Crossing Word Boundaries: Constraints for Misaligned Syllabification[1]

Caroline R. Wiltshire
University of Florida, Gainesville

1 Introduction

In this work, I examine a set of languages which appear to require resyllabification post-lexically; in less derivational terms, a word's syllabification in isolation differs from its syllabification in a phrase-internal context. Although many people, myself included, have been looking at such cases in isolation over the years, I bring together several examples here to see what features they share and how an Optimality Theory analysis improves upon rule-based derivational approaches.

I show that the interaction of word edges in phrases can be analyzed using alignment constraints in a monostratal Optimality Theory framework (henceforth OT, Prince and Smolensky 1993, McCarthy and Prince 1993a). Across-word syllabification results when constraints aligning word boundaries with syllables edges are outranked by constraints on well-formed syllable structures. By submitting entire phrases as input to syllabification, multiple levels of syllabification are unnecessary, in contrast to multi-level theories such as lexical phonology (Kiparsky 1982, Mohanan 1982) and multi-level OT (McCarthy and Prince 1993b). Furthermore, I show an advantage of the OT perspective: constraints for word-edge syllabification are not turned off, but merely overridden in cases in which phrasal position plays a role in syllabification. Such constraints can still exert themselves in the grammar in other circumstances, despite being outranked, which is exactly the prediction of the OT architecture.

2 Optimality Theory

I assume basic familiarity with a correspondence version of Optimality theory (McCarthy and Prince 1995), and will mention only a few relevant points here. The correspondence constraints include those in (1), parametrized for consonants, vowels, and features, which penalize any deviations between input and output forms.

(1) Correspondence Constraints (McCarthy and Prince 1995, p. 264)

(a) MAX-IO:	Every segment of input has a correspondent in output.
(b) DEP-IO:	Every segment of output has a correspondent in input
(c) IDENT(F)-IO:	Correspondent segments are identical in feature F.

[1] I thank the audience at the DGfS annual meeting, March 2000, in Marburg, for their comments on the oral version of this paper. In particular, I also thank T. Alan Hall and Bożena Cetnarowska for discussion afterwards. Any and all remaining errors of fact or interpretation are my own.

I formulate the interaction of words and syllables in terms of alignment constraints; the general form of such constraints is shown in (2):

(2) Generalized Alignment Constraint (McCarthy and Prince 1993a:2)
ALIGN (Cat1, Edge1, Cat2, Edge2) $=_{def}$
$\forall$ Cat1 $\exists$ Cat2 such that Edge1 of Cat1 & Edge2 of Cat2 coincide.
Where Cat1, Cat2 $\in$ Pcat U Gcat & Edge1, Edge2 $\in$ {Right, Left}

Alignment constraints are parametrized for various categories, whether prosodic or grammatical, and edges, either left or right. In languages whose words are syllabified without reference to phrasal context, the constraints given in (3) rank high.

(3) Constraints against Cross-word syllabification

(a) ALIGN(Wd,L,σ,L) = ALIGN-L(Wd, σ):	the left edge of each word aligns with the left edge of a syllable.
(b) ALIGN(Wd,R, σ,R) = ALIGN-R(Wd, σ):	the right edge of each word aligns with the right edge of a syllable.

When highly ranked, these constraints enforce the alignment of word and syllable boundaries, so that syllables do not straddle word boundaries. When lower ranked than other phonological constraints, however, constraints (3a) and (3b) can be violated, resulting in syllabification across words as the optimal output. The ranking of (3a-b) in an OT grammar will allow us to do both word and phrasal syllabification in a single stage of parallel constraint evaluation. In section 3, I examine four languages as case studies of the factors causing syllable/word misalignment; these show that constraints on each part of the syllable (onset, nucleus, coda) can be responsible for misalignment. In section 4, I show that two of these languages give evidence that the syllable/word alignment constraints play a role in phrasal syllabification, though they are outranked, as predicted in OT.

3 Phrasal Syllabification

3.1 Misalignment in Spanish

Spanish provides a straightforward example of syllabification across word boundaries caused by the requirement that syllables should have onsets (Harris 1983, 1993; Hualde 1992). The constraint requiring onsets (4) is familiar from the literature, and is widely attested cross-linguistically.

(4) ONSET: Every syllable begins with a consonant (McCarthy and Prince 1993a: 20)

In phrase-initial position, an onsetless syllable is tolerated, as in (5a). Word-internally, ONSET ensures that a single intervocalic consonant appears in onset rather than coda (5a-c). Phrase-internally, ONSET plays the same role, ensuring that a single intervocalic consonant is in onset position, as in (5d-f).

(5)	a. /asules/	[.a.su.les.][2]	"blue"
	b. /komida/	[.ko.mi.ða.]	"food"
	c. /kopa/	[.ko.pa.]	"cup, goblet"
vs.	d. /grandes#ojos#asules/	[gran.de.so.jo.sa.su.les.]	"big blue eyes"
	e. /asul#oskuro/	[.a.su.los.ku.ro.]	"dark blue"
	f. /klub#elegante/	[.klu.ße.le.ɣan.te.]	"elegant club"

As onsetless syllables are tolerated phrase-initially, correspondence constraints such as MAX-IO(V) and DEP-IO(C) outrank ONSET; otherwise, we would have vowel deletion (a MAX-IO(V) violation) or consonant epenthesis (a DEP-IO(C) violation) to resolve the lack of onset. In Tableau 1a, we see that the presence of the ONSET constraint favors the parsing of single intervocalic consonants into onset rather than coda position. The constraints aligning word and syllable boundaries play no role in word-internal evaluation; at phrase edges, syllable and word edges align.

Tableau 1a Phrase-initial and word-internal /asules/ "blue"

Candidates	Max-IO(V)	Dep- IO(C)	ONSET	Align-L (Wd, σ)	Align-R (Wd, σ)
.su.les.	*!<a>				
.ʔa.su.les.		*!(ʔ)		*	
.as.u.les.			**!		
☞.a.su.les.			*		

However, phrase-internally, syllabification crosses word boundaries in order to satisfy ONSET, resulting in violations of the constraints from (3), which are outranked. Thus, in Tableau 1b, we see that ranking these constraints lower than ONSET gives syllabification across words as the optimal result, despite the misalignments of words and syllables.

[2] I use '.' for syllable boundaries, and '#' for word boundaries. Underlining in (5d-f) is meant only to draw the eye to the crucial syllabification crossing word boundaries.

Tableau 1b Phrase-internal /grandes#ojos#asules/ "big blue eyes"

Candidates	Max-IO(V)	Dep- IO(C)	Onset	Align-L (Wd, σ)	Align-R (Wd, σ)
.gran.des.#jos.#su.les.	*!*<o,a>				
.gran.des. ʔ#o.jos. ʔ#a.su.les.		*!*(ʔʔ)		**	
.gran.des#.o.jos.#a.su.les.			*!*		
☞.gran.de.s#o.jo.s#a.su.les.				**	**

The overall ranking for Spanish is therefore the one shown in (6).

(6) { MAX-IO(V), DEP-IO(C) } >> ONSET >> { ALIGN-L(Wd, σ), ALIGN-R(Wd, σ) }

Syllabification within a phrase resembles syllabification within a word. Though the /s/ of *grandes* may be in a coda when the word is spoken in isolation, that is because a word in isolation is a phrase, and misalignment at the edges is not an option, even when important syllable phonotactics are at stake. The same /s/ of *grandes* can be an onset phrase-internally without resyllabification, so long as constraints evaluate entire phrases in parallel. Hence, distinct syllabifications do not entail multiple syllabification; both are attempting to satisfy the most important phonotactics at the expense of the least.

3.2 Misalignment in Italian

The second case study, Italian, offers two examples of the misalignment of syllables across word boundaries. Like Spanish, cross-word syllabification results from the interaction of syllable phonotactics with alignment constraints; unlike Spanish, where the requirement of a consonant in onset position was at issue, in Italian the dominant phonotactics limit the permissible onsets and codas. The analysis here is based on Wiltshire and Maranzana (1999).

The first example involves geminate consonants. Consonant length is generally distinctive word internally in Italian, but a few segments (e.g., [ts], [ʃ] and [ʎ], are always long except phrase-initially, as shown in (7a-c).

(7)	a.	[.faʃ.ʃa.]	"bandage"	vs.	*[faʃa]
	b.	[.ʃu.pa.to.]	"ruined"		*[ʃʃupato]
	c.	[.ca.saʃ.ʃu.pa.ta.]	"ruined house"		

I treat geminates as two consonantal slots here, though a moraic analysis is also possible (see Davis 1999 for a discussion of the representation of geminates). Geminates and other two obstruent clusters, when they appear word and phrase-internally, are syllabified as a coda plus onset. This results from the fact that standard Italian onsets are limited to clusters such as obstruent+glide and obstruent + liquid.[3] Italian onsets thus require an increase in sonority, based on a scale of sonority such as that in (8), proposed by Davis (1990) in his analysis of Italian onsets:

(8) Sonority Hierarchy for Italian (Davis 1990)

voiceless		voiced		non-cor		cor						glides		
stops	<	stops	<	frics	<	frics	<	n	<	m	<	liq	<	vow
1		2		3		4		5		6		7		8

Sonority sequencing refers to the increasing sonority before the syllable peak and decreasing sonority following the peak. Languages may impose a minimal sonority distance requirement on these increases and decreases (Steriade 1982, Selkirk 1984), such as a requirement that onset segments differ in sonority by some minimal amount. In OT, the minimal sonority distance requirements can be seen as a set of constraints universally ranked from least to most strict, with different languages differing in which of these constraints can be violated due to other constraints ranked in between.

(9) a. *EQUALSON — Syllable margins do not contain segments of equal sonority. (aka*<1DIFSON)

b. *<2DIFSON — Syllable margins do not contain segments that differ in less than 2 degrees of sonority.

c. *<4DIFSON — Syllable margins do not contain segments that differ in less than 4 degrees of sonority.

Universal ranking: *EQUALSON (*<1DIFSON)>>*<2DIFSON >>*<4DIFSON

In Italian, stops followed by liquids and glides make good onsets, but consonants of equal sonority are never permitted in onset position. The constraint *EQUALSON therefore ranks high, and, by this ranking, geminates, which consist of two consonants of equal sonority, are associated with the coda of one syllable and the onset of another word and phrase-internally.

In order for such syllabifications to be chosen as optimal word and phrase-internally, *EQUALSON must outrank the widely attested NOCODA constraint from (10), as one half of the geminate is forced into coda position.

(10) NOCODA: Syllables end with a vowel.

[3] Italian has a few other rare but permissible onset clusters, such as /pn/ and /kn/; the minimal sonority distance requirement of +4 will allow these as onsets according to Davis's scale in (8).

Tableau 2a illustrates that two other correspondence constraints must also rank high in Italian: DEP-IO(V), which bans epenthesis as a solution to an onset or coda cluster of consonants with equal sonority, and MAX-IO(C), which bans consonant deletion. Note that the alignment constraint on words and syllables is inactive in the word-internal case.

Tableau 2a Word-internal *fascia* "bandage" = [.faʃ.ʃa.]

Candidates	*Equal Son	Dep-IO (V)	Max-IO (C)	No Coda	Align-L/R (Wd,σ)
fa. ʃʃa.	*!				
faʃʃ.a.	*!			**	
fa. ʃa. ʃa.		*!(a)			
.fa. ʃa.			*!<ʃ>		
☞ faʃ.ʃa.				*	

Phrase-internally, as in Tableau 2b, we have evidence that the word/syllable alignment constraints are violable in order to satisfy *EQUALSON. The result is that word boundaries are ignored, and phrase-internal and word-internal syllabification look identical.

Tableau 2b Phrase-internal *casa sciupata* "house ruined" = casa[ʃ.ʃ]upata

Candidates	*Equal Son	Dep-IO (V)	Max-IO (C)	No Coda	Align-L/R (Wd, σ)
kasa.# ʃʃupata	*!				
kasa#ʃʃ.upata	*!			**	**
kasa.# ʃa. ʃa...		*!(a)			
kasa.# ʃa...			*!<ʃ>		
☞kasa#ʃ.ʃupata				*	*

A different result is seen in phrase-initial position, where there is no option for the word-initial geminate to be realized with its first half in a coda. Here the ranking gives us deletion of the word-initial consonant, so that MAX-IO(C) is clearly outranked by *EQUALSON and DEP- IO(V). Again, (mis)-alignment of words and syllables is not an option here, since there is no previous word to syllabify with.

Tableau 2c Phrase-initial *sciupata* "ruined" = [.ʃu.pa.ta.]

Candidates	*Equal Son	Dep-IO (V)	Max-IO (C)	No Coda	Align-L/R (Wd,σ)
.ʃʃu.pa.ta.	*!				
.iʃ.ʃu.pa.ta.		*!(i)			*
☞ .ʃu.pa.ta.			*<ʃ>		

Thus the behavior of geminates in phrase-internal vs. phrase-initial position appears different, but can be handled by the same set of constraints. The same high ranked syllable phonotactic, which limits onset and coda clusters, ensures that geminates are either split between two syllables (both within and across words) or shortened to a single consonant (phrase-initially).

The second example of misalignment in Italian involves word-initial clusters of /s/ plus a consonant and the doubling of initial consonants known as *raddoppiamento sintattico* (Chierchia 1986, Saltarelli 1970, Vogel 1977). When a word ends in a stressed vowel, the consonant beginning the following word may be doubled (11a-b). I will return to the truly doubling types later, but first note that in (11c-d), misalignment makes the standard doubling of an initial consonant unnecessary.

(11) a. /pulita/	[p]ulita	"clean"	cittá[pp]ulita	"a clean city"
b. /triste/	[t]riste	"sad"	cittá[tt]riste	"a sad city"
c. /ʃupata/	[ʃ]upata	"ruined"	cittá[ʃʃ]upata	"a ruined city"
d. /sporko/	[sp]orco	"filthy"	cittá[sp]orca	"a filthy city"

In Wiltshire and Maranzana (1999), we analyzed the phenomenon as the effect of the constraint PKPROM, which motivates misalignment or insertion of an initial consonant in order to make a stressed syllable heavy:

(12) PKPROM: x is a more harmonic stress peak than y if x is heavier than y.
(Prince and Smolensky 1993: 39)

Thus, a form like **cit.tá.pulita*, with a light stressed syllable *tá*, is less harmonic than one like *cittáp.pulita,* which has the stressed syllable closed with a consonant; in some cases the consonant closing the stressed syllable is epenthetic, as in (11a-b), and in other cases, it is underlying, as in (11c-d).

Looking more closely at the (11d) case, we see that /sC/ clusters are tolerated phrase-initially, despite violating the constraint *<4DIFSON from (9). Thus, *<4 DIFSON must be outranked by the correspondence constraints MAX-IO(C) and DEP-IO(V), leaving its violation as optimal to the alternatives, phrase-initially.

Tableau 3a: Phrase-initial sC *specchio* "mirror" = [.spek.kj.o]

Candidates	*Equal Son	Dep-IO (V)	Max-IO (C)	*<4Dif Son	No Coda
.is.pek.kjo		*!(i)			**
.pek.kjo			*!<s>		*
☞ .spek.kjo.				*	*

In phrase-internal position, however, *<4DIFSON can be satisfied where possible by syllabification of the /s/ into coda position with a preceding word-final vowel. The word-syllable alignment constraint, as well as NOCODA, therefore ranks lower than *<4DIFSON.

Tableau 3b: Phrase-internal sC *cittá sporca* "filthy city" = [cittás.porka]

Candidates	*Eq Son	Pk Prom	Dep-IO	*<4Dif Son	No Coda	Align-R/L (Wd,σ)
cittá.#sporka		*!		*		
cittás#.sporka			*!	*		
☞ cittá#s.porka					*	*

PEAKPROM is included in Tableau 3b to show that it is satisfied in such cases, so that the doubling seen in *raddoppiamento sintattico* is unnecessary. I return to the doubling of (11a-b) in section 4.2.

Thus, we have seen that, in Italian, syllables cross word boundaries in order to satisfy onset restrictions against geminates and /sC/ clusters, as well as correspondence constraints DEP-IO(V) and MAX-IO(C). In both Italian and Spanish, we are seeing syllable phonotactics on the onset, whether requiring or restricting them, drive syllabification across words. These two cases also involve the syllabification of an entire segment with material from a different word; the segment can be any consonant in Spanish and the /s/ of /sC/ clusters in Italian. For the geminates in Italian, it is possible that less than a full segment is spread, depending on whether a geminate is considered to be two C-slots or a consonant with a mora. In the next two case studies, we see clearer cases of subsegmental misalignment across words, involving moras and features.

3.3 Mora Misalignment in Luganda

Luganda shows two types of compensatory lengthening (CL) which apply within morphemes, across morpheme boundaries, and across word boundaries within a phonological phrase (Clements 1986, Herbert 1975, Tucker 1962). In the first, prenasalization lengthening, nasals which are preceded by vowels and followed by stops or fricatives surface as prenasalization on the following consonant, while the preceding vowel is realized as long, as shown in (13):

(13) Prenasalization lengthening

	a. /ku+linda/	[kulii^{n}da]	"to wait"
	b. /mu+lenzi/	[mulee^{n}zi]	"boy"
	c. /mu+ntu/	[muu^{n}tu]	"person"
	d. /ba+ntu/	[baa^{n}tu]	"people"
	e. /#buta+lab+a# #njovu/	[butalabaa^{n}jovu]	"to not see elephants"
	f. /#si+agala##mva/	[saagalaa^{m}va]	"don't like vegetable relish'
c.f.	g. /mva/	[m̩va]	"vegetable relish"

The second type of CL, glide formation lengthening, results when a high vowel is followed by a vowel in another morpheme; the first vowel is realized as the corresponding glide, while the second is realized as a long vowel, as in (14).

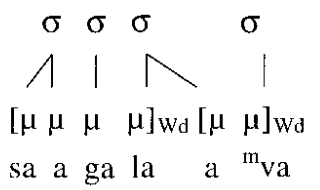
(14) Glide Formation lengthening

a. /li+ato/	[l^{y}aato]	"boat"
b. /ki+uma/	[k^{y}uuma]	"metal object"
c. /mu+oyo/	[m^{w}ooyo]	"soul"
d. /mu+iko/	[m^{w}iiko]	"trowel"
e. /o+lu+naku##o+lu+o/	[olunak^{w}ool^{w}o]	"that day"
f. /a+ba+kulu##a+ba+o/	[abakul^{w}aabo]	"those elders"

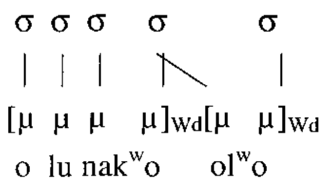
To see how the two forms of compensatory lengthening involve misalignment of a mora across a word boundary, consider the structures in (15):

(15) Subsyllabic segment crosses word boundaries:

a.



b.



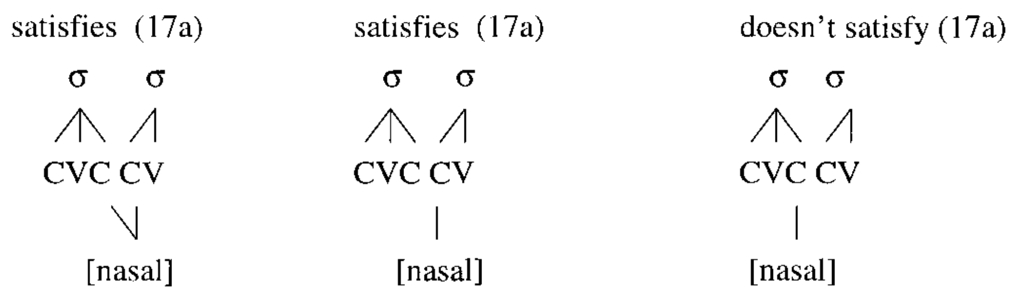
In (15a), the mora from the /m/ of /mva/ is syllabified with the preceding vowel of the preceding word, making it long, though the mora is part of the underlying form of the second word. Similarly, in (15b), the mora of the word-final vowel is realized in a syllable that contains the initial vowel of the following word.

In Wiltshire (1999), I connected the two types of CL in an OT account involving the satisfaction of the correspondence constraints in (16a-c), while violating the constraint (16d). That is, moras and features are preserved at the expense of the input location of some mora.

(16)		
	a. MAX-IO(μ)	Every mora of the Input has a correspondent in the Output. (Rosenthall 1997)
	b. MAX-IO([nas])	Every instance of [nasal] in the Input has a correspondent in the Output.
	c. MAX-IO([V-feat])	Every instance of [V-feat] in the Input has a correspondent in the Output.
	d. IDEN-IO(μ)	Correspondent segments in Input and Output have identical values for weight. (Rosenthall 1997)

The driving force behind the prenasalization is (17a); a coda condition to capture the fact that Luganda has no coda nasals unless they are in the first half of a geminate. This constraint is to be understood as satisfied by non-crisp alignment (Itô and Mester 1994), which means that so long as the feature [nasal] does align with the left edge of some syllable, it may also be associated with other segments not at the left edge. Examples of structures satisfying and violating (17a) are given below in (17b). Violation occurs when a [nasal] feature attaches only to a segment at the right edge of a syllable, i.e., in the coda.

(17) a. Align-Left ([nasal], σ): The feature [nasal] is aligned with the left edge of a syllable (i.e., onset position licenses the feature [nasal]).



Prenasalized stops in the output satisfy ALIGN-L(nasal), since the feature [nasal] is associated with the initial segment of a syllable. The high ranking of this constraint, along with MAX- IO(nasal), forces the nasal of the input to attach itself to the following onset. Ranking the correspondence constraint MAX-IO(μ) above IDEN-IO(μ) preserves the mora from the input nasal, but allows it to be attached to the preceding vowel. In the word-internal and phrase-initial cases, once again alignment of words with syllables plays no role.

Tableau 4 Word-internal /mu+ntu/ "person" = [muu^{n}tu]

Candidates	Align-L (nasal)	Max-IO (μ)	Max-IO (nasal)	Iden-IO (μ)	Align-L/R (Wd, σ)
.mun.tu.	*!				
.mu.^{n}tu		*!			
.muu.tu			*!	*	
☞ .muu.^{n}tu.				*	

Tableau 5a Phrase-initial /mva/ "vegetable relish" = [m̩va][4]

Candidates	Align-L (nasal)	Max-IO (μ)	Max-IO (nasal)	Iden-IO (μ)	Align-L/R (Wd, σ)
.^{m}va.		*!			
☞ .m.va.					

In the phrase-internal case, however, in order to have the same prenasalization, the alignment of words and syllables must be violated. The ranking is illustrated in Tableau 5b:

Tableau 5b Phrase-internal /si+agala # mva/ "don't like vegetable relish" = [saagalaa^{m}va]

Candidates	Align-L (nasal)	Max-IO (μ)	Max-IO (nasal)	Iden-IO(μ)	Align-L/R (Wd, σ)
saa.ga.la#m.va	*!				*
saa.ga.la.^{m}#va		*!			
saa.ga.laa.#va			*!	*	
☞ saa.ga.laa.m#va				*	*

Glide-formation compensatory lengthening follows basically the same logic, with the major difference being that the driving force is a constraint against diphthongs, as in (18).

---

[4] While the syllabic nasal wins phrase-initially, where there is no option of preservation of the mora by association with a preceding vowel, presumably a high ranking constraint against syllabic nasals prevents this option from winning phrase-medially.

```
(18) No Diphthongs (NoDiph)        * σ          (Rosenthall 1997)
                                    / \
                                   μ   μ
                                   |   |
                                   Vi  Vj
```

By the ranking of this constraint above IDEN-IO(μ), when two vowels are in hiatus word-internally, the diphthong is avoided but input moras are preserved. As shown in Tableau (6a), alignment is vacuously satisfied in the word-internal case; however, Tableau (6b) reveals that the constraint against misalignment must again rank low so that the same result is found phrase-internally.

Tableau 6a Word-internal /li+ato/ "boat" = [l^yaato]

Candidates	No Diph	Max-IO (μ)	Max-IO (V-feat)	Iden-IO (μ)	Align-R/L (Wd, σ)
.lia.to.	*!				
.l^ya.to.		*!			
.laa.to.			*!	*	
☞ .l^yaa.to.				*	

Tableau 6b Phrase-internal /o+lu+naku # o+lu+o/ "that day" = [olunak^{w}ool^{w}o][5]

Candidates	No Diph	Max-IO (μ)	Max-IO (V-feat)	Iden- IO(μ)	Align-R/L (Wd, σ)
.na.ku#o.l^{w}o	*!*				*
na.k^w#o.l^{w}o		*!*			*
.na.ko#o.l^{w}o			*!*	**	*
☞ na.k^w#oo.l^{w}o				**	*

Thus, for both types of compensatory lengthening, the phonotactic constraints on well-formed syllables, NODIPH and ALIGN-L(nas) rank high to motivate the difference between input and output. The relative ranking of MAX-IO(μ) above IDEN-IO(μ) allows for the preservation of the mora in a new location, while the low ranking of ALIGN-L/R(Wd, σ) allows for that

[5] A high ranking constraint prevents long-vowels from appearing phrase-finally.

preservation even at the cost of misalignment of the syllable and word boundaries. The overall ranking is thus:

(19) {MAX-IO(μ), NODIPH, ALIGN-L(nas), MAX-IO(nas),MAX-IO(V-feat)}
>> {IDEN-IO(μ), ALIGN-L/R(Wd, σ)}

Note that both types of compensatory lengthening require only general cross-linguistically motivated constraints. In each case, syllabification crosses words to satisfy MAX-IO(μ) plus a syllable well-formedness constraint, either on the coda (ALIGN-L(nas)) or nucleus (NODIPH). The following example shows similarly that sub-segmental units can be syllabified with a different word due to coda constraints; in this case, features rather than moras are misaligned.

3.4 Feature Misalignment in Tamil

Features are a second kind of subsegment that can be shared across word edges due to syllable phonotactics. In Tamil, coda constraints force adjacent word-final and word-initial consonants to share features of place of articulation.[6] In examples (20a-d), we see words with plural suffixes or emphatic clitics; examples (20e-g) are compounds (Christdas 1988). In both cases, word-final nasals assimilate in place to the following obstruent.

(20)	a. /maram+kaɭ/	tree + pl	[mʌrəŋgə]	"trees"
	b. /maram+ taan̲/	tree + emph	[mʌrən̲dã]	"tree" (emph)
	c. /pasan + kaɭ/	child + pl	[pʌsəŋgə]	"children"
	d. /vayal + taan/	field + emph	[vʌjəl̲dã]	"field" (emph)
	e. /paɳam#kaas/	money # cash	[pʌɳəŋkasɯ]	"money"
	f. /maram# tʃeʈi/	tree # plant	[mʌrəɲtʃɛɖɪ]	"vegetation"
	g. /koɭam#tooɳʈi/	pond # dredge	[kɔɭən̲toɳɖɪ]	"tool for dredging ponds"

In phrases, we see the same phenomenon of nasal place assimilation across words, though phrase-final nasals are deleted (Wiltshire 1998).

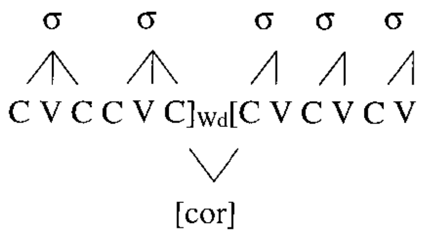
(21)	a. mt → nt	/kontʃam/ /terijum/	[kɔɲdʒəntɛrɪujũ]
		little knows	"knows a little"
	b. mk →ŋk	/neeram//kaalam//kiʈajaataa/	[n̲erəŋkaləŋkɪɖəjada]
		time season there-isn't-q̲u̲	"'isn't there a proper time?"
	c. np → mp	/en/ /peer/	[jɛmperɯ]
		my name	"my name"

[6] Note that voice and place assimilation act differently, as voicing assimilation occurs only word-internally. I deal only with place here, since it acts the same in both word and phrase internal positions.

d. nk →ŋk	/avan/ /keekkiraan/	[ʔʌvəŋkekkɪrã]
	he hear-pres-he	"he hears"

To see how this assimilation results in a subsegment being shared across word boundaries, consider the diagram in (22):

(22) Subsyllabic segment crosses word boundaries:



[kɔɲdʒəntɛrɪjũ]

Here the place features from the second word are linked to a coda consonant syllabified with the first word. As in Luganda prenasalization compensatory lengthening, the sharing of a sub- segment is motivated by a coda restriction, here NOCODA$_{Pl}$. This constraint, which is also evaluated to allow non-crisp alignment, requires that each consonantal place of articulation be linked at the left edge of a syllable; hence, a coda consonant may not have a place of articulation distinct from that of the following onset consonant.

(23) NOCODA$_{Pl}$, a.k.a. ALIGN-L(C-Place, σ): each instance of consonantal place aligns with the left edge of some syllable

Place assimilation requires that NOCODA$_{Pl}$ outranks a correspondence constraint on the features of consonantal place, MAX-IO(C-Pl). The overall ranking appears in (24), and includes the correspondence constraints MAX-IO(C) and DEP-IO(V) .

(24) { NOCODA$_{Pl}$, MAX-IO(C), DEP-IO(V) } >> MAX-IO(C-PL).

By this ranking, the consonant is preserved in the output, to satisfy MAX-IO(C), but its place features may be deleted, to satisfy NOCODA$_{Pl}$. By ranking DEP-IO(V) high, no epenthetic vowels appear in the output in order to rescue the place features from appearing in the coda. Tableau 7a shows this ranking for stem-final nasals; the alignment constraint on words and syllable edges is not violated by anything involved in the internal assimilation and is left unmarked, though the question of the right edge of phrases is an interesting one, discussed in Wiltshire (1998).

Tableau 7a Word-internal /maram + taan/ "tree" (emph) = [mʌrən̠dã]

Candidates	No CodaPl	Max-IO (C)	Dep-IO (V)	Max-IO (CPl)	Align-L/R (Wd, σ)
.ma.ram.dãã.	*!				
.ma.rã.dãã.		*!			
.ma.ra.mu.dãã.			*!		
☞ .ma.ran.dããn.				*	

As in word-internal assimilation, the same ranking results in phrase-internal assimilation, giving the feature misalignment as shown in (22) above, with the coronal of the second word associated with a consonant at the end of a syllable in the first word.

Tableau 7b Phrase-internal /kontʃam/ /terijum/ "knows a little" = [kɔɲdʒən#tɛrɪjũ]

Candidates	No CodaPl	Max-IO (C)	Dep-IO (V)	Max-IO (CPl)	Align-R/L (Wd, σ)
.koɲ.tʃam.#te.ri...	*!				
.koɲ .tʃa.#te.ri...		*!<m>			
.koɲ.tʃa.mu.#te.ri...			*!(u)		
☞ .koɲ.tʃan̠.#teri...				*	*

Thus, the sub-segmental features of place of articulation cross the word boundary to satisfy NOCODA$_{Pl}$ in Tamil.

We have seen throughout section 3 that constraints on any part of the syllable may be responsible for misalignment of word and syllable edges. Onset phonotactics, either requiring or limiting onsets, force segments to cross into a syllable with segments of another word in Spanish and Italian. A constraint on the rhyme, NODIPH, results in material from two words sharing a syllable at the boundary in Luganda. Finally, constraints on the coda play a role in sharing subsegmental features from one word into a syllable of an adjacent word, for moras in Luganda and place features in Tamil. All of these analyses have used syllable constraints which are widely attested cross-linguistically, which is one benefit of an Optimality Theory analysis. In the following section, I will propose that the fully parallel version of OT used thus far has another advantage; it predicts that the word-alignment constraints are present even in grammars in which they are violated because they rank lower than some syllable phonotactic constraint.

4 Comparison with alternatives

Although the observations of the preceding section could be formulated in rule-based or constraint-based accounts in which word-level syllabification precedes phrasal resyllabification, I want to show now how fully parallel OT captures an aspect of word-edge alignment that other such accounts would miss. That is that the constraints in (3) are not turned off, but merely overridden. I will illustrate using Spanish (4.1) and Italian (4.2) examples.

4.1 Spanish alignment in action

In Spanish, word-edge alignment plays a role in phrasal syllabification even though it is violated in some cases. In word-internal cases, we saw that Spanish prefers to syllabify single intervocalic consonants as onsets rather than codas (section 3.1). In fact, word-internally, Spanish prefers to maximize onsets rather than tolerate codas, so that clusters of consonants will be parsed in the onset rather than coda plus onset, if possible.

(25)	a. /soplo/	[.so.plo.]	*[.sop.lo.]	"breath"
	b. /ablar/	[.a.βlar.]	*[.aβ.lar.]	"talk"
	c. /peregrino/	[.pe.ɾe.ɣɾi.no.]	*[.pe.ɾeɣ.ɾi.no.]	"pilgrim"

We also saw that a word-final consonant would syllabify with a following word-initial vowel, so that an intervocalic consonant is always realized as an onset, whether or not the syllable has to cross a word boundary. However, a word-final consonant does not syllabify across a word- boundary if the following word has an onset, even though a well-formed onset would result. Instead, the word-final consonant is parsed in the coda, violating NoCoda.

(26)	a. /klub#lindo/	[.kluβ.lin.do.]	*[klu.βlin.do.]	"beautiful club"
	b. /čef#loko/	[.čef.lo.ko.]	*[če.flo.ko.]	"crazy chef"
	c. /benid#rapido/	[.be.nið.ra.pi.ðo.]	*[be.ni.ðra.pi.ðo.]	"come (pl. imp.) quickly"

Since the following word is consonant-initial and already has an onset, syllabification aligns with the word boundaries, and any word-final consonant is in the coda. This gives the appearance of different rules of syllabification in phrases than word-internally, since a cluster such as /bl/ is treated as a good onset within a word ([.a.βlar.]), but as a coda plus onset in a phrase ([.kluβ.lin.do.]). Accounts which use different levels of syllabification for words and phrases have to postulate distinct syllabification rules (Hualde 1992). However, the generalization is that word edges coincide with syllable edges unless a syllable would lack an onset. Interestingly, the same generalization holds across prefix-edges. Unless a syllable would lack an onset (27a-c), prefix edges coincide with syllable edges (27d-f).

(27)	a. /des+igual/	[.de.si.gwal.]	*[.des.i.gwal.]	"unequal"
	b. /sub+alterno/	[.su.ßal.tér.no.]	*[.suß.a.l.tér.no.]	"subordinate"
	c. /in+esperado/	[.i.nes.pe.ɾa.ðo.]	*[.in.es.pe. ɾa.ðo.]	"unexpected"
	d. /sub+lu.nár/	[.suß.lu.nár.]	*[.su.ßlu.nár]	"sublunar"
	e. /sub+lingwal/	[.suß.liŋ.gwal.]	*[.su.ßliŋ.gwal.]	"sublingual"
	f. /ad+risar/	[.að.ri.sar.]	*[.a.ðri.sar.]	"to right (nautical)"

Although ONSET must outrank the constraints aligning word and syllable edges, these constraints do assert themselves when onset is satisfied, even at the expense of NOCODA. Thus the ranking that simultaneously gives us maximal onsets within words and syllabification across words only in cases in which a word would otherwise be onsetless is shown in (28).

(28) ONSET >> ALIGN-R, ALIGN-L >> NOCODA

Though the word alignment constraints are outranked, they assert themselves if ONSET is already satisfied, as shown in Tableau 8. Word-internal clusters form maximal onsets because of the ranking of ONSET above NOCODA, while word-final consonants do not cross word boundaries to form maximal onsets because of the ranking of the alignment constraints. I analyze prefixes as separated from the base by a prosodic word bracket; arguments for this analysis can be found in Wiltshire (to appear).

Tableau 8 Partial Analysis

Inputs	Candidates	Onset	Align-R (Wd,σ)	Align-L (Wd, σ)	No Coda
ablar	☞ a.ßlar.	*			
'to talk'	aß.lar.	*			*!
klub #~~ele~~gante	☞ klu.ß # e.le. ɣan.te.		*	*	*
'elegant club'	kluß# e.le. ɣan.te	*!			**
klub # lindo	klu.ß # lin.do		*!	*	*
'beautiful club'	☞ kluß.# lin.do				**
des+igual	☞ .de.s # i.gual.			*	*
'unequal'	.des. # i.gual.	*!			**
sub+lu.nár	.su.ß # lu.nár			*!	*
'sublunar'	☞ .suß. # lu.nár				**

Thus, though a word-final consonant may be syllabified differently in different phrasal contexts, rules of resyllabification are not required. Furthermore, we do not need rules of syllabification across word boundaries that differ from those within words, as in the different syllabifications of /bl/ word-internally vs. across words. Instead, the presence of the alignment constraints on word and syllable boundaries provides for different syllabifications in different contexts, although it is overruled if the high ranked ONSET constraint is at stake.

4.2 Italian alignment in action

I now return to *raddoppiamento sintattico* in Italian, illustrated in section 3.2, in which word-final stressed syllables have to be heavy, and use a consonant from the following word if necessary. We saw that if a word began with a geminate or sC cluster, word alignment was violated; these sequences did not form ideal onsets, so the first consonant was syllabified into the coda of the final syllable of the preceding word. However, words that begin with a single consonant or a good onset cluster have a consonant doubled to satisfy PKPROM, the requirement that a stressed syllable is heavy; these were the examples in (11a-b), such as *triste* 'sad', [t͡ʃittáttriste] 'a sad city'. Thus, when the word-initial onset is already acceptable in Italian, word- alignment is satisfied at the expense of DEP-IO(C), the constraint against epenthetic consonants on the surface.

As shown in Tableau 9, due to the ranking of alignment above DEP-IO(C), a good cluster is not broken across boundaries, nor are the word edges realigned or shifted.

Tableau 9 Input: /t͡ʃittá triste/ "a sad city"

Candidates	PkProm	*<4DifSon	Align-L(Wd, σ)	Dep-IO(C)
t͡ʃittá]$_\sigma$[triste	*!			
t͡ʃittát]$_\sigma$[riste			*!	
☞ t͡ʃittát]$_\sigma$[triste				*(t)

However, while /sC/ is treated as a tolerable cluster, phrase-initially, it is only tolerated when nothing better is available. Phrase-medially, following a vowel-final word, the constraint ranking determines that a better option is to break the cluster across words.

Tableau 10 Input: /t͡ʃittá sporka/ "a filthy city"

Candidates	PkProm	*<4difson	Align-L(Wd, σ)	Dep-IO(C)
t͡ʃittá]$_\sigma$[sporka	*!	*		
t͡ʃittás]$_\sigma$[sporka		*!		*(s)
☞ t͡ʃittás]$_\sigma$[porka			*	

Thus, the logic of the ranking is that /tr/ and /sp/ are different because one satisfies *<4DIFSON and the other does not. While the word/syllable alignment constraints are ranked low enough to be violated in order to improve the satisfaction of *<4DIFSON, they still play a role in Italian by encouraging clusters such as /tr/ to stay together, with a DEP-IO(C) violation resulting instead.

Compare this account to Peperkamp (1997), who appeals to levels of syllabification. In her account, the resyllabification of *triste* is blocked by a kind of FAITHFULNESS to previously built lexical syllabifications, so that at the phrasal level, the initial consonant must be doubled to satisfy the weight requirement of the preceding stressed syllable. Such an account would then have difficulty with handling /#sC/ cases, where the equivalent "resyllabification" does happen, misaligning the word boundary by putting a word-initial /s/ into the stressed syllable of the preceding word. That is, if we use two levels and faithfunless to lexical structure and rank it high, we can get [t͡ʃittat.triste] but also *[[t͡ʃittas.sporka], while if we rank faithfulness low, we can get [[t͡ʃittas.porka] but also *[[t͡ʃittat.riste]. A possible alternative analysis to preserve Peperkamp's approach would be to treat the /s/ in an /sC/ cluster as at least temporarily extraprosodic, though extraprosodicity is generally avoided in OT. In this case, it seems to be merely a way to look ahead to the phrasal context, since a special structure is being built lexically for /sC/ clusters in order to accomodate their phrasal syllabification.

The account here, which is also based on Wiltshire and Maranzana (1999), uses independently motivated onset sonority sequencing constraints (Davis 1990), which capture the different behavior of word-initial clusters in the *raddoppiamento sintattico* contexts in Italian. Furthermore, as with Spanish syllabification, we do not require levels of syllabification or re-syllabification as in previous rule or constraint-based analyses. Finally, the use of the word/ syllable alignment constraints shows a phenomenon that is an essential claim of OT: though constraints may be outranked in a grammar, they will express themselves when the higher ranked constraints are tied.

5 Conclusions

In each case discussed here, the syllabification of words in isolation and in phrases has been shown to result from the same ranked set of constraints within each language; hence resyllabification at word-edges is shown to be unnecessary in a constraint-based account. Syllabification crosses word boundaries to satisfy constraints on syllable markedness (onset requirement, onset, nucleus, and coda restrictions) and IO correspondence at the expense of alignment. In each analysis, the markedness constraints involved are justified cross linguistically, language specifically, and word-internally, so that it should be no surprise to see the role they play across words in phrases.

An OT account is best able to capture the role of the constraints aligning word and syllable edges even in languages in which they are sometimes violated due to higher ranking constraints. Rather than requiring a set of rules ordered with resyllabification, constraints at the edges account for the limitations of cross-word syllabification, and provides for syllabification with independently motivated constraints on prosodic structures, so that OT need not resort to multiple levels, but instead can be a truly parallel system.

References

Chierchia, G. (1986). Length, syllabification, and the phonological cycle in Italian. *Journal of Italian Linguistics* 8: 5-34.

Christdas, P. (1988). *The Phonology and Morphology of Tamil*. Unpublished doctoral dissertation, Cornell University.

Clements, G. N. (1986). Compensatory lengthening and consonant gemination in Luganda. In L. Wetzels and E. Sezer, eds., *Studies in Compensatory Lengthening*. Dordrecht: Foris, pp. 37-76.

Davis, S. (1999). On the representation of initial geminates. *Phonology* 16.

Davis, S. (1990). The Onset as a Constituent of the Syllable: Evidence from Italian, in M. Ziolkowsky, M. Noske, and K. Deaton (eds.) *Papers from the 26th Regional Meeting of the Chicago Linguistic Society*, CLS: Chicago, 71-80.

Harris, J. (1983). *Syllable Structure and Stress in Spanish: A Non-Linear Analysis*. Cambridge, MA: MIT Press.

Harris, James W. (1993). Integrity of prosodic constituents and the domain of syllabification rules in Spanish and Catalan. *The View from Building 20*, Ken Hale and S. Jay Keyser, Eds., 177-194. Cambridge: MIT Press.

Herbert, R. (1975). Reanalyzing prenasalized consonants. *Studies in African Linguistics* 6.2: 105-123.

Hualde, José Ignacio. (1992)."On Spanish syllabification. in Hector Campos and Fernando Martínez-Gil, Eds., 475-493. *Current Studies in Spanish Linguistics*. Washington, DC: Georgetown.

Itô, J. and A. Mester. (1994). Realignment. *Proceedings of the 1994 Utrecht Prosodic Morphonology Workshop*, Utrecht, Holland.

Kiparsky, P. (1982). Lexical morphology and phonology. *Linguistics in the Morning Calm*. Linguistic Society of Korea (eds.) Seoul: Hanshin Publishing Co., pp. 3-91.

McCarthy, J. and A. Prince. (1993a). Generalized Alignment. Technical Report #7 of the RuCCs, Rutgers, NJ.

________. (1993b). Prosodic Morphology. Technical Report #3 of RuCCS, Rutgers, NJ.

________. (1995). Faithfulness and Reduplicative Identity. In Beckman et al. (eds) *Papers in Optimality Theory*, GLSA, U Mass., Amherst: pp. 249-384.

Mohanan, K.P. (1982). *The Theory of Lexical Phonology*. Unpublished doctoral dissertation, MIT.

Peperkamp, S. (1997). *Prosodic Words*. PhD dissertation: University of Amsterdam. Den Haag: Holland Academic Graphics, HIL dissertation: 34.

Prince, A. and P. Smolensky. (1993). Optimality Theory. Technical Report #2 of the RuCCS, Rutgers, NJ.

Rosenthall, S. (1997). The distribution of prevocalic vowels. *Natural Language and Linguistic Theory*. 15:1, pp. 139-180.

Saltarelli, M. (1970). *A Phonology of Italian in Generative Grammar*. The Hague: Mouton.

Steriade, D. (1982). *Greek Prosodies and the Nature of Syllabification.* Ph.D. Dissertation, MIT, Cambridge.

Selkirk, E.(1984). *Phonology and Syntax: The Relation between Sound and Structure.* Cambridge, Ma:MIT Press.

Vogel, I. (1977). *The Syllable in Phonological Theory; with Special Reference to Italian.* PhD dissertation. Stanford, CA.

Tucker, A.N. (1962). The syllable in Luganda: A prosodic approach. *Journal of African Languages* 1.2: 122-66.

Wiltshire, C. (1998). "Extending ALIGN constraints to new domains." *Linguistics* 36.423-467.

________. (1999). "The Conspiracy of Luganda Compensatory Lengthening," *New Dimensions in African Linguistics and Languages,* Paul Kotey, ed. Vol 3, p. 131-147.

________. (to appear), "Variation in Spanish and prosodic boundary constraints", in *Formal Perspecives on Romance Linguistics*, T. Sattersfield, C. Tortora, and D. Cresti, eds., Amsterdam and Philadelphia: John Benjamins.

Wiltshire, C. and E. Maranzana. (1999). Geminates and clusters in Italian and Piedmontese: a case for OT ranking *Formal Perspectives on Romance Linguistics*, M. Authier, B. Bullock, and L. Reed, eds., Amsterdam and Philadelphia: Benjamins, pp.289-303.