair strategies. *The Linguistic Review* 6. 71–97. https://doi.org/10.1515/tlir.1987.6.1.71.

Paramore, Jonathan Charles. 2025. Codas are universally moraic. *Phonology* 42(e1). 1–36. https://doi.org/10.1017/s0952675724000204.

Pater, Joe. 2009. Weighted constraints in generative linguistics. *Cognitive Science* 33. 999–1035. https://doi.org/10.1111/j.1551-6709.2009.01047.x.

Pater, Joe. 2016. Universal grammar with weighted constraints. In John J. McCarthy & Joe Pater (eds.), *Harmonic grammar and harmonic serialism*, 1–46. Bristol, CT: Equinox.

Pizzo, Presley. 2015. *Investigating properties of phonotactic knowledge through web-based experimentation*: University of Massachusetts, Amherst dissertation.

Podesva, Robert J. 2002. Segmental constraints on geminates and their implications for typology. Paper presented at the Annual Meeting of the Linguistic Society of America.

Prince, Alan & Paul Smolensky. 1993/2004. *Optimality Theory: Constraint interaction in Generative Grammar*. Malden, MA: Blackwell. https://doi.org/10.1002/9780470759400.

Rajam, V. S. 1992. *A reference grammar of classical Tamil poetry*. Philadelphia, Pennsylvania: American Philosophical Society. https://doi.org/10.70249/9798893981834.
Remijsen, Bert, Otto Gwado Ayoker & Signe Jørgensen. 2019. Ternary length in Shilluk. *Phonology* 36(1). 91–125. https://doi.org/10.1017/S0952675719000058.
Robinson, Stuart. 2006. The phoneme inventory of the Aita dialect of Rotokas. *Oceanic Linguistics* 45(1). 206–209. https://doi.org/10.1353/ol.2006.0018.
Ryan, Kevin M. 2011. Gradient syllable weight and weight universals in quantitative metrics. *Phonology* 28(3). 413–454. https://doi.org/10.1017/s0952675711000212.
Ryan, Kevin M. 2014. Onsets contribute to syllable weight: Statistical evidence from stress and meter. *Language* 90(2). 309–341. https://doi.org/10.1353/lan.2014.0029.
Ryan, Kevin M. 2017. Attenuated spreading in Sanskrit retroflex harmony. *Linguistic Inquiry* 48(2). 299–340. https://doi.org/10.1162/ling_a_00244.
Ryan, Kevin M. 2019. *Prosodic weight: Categories and continua*. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198817949.001.0001.
Ryan, Kevin M. 2020. VV > VC > V for stress: Coercion vs. prominence. *Linguistic Inquiry* 51(1). 124–140. https://doi.org/10.1162/ling_a_00325.
Ryan, Kevin M. 2021. The development of diphthongs in Vedic Sanskrit. *Journal of the American Oriental Society* 141(2). 289–297. https://doi.org/10.7817/jameroriesoci.141.2.0289.
Sampath Kumar, Srinivas. 2016. *The mora-constituent interface model*: Hong Kong Baptist University dissertation.
Sayeed, Ollie. 2017. Osthoff's law in Latin. *Indo-European Linguistics* 5. 147–177. https://doi.org/10.1163/22125892-00501005.
Scheer, Tobias & Péter Szigetvári. 2005. Unified representations for stress and the syllable. *Phonology* 22. 37–75. https://doi.org/10.1017/S0952675705000436.
Shih, Stephanie S. 2017. Constraint conjunction in weighted probabilistic grammar. *Phonology* 34(2). 243–268. https://doi.org/10.1017/s0952675717000136.
Smith, Brian W. & Joe Pater. 2020. French schwa and gradient cumulativity. *Glossa* 5(1). 24. https://doi.org/10.5334/gjgl.583.
Smolensky, Paul. 1993. Harmony, markedness, and phonological activity. Handout from Rutgers Optimality Workshop 1, Rutgers University, ROA-87.
Smolensky, Paul. 2006. Optimality in phonology II: Harmonic completeness, local constraint conjunction, and feature domain markedness. In Paul Smolensky & Géraldine Legendre (eds.), *The harmonic mind: From neural computation to Optimality-Theoretic grammar*, vol. 2, 27–160. MIT Press.
Smolensky, Paul & Géraldine Legendre. 2006. *The harmonic mind: From neural computation to Optimality-Theoretic grammar*. Cambridge, MA: MIT Press.
Steriade, Donca. 2019. CiV lengthening and the weight of CV. In Margit Bowler, Philip T. Duncan, Travis Major & Harold Torrence (eds.), *Schuhschrift: Papers in honor of Russell Schuh*, 161–177. eScholarship Publishing.
Tirumurukaṉ, Irā. 1987. *Kampaṉ pāṭiya vaṇṇaṅkaḷ*. Chennai: Vāṉati Patippakam.
Topintzi, Nina & Eva Zimmermann. 2025. Reconciling "heavy" and "long": The typology of lexical geminates. *Linguistic Inquiry* 56(2). 311–351. https://doi.org/10.1162/ling_a_00499.
Törkenczy, Miklós. 2004. *The phonotactics of Hungarian*: Hungarian Academy of Sciences dissertation.

Trubetzkoy, Nikolai S. 1939. *Grundzüge der Phonologie*. Prague: Travaux du Cercle Linguistique de Prague. https://doi.org/10.1515/9783110226300.4.388.
Tyler, Stephen. 1969. *Koya: An outline grammar*. Berkeley, CA: University of California Press.
Venkatakrishnan, Ramprashanth & Somdev Kar. 2025. Sonorant gemination in Old Tamil. *Bhasha* 4(1). 95–122. https://doi.org/10.30687/bhasha/2785-5953/2025/01/005.
Walker, Rachel. 2017. Cumulative constraint interactions: Violations and domains encompassing segmental spans. In Aaron Kaplan, Abby Kaplan, Miranda K. McCarvel & Edward J. Rubin (eds.), *Proceedings of the 34th West Coast Conference on Formal Linguistics*, 1–15. Somerville, MA: Cascadilla Proceedings Project.
Wentworth, Blake (translator). 2022. *Youth: The Balakandam of Kampan's Ramayana*. New Delhi: Speaking Tiger Books.
Wilden, Eva. 2018. *Grammar of Old Tamil for students*. Pondichéry: Institut Français de Pondichéry and École française d'Extrême-Orient.
Wolf, Matthew. 2007. What constraint connectives should be permitted in OT? *University of Massachusetts Occasional Papers in Linguistics* 36. 1–29. https://doi.org/10.7282/T33T9F5B.
Zec, Draga. 1995. Sonority constraints on syllable structure. *Phonology* 12(1). 85–129. https://doi.org/10.1017/s0952675700002396.
Zuraw, Kie & Bruce Hayes. 2017. Intersecting constraint families: An argument for Harmonic Grammar. *Language* 93. 497–548. https://doi.org/10.1353/lan.2017.0035.
Zvelebil, Kamil. 1989. *Classical Tamil prosody: an introduction*. Madras: New Era Publications.