Infixation and segmental constraint effects: UM and IN in Tagalog, Chamorro, and Toba Batak*

In press: *Lingua* (2004)

Thomas B. Klein

Georgia Southern University
Department of Writing and Linguistics, P.O. Box 8026
Statesboro, GA 30460, USA

Phone: +1-912-681-0986
Fax: +1-912-681-0783

tklein@georgiasouthern.edu
www.georgiasouthern.edu/~tklein

Abstract
The proper understanding of infixation continues to be a matter of debate among linguists. The data presented in this paper show that infixation in Austronesian languages is not exclusively due to prosodic morphology, but instead is significantly influenced by the segmental phonology as well. Specifically, infixation may be blocked if it would create dissimilation environments in the first bimoraic foot of the morphological base. To capture this asymmetric effect, it is proposed that OCP-type markedness constraints may be sensitive to positional domains. The analysis also accounts for the alternatives to infixation that individual languages employ. Blocking creates a morphological gap in Tagalog. In contrast, infixation in Chamorro competes with prefixation plus metathesis, whereas prefixation with assimilation is observed in Toba Batak. The investigation of Chamorro also uncovers the phonological conspiracy that infixation and metathesis are both driven by the prosodic requirement that syllables must have onsets.

*This contribution was supported in part by awards from the British Arts and Humanities Research Board (AHRB) (Research Award MRG-AN5855/APN9404), the British Academy (Grant OCG-33464) and the Research Support Fund of the Research and Graduate Support Unit (RGSU) of The University of Manchester. Particular gratitude is extended to the native speakers and scholars of Chamorro whose data is represented here. Special thanks to Meta Y. Harris for her assistance with the Chamorro data. Fragments of earlier versions of this material were presented at a colloquium at the University of Leeds, at the 9th International Conference on Austronesian Linguistics in Canberra and at the 9th and the 10th Manchester Phonology Meeting. I am grateful to the audiences there for useful comments and in particular to Ricardo Bermúdez-Otero, Marc Brunelle, Colleen Fitzgerald, Anthony Fox, Anthony Grant, Barry Heselwood, Larry Hyman, Ghada Khattab, Martin Krämer, April McMahon, Diane Nelson, Charles Reiss and an anonymous reviewer. None of these institutions and individuals are to be held responsible for any of the shortcomings of this research.

1. Introduction

Infixation has been a central focus of research in Optimality Theory (OT) from the start (McCarthy & Prince, 1993a, b; Prince & Smolensky, 1993). The analysis of Tagalog -um- infixation in McCarthy & Prince (1993a) especially has been regarded as a significant achievement. It offers a straightforward understanding of this type of infixation as the interaction of edge-bound morphological concatenation with simple prosodic constraints such as the requirement that syllables avoid codas (NOCODA). According to this analysis, Tagalog -um- is placed as far left in the word as possible without creating offending syllabic structures. This implies a causal connection between the segmental form of the affix and its placement. Affixes consisting of a vowel-consonant sequence may be infixes because their placement as prefixes or suffixes would create illicit onsetless syllables or codas.

Given that constraints regulating concatenation and syllable markedness are needed independently to account for other morphological and phonological phenomena, the OT approach predicts that there should be a sizeable number of languages which show infixation similar to Tagalog. This prediction is broadly confirmed, given that numerous Austronesian languages have affixes which exhibit um-type infixation (see, e.g., Tryon, 1995a and the contributions in Tryon, 1995b; Barsel, 1994; Hayes & Abad, 1989). Subsequent authors such as Yu (2002), Blevins (1999), Stemberger & Handford Bernhardt (1999), and Golston (1996) have argued that certain infixation patterns are a genuine morphological operation without prosodic interaction. Yet, there is a broad consensus that a significant set of infixation phenomena exists which clearly shows interactions with the prosodic phonology.

The thesis that infixation of the Tagalog type is the result of Prosodic Morphology goes back to McCarthy & Prince (1986). Prosodic Morphology in the OT framework results by ranking prosodic constraints higher than morphological ones. This approach predicts that the segmental phonology plays no role in infixation. However, Crowhurst (1998) has concluded correctly that um-infixation in Toba Batak, an Austronesian language of Sumatra, is influenced by segmental factors. In particular, UM in Toba Batak surfaces as a prefix when the initial consonant of the stem matches the labial or nasal quality of /m/ in um. For example, infixation after the initial labial in paddita 'preach' is not possible; instead, the attested form is up-paddita 'preacherlike', where UM surfaces as a prefix and /m/ is fully assimilated to the stem-initial C.

Recent data elicited by Orgun & Sprouse (1999) have confirmed Schachter & Otanes's (1972) observation that um-infixation in Tagalog is subject to segmental restrictions as well. Similar to Toba Batak, um-infixation in Tagalog is not possible after labial sonorants. In sharp contrast to Toba Batak, an alternative placement of UM is not available in Tagalog. Thus, the morphosyntactic functions of UM must be expressed through other means for stems such as mahal 'expensive'. The data from the Austronesian language Chamorro, spoken indigenously in the Mariana Islands of Guam, Saipan, Rota and Tinian in the Western Pacific, presented in this paper show further segmental constraints on infixation. In particular, Chamorro infixation after obstruents is analogous to Tagalog, but infixation after sonorants competes with metathesized prefixation. Thus, UM-affixation in stems such as li'e' 'to see' may surface as l-um-i'e' or mu-li'e'. Considering the data from Tagalog, Chamorro and Toba Batak, it appears that there is a family of dissimilatory effects on the availability and placement of UM-affixation which has not been noticed before. Consequently, a major aim of this paper is to propose a unified account of the interaction of infixation with dissimilatory segmental environments in the OT framework.

The OT approach is also useful in the account of a phonological conspiracy effect that had gone unnoticed previously. As is shown below, Chamorro infixation and metathesis are driven by the surface requirement that syllables have onsets. This is a phonological conspiracy in the sense of Kisseberth (1970) in that heterogeneous processes converge to produce a homogeneous output target. As shown in McCarthy (2000), among others, conspiracy effects fall out from the central organizing OT principle that a grammar is a language-particular ranking of violable, but universal constraints. This approach to phonological conspiracies has been one of the major successes of the OT framework. Thus, the analysis of Chamorro presented in this paper is further support for the value of the OT approach.

Orgun & Sprouse (1999) have argued that the Tagalog facts make necessary a significant change in the OT architecture to include a CONTROL component which contains only inviolable constraints that cause ungrammaticality. McCarthy (2002a), however, has revisited the OT analysis of Tagalog -um- and has defended the conventional OT architecture. McCarthy has shown that the need for CONTROL is obviated if the gradient violation of alignment constraints in McCarthy & Prince's (1993a) account is replaced by quantized, categorical alignment. In addition, he has also argued that quantized alignment is necessary because alignment by syllables (ALIGN-BY-σ) must be distinguished from alignment by segments (ALIGN-BY-SEG) for the analysis of Tagalog. This move would not have been possible in the gradient alignment framework. However, it is demonstrated below that ALIGN-BY-SEG may simply be omitted in the analysis of Tagalog without compromising the result. Thus, this part of McCarthy's argument for quantized alignment is called into question. However, the analysis of variation in Chamorro infixation presented below shows that the split between ALIGN-BY-σ and ALIGN-BY-SEG is indeed necessary. Thus, a new, strong argument for the categorical alignment framework in the analysis of infixation is gained through the analysis of Chamorro.

According to Crowhurst (2001), the data from Toba Batak provide an argument for crucial non-ranking, or *co-ranking*, of constraints. Crucially non-ranked constraints occupy the same position in the constraint hierarchy, yet none of them dominates other constraints in the set. Violations of crucially non-ranked constraints count equally for the purposes of optimization. Crucial non-ranking of constraints has been considered as a logical possibility early on (Prince & Smolensky, 1993: 51), but has not been explored further in foundational OT works. Even though it emerges as an option, it is obvious that the admission of crucial non-ranking is a fundamental proliferation of the constraint component and, more importantly, runs counter to the OT thesis that grammars are hierarchies of ranked constraints. In the words of Tesar & Smolensky (1998: 232): "The ranking defining a grammar is total: the hierarchy determines the relative dominance of every pair of constraints". Furthermore, co-ranking would increase the generative power of the model exponentially because for every constraint there would be the option of it being ranked or not ranked with respect to the other constraints. Given these significant arguments against crucial non-ranking, it is important that I demonstrate in my analysis of Toba Batak below that it is not necessary. Instead, it is shown that ranked constraints are sufficient to capture the behavior of UM in Toba Batak.

The remainder of this paper is organized as follows. §2 provides a critical review of the analyses of Tagalog um-infixation in Orgun & Sprouse (1999) and McCarthy (2002a). Orgun & Sprouse's particular approach to dissimilation is shown to be in need of revision. It is proposed that the dissimilatory segmental environments affecting infixation are best understood via a family of domain-specific OCP-type constraints. A full data set of Chamorro infixation from the author's original fieldwork showing the segmental effects and age-graded variation is presented in §3. It is shown in this section that quantized alignment and the OCP-type constraints are necessary for the analysis of UM/IN-affixation in this language. The analysis of Toba Batak in §4 presents further support for the approach to dissimilation invoked in this paper and shows that the conventional constraint architecture suffices to analyze infixation in this language. In §5, Halle's (2001) proposal of infixation as rule-based metathesis is examined and consequently rejected. §6 offers some conclusions.

2. Alignment, prosody and dissimilatory blocking in Tagalog

This section critically reviews the shift from gradient to categorical alignment and revisits the role of prosodic constraints in driving infixation. McCarthy (2002a) invokes categorical alignment in part to obviate the CONTROL component proposed in Orgun & Sprouse (1999). I demonstrate that categorical alignment is not required to capture Tagalog UM-infixation; however, it is necessary to distinguish which phonological category incurs alignment violations, e.g. segment versus syllable. The second part of this section reviews and critiques Orgun & Sprouse's (1999) version of the OCP that they use to account for the blocking of UM-infixation in dissimilatory segmental environments. To capture the segmental effects concerning the availability of UM-infixation in the three languages under discussion, I invoke a family of OCP-type constraints that is specified for the domain of the base-initial bimoraic foot.

The initial OT analyses by McCarthy & Prince (1993a, b) did not take into account that all Tagalog words must begin with a consonant, and that UM-infixation is blocked in certain segmental environments. Orgun & Sprouse (1999) and McCarthy (2002a) have explored the consequences of these facts for the original analytical framework.

Note that all potential infixes in the data to follow are underlined for the convenience of the reader. Consider now the Tagalog data from Orgun & Sprouse (1999).

(1) UM-infixation in Tagalog

a. C-initial words

sulat	sumulat	'to write'
ʔabot	ʔumabot	'to reach for'

b. CC-initial words

gradwet	grumadwet ~ gumradwet	'to graduate'
preno	prumeno ~ pumreno	'to brake'

c. *m*, *w*-initial words

mahal	*mumahal	'expensive'
mura	*mumura	'cheap'
walow	*wumalow	'to wallow'

d. *s* + *m*, *w*-initial words

smajl	*summajl, *smumajl	'to smile'
swiŋ	sumwiŋ, *swumiŋ	'to swing'

First, this data set shows that there are no vowel-initial words or syllables at the surface in Tagalog (see also Zorc, 1995, Schachter & Otanes, 1972 and Pittman, 1966). Instead, words that begin with a vowel in the orthography such as *abot* have an initial glottal stop in the phonology. Thus, -um- in Tagalog never occurs as a prefix on the surface. Instead, -um- is infixed after a single initial consonant as in (1) (a). With cluster-initial bases, -um- is variably infixed within or after the cluster, at least for some speakers, as shown in (1) (b). Interestingly, um-infixation does not occur, as seen in (1) (c), when the initial consonant of the base matches the labial sonorant quality of /m/ in um, that is, there is no infixation after /m/ or /w/ (see also Schachter & Otanes, 1972: 292). As the ungrammatical items in (1) (c) show, there is nothing wrong prosodically with infixation in this context. Unmarked CV syllables are created in *m+um+ahal just as in s+um+ulat. Thus, the Prosodic Morphology hypothesis cannot explain why infixation is not available with initial labial sonorants. Instead, we observe a segmental, dissimilatory blocking effect on affixation; -um- cannot appear with *m* and *w* because an illicit sequence of labial sonorants would be created.

The dissimilatory blocking of –um- in Tagalog creates a gap in that no alternative realization of it is grammatical in these environments. Consequently, the morphosyntactic functions of -um- in the ungrammatical forms must be fulfilled by other means. As pointed out by an anonymous reviewer, similar cases are found in other languages. In Hungarian, for example, the suffix –asz must be replaced by –ol after stems ending in coronal sibilants (see Carstairs-McCarthy, 1995).

2.1 Gradient versus categorical alignment

Generalized Alignment (GA) constraints are pivotal to the analysis of infixation in McCarthy & Prince (1993a). These constraints demand that a designated edge of a prosodic or grammatical category coincides with a designated edge of some other prosodic or grammatical category. The evaluation of GA constraints in McCarthy & Prince (1993a) is gradient in that the degree of violation is measured incrementally by the distance of the category at issue from the designated edge. Thus, given the GA constraint in (2) and the NOCODA constraint in (3), candidates for the Tagalog form gr+um+adwet, from gradwet 'to graduate', are evaluated as in the tableau in (4). The syllabification is indicated by the periods. The comma separating the affix and the base in the input encodes the absence of linear ordering between the two elements.

(2) ALIGN-um
Align([um]$_{Af}$, L, stem, L)

(3) NOCODA
Syllables must be open.

(4)

	Input: /um, gradwet/	NOCODA	ALIGN-um
	a. um.grad.wet.	***!	
	b. gum.rad.wet.	***!	g
☞	c. gru.mad.wet.	**	gr
	d. grad.wu.met.	**	gra!dw

Prefixation of -um- to this base as in (4) (a) incurs three NOCODA violations, given the two codas in this particular base plus another one in the VC prefix. Note that -um- in this candidate is perfectly aligned with the beginning of the stem. Putting -um- between the members of the consonant cluster as in (4) (b) incurs the same number of NOCODA violations as the prefixing candidate. This candidate is worse than prefixation, however, because the left edge of -um- and the left edge of the stem are misaligned by one consonant. In (4) (c) -um- is placed before the first vowel. In that way the number of codas in the stem is reduced by one compared to candidates (4) (a) and (4) (b). The increase in alignment violations is not decisive, however, because of the subordinate ranking of the alignment constraint. Placing -um- before the second vowel as in (4) (d) incurs the same number of NOCODA violations as the placement before the first vowel, but this candidate is much worse in terms of alignment. Candidate (4) (c) incurs two gradient alignment violations, one for each member of the initial consonant cluster. The hyperinfixed candidate in (4) (d), however, incurs a gradient alignment violation for each segment separating the affix from the beginning of the stem. A neat consequence of McCarthy & Prince's (1993a) analysis is that the location of the affix need not be specified in the input because it falls out from the constraint hierarchy. As a result, affix placement information in the input is redundant and, hence, should not be encoded there.

Orgun & Sprouse (1999) have argued that McCarthy & Prince's (1993a) approach cannot account for the full data set in (1) without a constraint ranking paradox. In particular, they find the approach to ungrammaticality in Prince & Smolensky (1993), which relies on the constraint MPARSE, to be inadequate to account for um-blocking in Tagalog. Recall that no um-form surfaces when it would be placed after a labial sonorant. Thus, according to Orgun & Sprouse (1999), forms such as *sw+um+iŋ are disallowed because of a prohibition against labial sonorants in consecutive onsets applying to the affix -um-. This constraint is given in (5).

(5) OCP-um (Orgun & Sprouse, 1999)
*mum, *wum: Sonorant labials are not allowed in consecutive onsets.

Having established the constraint rankings shown in tableau (6), Orgun & Sprouse (1999: 206) point out that hyperinfixation is incorrectly predicted to be optimal, as shown by the sad face symbol, in cases where complete blocking of um-affixation, i.e. the null parse, is observed at the surface.

(6) Incorrect optimization of hyperinfixation

	Input: /um+mahal/	OCP-um	MPARSE	NOCODA	ALIGN
	a. mumahal	*!		*	m
	b. ummahal			**!	
	c. Null Parse		*!		
☹	d. mahumal			*	mah

Given that surfacing candidates may violate NOCODA and ALIGN, as seen in (4), MPARSE must outrank both constraints. The problem for McCarthy & Prince's (1993a) infixation theory is that it predicts that hyperinfixation should be permitted to avoid structures offending OCP-um, given the high ranking of MPARSE, the constraint demanding the realization of morphemes, and the low ranking of gradient ALIGN. The exclusion of um-infixation with sonorant labial-initial bases would require the ranking of the constraints OCP-um and ALIGN above MPARSE, whereas MPARSE must outrank ALIGN elsewhere. Orgun & Sprouse (1999) propose to resolve this ranking paradox by significantly restructuring the architecture of the OT constraint system. They propose to add a non-optimizing constraint component called CONTROL; it contains only inviolable constraints that cause ungrammaticality. According to Orgun & Sprouse, there is no MPARSE constraint and OCP-um is in CONTROL in Tagalog. Consider the tableau in (7).

(7) Ungrammaticality via CONTROL (Orgun & Sprouse, 1999: 208)

	/um+mahal/	NOCODA	ALIGN
a. EVAL	☞ mumahal	*	m
	ummahal	**!	
	mahumal	*	ma!h
		OCP-um	
b. CONTROL	✂ mumahal	*!	

On the pass through EVAL, the prefixed and the hyperinfixed candidate lose. The infixed candidate winning in EVAL is passed on to the CONTROL component, however, and loses there. Thus, the null output is the only candidate which survives EVAL and CONTROL. Consequently, forms like m+um+ahal are rendered ungrammatical and the lack of um-infixed surface forms with initial labial sonorants is accounted for.

McCarthy (2002a) has shown that the CONTROL component is unnecessary in Tagalog when the alignment constraint schema is refined. He argues that the issue lies in the gradient assignment of violation marks and proposes to replace gradient GA constraints by a quantized, categorical alternative.

(8) Quantized, categorical ALIGN (-um-, Wd, L) (McCarthy, 2002a)
a. ALIGN-BY-□(-um-, Wd, L)
No syllable stands between the left edge of -um- and the left edge of Wd.

b. ALIGN-BY-SEG(-um-, Wd, L)
No segment stands between the left edge of -um- and the left edge of Wd.

The constraints in (8) (a) and (8) (b) demand that -um- stands in initial position in the word. Categorical alignment is calculated in syllables for (8) (a) and in segments for (8) (b). (8) (a) assigns a single violation mark if -um- is one or more syllables away from the beginning of the word, whereas (8) (b) assigns one violation mark if -um- is one or more segments short of starting the word. In contrast to gradient alignment, no multiple marks result if -um- in a given candidate is more than one syllable or segment away from the word edge. Constraints may assign multiple marks, but only if there are multiple violating structures in the candidate under evaluation. For instance, given the constraint LINEARITY which demands that the input is consistent with the precedence structure of the output (McCarthy & Prince, 1995) and the mapping $/C_1V_2C_3V_4C_5/ \rightarrow [C_1V_2C_5C_3V_4]$, one violation mark is assigned to each of the offending precedence relations, that is, $C_3 > C_5$ and $V_4 > C_5$.

Note that quantized alignment constraints are much more specific and finely tuned than conventional GA constraints as to which structures and categories are at issue for the calculation of violation marks. As a result, Tagalog um-infixation may be analyzed without incurring a ranking paradox, as recapitulated below.

McCarthy (2002a) views blocking as ineffability in that there are no surface -um- forms for words with a certain phonological shape, even though there is no restriction on the input to provide the elements that would otherwise license them. Blocking qua ineffability is viewed by McCarthy as the optimization of the null output. This output, symbolized by '⊙', is a candidate that does not have any phonological structure and whose correspondence relation to the input is undefined. Thus, ⊙ satisfies all faithfulness constraints and phonological markedness constraints such as NOCODA. However, ⊙ violates constraints of the M-PARSE family. Thus, according to McCarthy (2002a), the constraint M-PARSE(-um-) is violated by the candidate ⊙ whenever the input contains the morpheme -um-.

(9) -um- infixation: initial non-labials (McCarthy, 2002a)

	Input: /um+sulat/	M-PARSE(-um-)	ALIGN-BY-SEG(-um-, Wd, L)
☞	a. sumulat		*
	b. ⊙	*!	

In (9), the null output correctly loses because M-PARSE(-um-) is ranked high. Note the replacement of the gradient ALIGN($[um]_{Af}$, L, stem, L) constraint through the categorical constraint ALIGN-BY-SEG(-um-, Wd, L).

McCarthy (2002a) invokes the quantized constraint ALIGN-BY-□(-um-, Wd, L) and Orgun & Sprouse's OCP-um constraint to account for the blocking of -um- infixation with initial labial sonorants. Consider the tableau in (10).

(10) Null output: initial labial sonorants (McCarthy, 2002a)

	Input: /um+mahal/	OCP-um	ALIGN-BY-□	M-PARSE(-um-)	ALIGN-BY-SEG
☞	a. ⊙			*	
	b. mumahal	*!			*
	c. mahumal		*!		*

The ranking of M-PARSE(-um-) below the constraints OCP-um and ALIGN-BY-□(-um-, Wd, L) is crucial. Given that a violation of ALIGN-BY-□ is incurred only for the hyperinfixed candidate (10) (c), the null output in (10) (a) emerges as optimal because infixation before the first vowel in (10) (b) violates the OCP-type constraint. This is possible because infixation before the first vowel passes the new ALIGN-BY-□ constraint. Thus, instead of a powerful restructuring of the architecture of the constraint system, a refinement of the alignment constraint schema is sufficient to account for the blocking of -um- forms with initial labial sonorants.

According to McCarthy (2002a), the tableau in (10) contains an argument for quantized alignment because M-PARSE(-um-) is ranked between the two alignment constraints; hence, they must be understood as separate constraints. Splitting alignment constraints in this way is not possible in the gradient alignment theory. This argument for quantized alignment does not quite go through, however, because the constraint ALIGN-BY-SEG(-um-, Wd, L) can be shown to be redundant. Note that ALIGN-BY-SEG(-um-, Wd, L) plays no role in excluding or licensing any of the candidates in tableau (9) or (10). In other words, violations incurred by the constraint ALIGN-BY-SEG (-um-, Wd, L) are not decisive in choosing any of the candidates. Furthermore, it is the lowest-ranked constraint in McCarthy's (2002a) hierarchy for Tagalog. Thus, it can be omitted without compromising the results as shown in (11), hence undermining this part of the argument for categorical alignment.

Two issues need to be addressed before we proceed. First, McCarthy (2002a) seems to encode affix ordering in the input, for example as in /um+sulat/ in tableau (9) above. As we know from McCarthy & Prince (1993a), this is redundant given that the ordering information is encapsulated in the alignment constraints. Thus, I propose to continue to adhere to the view that affixes are unordered in the input with respect to the base. This is indicated by the commas separating affix and root in the tableaux to follow.

Secondly, note that the alignment constraints in McCarthy (2002a) and McCarthy & Prince (1993a) refer to the affix via a single phonological surface form. This is not advantageous, given that the cognate affixes in Chamorro and Toba Batak exhibit multiple surface forms. Thus, I propose to use archimorphemic notations such as UM as shorthand for the set of phonologically conditioned affixal allomorphs. Thus, ALIGN-BY-□(UM, Wd, L) and M-PARSE(UM) replace the corresponding constraints of McCarthy & Prince (1993a) and McCarthy (2002a). Consider now the tableau in (11).

(11) Simplification of McCarthy's (2002a) analysis

		OCP-um	ALIGN-BY-□(UM, Wd, L)	M-PARSE(UM)
	Input: /um, sulat/			
☞	a. sumulat			
	b. ⊙			*!
	Input: /um, mahal/			
☞	c. ⊙			*
	d. mumahal	*!		
	e. mahumal		*!	

The optimal candidates in (11) are the same as with McCarthy's (2002a) more complex constraint hierarchy. In other words, nothing is lost by omitting the constraint ALIGN-BY-SEG from consideration.

Importantly, the categorical calculation mode is also not necessary to achieve the required result. Hyperinfixation is forbidden because it places the infix a syllable or more inside the word. This contrasts with actually surfacing infixation which places the infix no more than one syllable onset inward. Thus, what is needed for the analysis of Tagalog is to be specific that syllables, not segments, are used to calculate the violations. This is demonstrated in tableau (12) where quantized ALIGN-BY-□ has been replaced by its would-be gradient counterpart.

(12) Gradient ALIGN-BY-□

		OCP-um	Gradient ALIGN-BY-□(UM, Wd, L)	M-PARSE(UM)
	Input: /um, sulat/			
☞	a. sumulat			
	b. ⊙			*!
	Input: /um, mahal/			
☞	c. ⊙			*
	d. mumahal	*!		
	e. mahumal		*!	
	f. mahalum		*!*	

As the correct optimization of the candidates in tableau (12) shows, a gradient alignment constraint could exclude hyperinfixation (see (12) (e) and (12)) (f)) if the violations are calculated in syllables. Thus, the only advantage of the quantized schema for the analysis of Tagalog is that the category used to calculate the violations is made an explicit part of the constraint.

The availability of the constraint split, however, is vindicated through the Chamorro variation data analyzed in §3. It is demonstrated there that an ALIGN-BY-SEG constraint is necessary in addition to ALIGN-BY-□. Thus, Chamorro actually provides the type of argument for quantized, categorical alignment intended in McCarthy (2002a) for Tagalog.

2.2 ONSET versus NOCODA constraints

Recall that Tagalog does not have any vowel-initial words at the surface. Consequently, the constraint ONSET must be ranked high, and it essentially drives infixation in the revised OT analysis of McCarthy (2002a). This is a significant change from the earlier analysis, where NOCODA was responsible for infixation. However, coda avoidance is still important because it contributes to the account of the variable placement of -um- with initial consonant clusters. Recall that gr+um+adwet and g+um+radwet 'to graduate' are possible in the expanded Tagalog data set. McCarthy (2002a) shows that ranking NOCODA and *COMPLEXONSET variably in the sense of Kiparsky (1993) and Anttila (1997) is able to account for this variation. In this framework, each ranking of non-crucially ranked constraints supplies variant surface candidates when the candidates satisfy all higher-ranking constraints. Consider the tableaux in (13) and (14).

(13)

	Input: /um, gradwet/	NOCODA	*COMPLEXONSET
☞	a. grumadwet	**	*
	b. gumradwet	***!	

(14)

	Input: /um, gradwet/	*COMPLEXONSET	NOCODA
	a. grumadwet	*!	**
☞	b. gumradwet		***

When the prosodic constraint NOCODA is ranked high, the placement of -um- after the initial consonant cluster is optimal, as shown in (13). In the reverse ranking, placement of -um- after the first consonant is optimal because this avoids consonant clusters in the onset, as shown in (14).

2.3 Dissimilatory environments

Recall from the discussion above that UM-infixation in Tagalog is blocked after a labial sonorant. Orgun & Sprouse (1999) have proposed to account for this restriction through an OCP-type constraint, repeated for convenience in (15). Recall also that McCarthy (2002a) has adopted this constraint in his analysis as well.

(15) OCP-um (Orgun & Sprouse 1999)
*mum, *wum: Sonorant labials are not allowed in consecutive onsets.

Orgun & Sprouse's OCP-um constraint singles out the morpheme -um-, specifies the segmental environment as sonorant labials, and invokes the peculiar prosodic environment of consecutive onsets. A broader view of this type of constraint is necessary, however, given that dissimilatory effects in the languages at hand is not limited to UM or to sonorant labials. The Chamorro case detailed further in section 3 shows that IN may be affected by dissimilation as well. Chamorro and Toba Batak show that the phonological environment may involve sonorants, nasals or labials so that interaction with dissimilation appears as a more general, feature-driven property of infixation. In addition, close inspection of other sources on Tagalog reveals that sonorant labials may occur in consecutive onsets if –um- is placed before a sonorant labial in the base. Consequently, the phonological domain specification in Orgun & Sprouse (1999) is empirically problematic, as shown below.

Any analysis of the Tagalog facts has to acknowledge that mum sequences are allowed in roots but not in the infixation environment, as pointed out in Orgun & Sprouse (1999); consider the data in (16).

(16) mum in Tagalog roots

mumo	'particles of cooked rice'
mumoʔ	'ghost'
mumugin	'gargle'

The comparison of the data in (16) with (1) (c) and (1) (d) above shows that a sequence of labial sonorants is disallowed in Tagalog only in the morphological environment of um-infixation.

Against the predictions of Orgun & Sprouse's OCP-type constraint, consecutive onsets of sonorant labial may be created through um-infixation, but only if -um- precedes the other sonorant labial.

(17) Tagalog -um- before sonorant labials in consecutive onsets

ʔuwi	ʔumuwi	'go home'	(Schachter & Otanes 1972: 293)
luwas	lumuwas	'go down-stream'	(Bloomfield 1917: 397)

Orgun & Sprouse's (1999) constraint incorrectly predicts that the forms in (17) should be ungrammatical, given that there are labial sonorants in consecutive onsets as the result of -um-infixation. The availability of -um- in (17) versus the dissimilatory blocking in (1) (c) and (1) (d) may be explained, however, if we understand the phonological domain of the OCP-type constraint as the base-initial bimoraic foot. The data in (18) show the data in question with the annotation for base-initial bimoraic feet.

(18) um-blocking within the base-initial bimoraic foot

(a)	*(muma)hal	*(mumu)ra	*(wuma)low
	*(smuma)jl		
(b)	(sum)wiŋ	'to swing'	
(c)	*(swumi)ŋ	'to swing'	
(d)	(ʔumu)wi	(lumu)was	

In the spirit of Hayes (1989) and others, I assume for the Austronesian languages at hand that onsets never carry moras, that short vowel nuclei carry one mora and that heavy syllables carry two moras. Note that the complete analysis of Austronesian languages such as Tagalog may well reveal that the kind of foot required by a language's prosodic morphology system is not necessarily the same as that required by its metrical system (cf. Hayes 1995: 47). Note furthermore that infixes are contained inside their morphological bases at the surface. Given that more than one sonorant labial in um-infixation is disallowed only in the base-initial bimoraic foot, the ungrammaticality of the forms in (18) (a) is explained because the foot in these forms contains more than one labial sonorant. Note that the bimoraic foot in (18) (b) is monosyllabic because –um- forms a heavy syllable rime. Hence, the potentially offending sonorant labial /w/ is outside of the domain of the initial foot, and the form is allowed. If –um- is placed after the consonant cluster containing /w/ as in (18) (c), ungrammaticality results because both sonorant labials are now inside the bimoraic foot. Word-medial or word-final sonorant labials are now unproblematic because they are outside the base-initial bimoraic foot, as shown in (18) (d). Thus, the seeming asymmetry between initial and medial sonorant labials is resolved because the dissimilatory effect is limited to the first foot. Note that I follow Orgun & Sprouse (1999) in assuming that forms such as *summajl are out because of a distinct constraint against labial geminates.

Given the preceding discussion, it seems clear that the phonological domain of the base-initial bimoraic foot is appropriate to produce the desired result. However, it is not sufficient to merely annotate Orgun & Sprouse's constraint with this domain. As we have seen in §1 and as it is discussed further in §3, the dissimilatory effects in Chamorro infixation obtain with sonorants regardless of place of articulation. Furthermore, the case of Toba Batak shows that um-infixation may be unavailable with labials or nasals. Given this variety of natural classes showing dissimilatory effects in the three languages under discussion, it appears that a family of OCP-type markedness constraints is at work, where individual constraints may refer to single phonological features as well. To capture blocking with sonorant labials, I invoke the following constraint for the analysis of Tagalog.

(19) OCP-UM ([SON], [LAB]$_{(\mu\mu)1\ base}$)
No consecutive sonorant labial consonants in the initial bimoraic foot in the base of UM-affixation.

The Obligatory Contour Principle (OCP) has been a central component in the analysis of dissimilation in autosegmental phonology. In its most general formulation, the OCP prohibits adjacent identical autosegments (McCarthy, 1986; Goldsmith, 1976). Subsequent research has shown that phonological features may be subject to the OCP in specific domains and that phonological and morphological processes may be blocked if they create an OCP violation (Yip, 1988; Ito & Mester, 1986, among others). Such dissimilatory blocking is clearly at work in infixation in the languages at hand. Concretely, -um- in Tagalog is blocked if infixation were to create a sequence of sonorant labials in the first bimoraic foot of the base.

The form of OCP-type constraints in OT has been the subject of discussion. One approach has been to translate the autosegmental OCP in more-or-less equivalent form as a family of representational, rankable and violable constraints. This line has been taken in accounts of phonological dissimilation (Suzuki, 1998; Myers, 1997; Myers & Carleton, 1997, among others) and morphological haplology (Plag, 1998; Yip, 1998). One alternative is to capture the OCP as locally self-conjoined markedness constraints (Ito & Mester, 1998; Alderete, 1997; but see Bakovic 2000: fn 23); another theory is to place the OCP in a distinct perception grammar (Boersma, 1998, 2000). Although I adopt a representational form of OCP-type constraints in this paper, it is worthwhile to note that nothing central hinges on this choice for the present purposes.

The underlying idea of reference to the initial bimoraic foot in the base is that the domain of OCP-type markedness constraints may be positional. This idea flows from recent work in OT on positional faithfulness. The central idea behind this notion is that faithfulness constraints may be relativized to certain prominent positions such as stressed or root-initial syllables (see McCarthy, 2002b; Beckman, 1997; Casali 1997, among others, for discussion). Positional domain specification has been argued to be advantageous for markedness constraints as well. Thus, according to de Lacy (2002), positional markedness constraints assign a violation for candidates that have a position x and a property y in combination (see also Smith, 2003, Lombardi, 2001, among others). The OCP-type constraint in (19) assigns violations when consecutive sonorant labials are in the first bimoraic foot of the morphological base. As is discussed below, not only Tagalog, but also the data from Chamorro and Toba Batak support the idea that OCP-type markedness constraints may be sensitive to position.

The constraint in (19) replaces Orgun & Sprouse's (1999) OCP-um in the constraint hierarchy of Tagalog. Consider the tableau in (20) for the near-minimal pair *m+um+ura vs. ʔ+um+uwi. The base-initial bimoraic foot is shown through the parentheses.

(20) Segmental blocking of -um- in Tagalog

		OCP-UM ([SON], $[LAB]_{(□□)1\ base}$)	ALIGN-BY-□ (UM, Wd, L)	M-PARSE(UM)
	Input: /um, mura/			
	a. (mumu)ra	*!		
	b. (muru)ma		*!	
☞	c. ⊙			*
	Input: /um, ʔuwi/			
☞	d. (ʔumu)wi			
	e. (ʔuwu)mi		*!	
	f. ⊙			*!

Infixing -um- after a sonorant labial within the base-initial bimoraic foot incurs a fatal violation of OCP-UM ([SON], $[LAB]_{(□□)1\ base}$), as seen in (20) (a), whereas no violation is incurred if the sonorant labial of the base is outside the initial foot, as in (20) (d). Hyper-infixation is excluded by the quantized ALIGN-BY-□ constraint, as seen in (20) (b) and (20) (e). Thus, the null output correctly wins for inputs like /um, mura/, as seen in (20) (c). Infixation before the first vowel is correctly optimal in comparison to the null output, however, when the sonorant labial is in medial or final position in the root, as shown in (20) (d) versus (20) (f).

In contrast to Chamorro, metathesis of <u>-um-</u> to <u>mu-</u> is not an option in Tagalog. Consequently, I propose that the constraint LINEARITY is ranked high to avoid metathesis in this language (however, see Horwood 2002 for a different approach to the role of LINEARITY constraints in the analysis of infixation). Given the preceding discussion, the constraint sub-hierarchy for Tagalog UM-infixation is as in (21). Note that constraints prohibiting insertion of consonants or vowels, although not shown in this hierarchy or below, are ranked high in Tagalog, Chamorro, and Toba Batak.

(21) Tagalog UM-infixation

LINEARITY, ONSET, OCP-UM ([SON], $[LAB]_{(\square\square)1\ base}$), ALIGN-BY-□(UM, Wd, L) » M-PARSE(UM) » NOCODA, *COMPLEXONSET

To sum up this section, I have recapitulated McCarthy's (2002a) argument that a refinement of the alignment constraint schema renders the CONTROL component unnecessary, at least for Tagalog. It has been shown, however, that the intended argument for the replacement of gradient alignment by quantized alignment does not quite go through. Tagalog <u>–um-</u> can be tackled if the evaluation of alignment constraints is done in terms of syllables instead of segments. As far as prosody is concerned, the prosodic constraint ONSET, not NOCODA, drives infixation primarily, given that Tagalog does not allow vowel-initial words. Nonetheless, NOCODA plays a subordinate, but important role in accounting for the placement of infixes when the base begins with a consonant cluster. Concerning the segmental blocking effect, the OCP-um constraint of Orgun & Sprouse (1999) has been shown to be empirically problematic. I have argued in the present account of Tagalog that OCP-type constraints may be sensitive to position, in particular to the first bimoraic stem of the base. Segmental conditioning of UM and IN infixation in Chamorro is addressed next.

3. Infixation and metathesized prefixation in Chamorro

The Chamorro affixes UM and IN are placed before the first vowel of the base when it begins with an obstruent. However, unlike Tagalog, sources on Chamorro have noted that UM and IN may occur as the metathesized CV prefixes <u>mu-</u> and <u>ni-</u>. This happens optionally with bases beginning with sonorant consonants (Chung, 1998; Gibson, 1992; Cooreman, 1987). This pattern has been confirmed in the author's fieldwork, but it has been found typically only for speakers aged 55 years or older. Representative data for this speaker group are presented and analyzed in section 3.1. Unlike in Tagalog, UM/IN-infixation with initial consonant clusters in Chamorro only allows placement after the cluster. The analysis of this data set is presented in §3.2. Another UM/IN pattern, used mainly by middle-aged speakers (35-55), shows infixation in all environments. This is treated in section 3.3. Finally, a third pattern, typically employed by younger speakers under age 35, shows a split between UM and IN depending on the type of initial consonant. UM may metathesize with any initial consonant, whereas IN shows metathesis only with sonorants. It is demonstrated that this pattern is elegantly accounted for if ALIGN-BY-SEG constraints are invoked in addition to ALIGN-BY-□. The necessity of this constraint split provides a compelling argument for the quantized, categorical alignment schema.

The number of speakers consulted for this study is too small to enable a viable quantitative sociolinguistic analysis. Thus, the age-grading reported here is to be understood as a broad trend only. Future sociolinguistic research may reveal solid statistical significance of the patterns at hand.

Chamorro, analogous to Tagalog, does not have vowel-initial words at the surface (Topping, 1973). Thus, words that are vowel-initial in the orthography have an initial glottal stop in the phonology. Glottal stop is phonemic in Chamorro and is represented orthographically in non-initial position through an apostrophe. This convention has been extended in this paper so that initial apostrophe stands for initial glottal stop in (22) and the Chamorro data below. Other grapheme-phoneme correspondences of note are: <ng> = /ŋ/, <y> = /dz/ and <ch> = [tʃ] or [ts]. Note that IN causes vowel fronting or umlaut in the base; this alternation is represented orthographically for the non-low vowels. Thus, for example, tungo' 'to know' and konne' 'to take' appear with umlaut as tiningo' and kinenne', respectively (see Klein, 2000; Halle & Vergnaud, 1987; Chung, 1983, among others, for further discussion of Chamorro umlaut).

3.1 Infixation and metathesis with initial sonorants

In the most frequently reported pattern, infixation varies with metathesized prefixation for both potential infixes if the base-initial consonant is a sonorant, as shown in (22).

(22) Infixation and optional metathesized prefixation

Obstruent-initial bases

(a)	tungo'		tiningo'		tumungo'	'to know'
(b)	faisen		finaisen		fumaisen	'to ask'
(c)	konne'		kinenne'		kumonne'	'to take'
(d)	'ayek		'inayek		'umayek	'to choose'
(e)	'espiha		'inespiha		'umespiha	'to seek'

Sonorant-initial bases

(f)	li'e'	(i)	lini'e'	(i)	lumi'e'	'to see'
		(ii)	nili'e'	(ii)	muli'e'	
(g)	risibi	(i)	rinisibi	(i)	rumisibi	'to receive'
		(ii)	nirisibi	(ii)	murisibi	
(h)	na'i		nina'i	(i)	numa'i	'to give'
				(ii)	muna'i	
(i)	nginge'	(i)	ngininge'	(i)	nguminge'	'to smell'
		(ii)	ninginge'	(ii)	munginge'	

IN and UM are placed as infixes after obstruents, as shown in (22) (a) through (22) (e). In contrast, sonorant-initial bases license infixal -in- and -um- as well as prefixal ni- and mu-, as shown in (22) (f) through (22) (i). Note that the shape and placement of IN and UM are ambiguous for n-initial and m-initial bases, respectively. Thus, items like nina'i 'to give' and mumaolek 'become good' are ambiguous between n+in+a'i versus ni+na'i and m+um+aolek versus mu+maolek, respectively.

The comparison between Tagalog and Chamorro infixation shows that there is a sharp distinction in response to the dissimilatory environments. Whereas we observe absolute ungrammaticality in Tagalog, Chamorro employs prefixation-cum-metathesis as an alternative realization. Importantly, the onset requirement demands consonant-vowel sequences in Chamorro infixation and in the metathesis alternative. Consequently, a major aim of this paper is to account for the distinct patterns in Chamorro and Tagalog.

Note that UM and IN serve multiple, distinct morphosyntactic functions in Chamorro. Thus, it is sensible to refer to these affixes through the archiphonemic notation employed in this paper. According to Gibson (1992), UM serves to mark singular agreement and ergative focus, whereas IN is used to mark passive and nominalization (see Chung, 1998; Cooreman, 1987 and Topping, 1973 for alternative views). These morphosyntactic details are abstracted away from in the glosses presented in this paper. Importantly, there seems to be no indication that the morphophonological patterning depends on the morphosyntactic functions of UM and IN in Chamorro.

It has been observed for Tagalog that um-infixation is blocked when there is a base-initial sonorant labial, but infixation occurs when the sonorant labial is farther inside the base. A similar effect is observed in Chamorro. Metathesized prefixation is an option only when the sonorant is base-initial, but not when it is inside the base. Analogous to Tagalog, I propose that the alternative to infixation may be invoked in Chamorro only within the first bimoraic foot of the base. This is illustrated in the data in (23).

(23) Infixation before sonorants in Chamorro

(a)	konne'	(kine)nne'	(kumo)nne'	'to take'
(b)	tungo'	(tini)ngo'	(tumu)ngo'	'to know'
(c)	sangan	(sina)ngan	(suma)ngan	'to tell'
(d)	tulaika	(tini)laika	(tumu)laika	'to change'

Given the foot structure in (23), variants to infixation are not attested when the base sonorant is outside the base-initial bimoraic foot. Furthermore, no metathesis is reported for sequences of sonorants in roots, as shown in (24).

(24) Chamorro roots with consecutive sonorants (data from Topping et al., 1975)

(a)	lumuhu	'April'	(c)	lumos	'drown'
(b)	munga	'not'	(d)	nira'	'leek'

For the formal analysis of the Chamorro pattern, I invoke the quantized, categorical alignment constraints in (25) and (26).

(25) ALIGN-BY-□(UM, Wd, L)
'No syllable stands between the left edge of UM and the left edge of Wd.'

(26) ALIGN-BY-□(IN, Wd, L)
'No syllable stands between the left edge of IN and the left edge of Wd.'

The constraint in (25) is familiar from the discussion of Tagalog. The constraint in (26) is different only in that it refers to the affix IN. For practical purposes, the two ALIGN-BY-□ constraints need not be distinguished in the tableaux to follow.

Note that UM and IN are the only segmentally pre-specified infixes in Chamorro and that they behave largely analogous with respect to infixation and metathesis. Thus, an OCP-type constraint in which general reference to an affix is made may be invoked in Chamorro. Note that a more complete analysis of Tagalog may show that this language gets by without the specific reference to UM in the OCP-type constraint as well.

(27) OCP-AFF ($[\text{SON}]_{(\square\square)1\ \text{BASE}}$)
No consecutive sonorant consonants in the initial bimoraic foot in the base of affixation.

We may now consider the infixation option with sonorant-initial bases. Infixation with obstruent-initial bases is optimized analogously. In addition to foot structure, syllabification is indicated by the periods.

(28) Infixation with sonorant-initial bases

	/um, na'i/	ONSET	ALIGN-BY-□	LINEARITY	OCP-AFF ($[\text{SON}]_{(\square\square)1\ \text{BASE}}$)
	a. um.(na.'i).	*!			
☞	b. (nu.ma).'i.				*
	c. mu.(na.'i).			*!	
	d. (na.'u).mi.		*!		

The high ranking of the constraint ONSET prohibits the prefixation of the affix in its VC input form as shown in (28) (a). The ALIGN-BY-□ constraint is essential in disallowing hyperinfixation a syllable or more away from the left edge of the word, as shown in (28) (d). In the tableau in (28), the constraint LINEARITY outranks the OCP-type constraint OCP-AFF ($[\text{SON}]_{(\square\square)1\ \text{BASE}}$). With this ranking, metathesized candidates such as (28) (c) incur a fatal violation of the constraint LINEARITY. The optimal candidate (28) (b) only incurs a violation of the low ranking constraint OCP-AFF ($[\text{SON}]_{(\square\ \square)1\ \text{BASE}}$). This is because infixation results in the placement of the affix after the base-initial sonorant within the first bimoraic foot. The metathesized candidate (28) (c), however, passes the constraint OCP-AFF ($[\text{SON}]_{(\square\square)1\ \text{BASE}}$) because the prefix mu- appears outside the base and so is not within the positional domain of this constraint. The infixed as well as the metathesized candidate passes the ALIGN-BY-□ constraint because the affix is just a segment away from the beginning of the word or flush with it, respectively.

Recent work on metathesis has highlighted the fact that this process is synchronically productive in a good number of languages. OT work has focused on metathesis driven by prosodic constraints (McCarthy, 2000; Hume, 1998), but it has also shown that perceptual, segmental conditions may contribute as well (Hume, 1999). Phonetic factors also play a role, as demonstrated in Blevins & Garrett (1998). Generally in the OT approach, the preservation of the linear order of segments in the mapping from input to output is demanded by the anti-metathesis constraint LINEARITY. Metathesis in this approach occurs when higher-ranking markedness constraints force LINEARITY to be violated. Note that, following Hume (1998) and McCarthy (2000), items subject to metathesis have a single input form with specified linear order. In the present case the input forms of UM and IN are um and in, but the surface forms of these affixes vary between -um- and mu- and -in- and ni-, respectively.

Analogous to the analysis of Tagalog variation above, the constraints LINEARITY and OCP-AFF ($[\text{SON}]_{(\square\square)1\ \text{BASE}}$) are non-crucially ranked to account for the observed variation in Chamorro. Recall that non-crucial ranking optimizes variant candidates if higher-ranking constraints are not involved in the decision. Consequently, the ranking LINEARITY » OCP-AFF ($[\text{SON}]_{(\square\square)1\ \text{BASE}}$) optimizes the infixed candidate with initial sonorants as seen in (28). The reverse ranking allows for the prefixed metathesized candidate, as shown in (29).

(29) Metathesized prefixation with sonorant-initial bases

	/um, na'i/	ONSET	ALIGN-BY-□	OCP-AFF ($[\text{SON}]_{(\square\square)1\ \text{BASE}}$)	LINEARITY
	a. um.(na.'i).	*!			
	b. (nu.ma).'i.			*!	
☞	c. mu.(na.'i).				*
	d. (na.'u).mi.		*!		

Analogous to (28) above, prefixation of um- and hyperinfixation incur fatal violations of ONSET and ALIGN-BY-□, respectively. The difference in tableau (29) is that the OCP-type constraint OCP-AFF ($[\text{SON}]_{(\square\square)1\ \text{BASE}}$) incurs a fatal violation for the infixed candidate because it is ranked above LINEARITY. The desirable result is that the metathesized prefixed candidate is optimal because it only incurs a violation of the low-ranking anti-metathesis constraint. In other words, metathesis emerges as optimal to satisfy the onset requirement in cases where infixation violates the OCP-type constraint, and the linearity constraint is ranked low.

The tableau in (30) shows how the constraints invoked explain the absence of metathesized prefixation with obstruent-initial bases.

(30) Infixation with obstruent-initial bases

	/um, tungo'/	ONSET	ALIGN-BY-□	LINEARITY	OCP-AFF ([SON]$_{(□□)1\ BASE}$)
	a. um.(tu.ngo)'.	*!			
☞	b. (tu.mu).ngo'.				
	c. mu.(tu.ngo)'.			*!	
	d. (tu.ngu).mo'.		*!		

As before, prefixation of um- and hyperinfixation are disallowed because they incur fatal violations of ONSET and ALIGN-BY-□, respectively. Note, however, that correctly infixed candidates such as t+um+ungo' pass the constraints LINEARITY and OCP-AFF ([SON]$_{(□□)1\ BASE}$). In particular, they pass the OCP-type constraint because there is no sequence of sonorants inside the base-initial bimoraic foot. Thus, any permutation of LINEARITY and OCP-AFF ([SON]$_{(□□)1\ BASE}$) optimizes infixation with obstruent-initial bases, whereas metathesis is never optimal, just as required.

The preceding discussion enables us to understand the phonological conspiracy in the optimization of infixation versus prefixal metathesis. The onset requirement results in infixation of VC affixes such as UM and IN, all other things being equal. In response to a potential OCP-type violation, metathesis, rather than ineffability as in Tagalog, is the next best option because it also satisfies the onset constraint and the M-Parse requirement, even though it violates the low-ranking linearity constraint.

3.2 Infixation with initial consonant clusters

Special note must be taken of Chamorro roots that begin with clusters of consonants. The Chamorro infixes are placed after initial consonant clusters throughout, that is, no Tagalog-style variation has been reported in this context. The data in (31) (a) are taken from Anderson (1992, 1993) who attributes them to Sandra Chung. (31) (b) is from Gibson (1992).

(31) Infixation after initial consonant clusters

(a)	tristi	trinisti	trumisti	'sad'
	planta		plumanta	'set the table'
(b)	brabu	brinabu	brumabu	'healthy'

There is no metathesis option reported for initial clusters, even though placement after a sonorant licenses metathesis elsewhere.

I propose to resolve this issue by locally conjoining the constraints LINEARITY and *COMPLEXONSET. Under local conjunction, as proposed in Smolensky (1993), two constraints are conjoined into one composite constraint. This composite constraint is violated if and only if both of its complements are violated in some domain. Local conjunction thus punishes the "worst of the worst" in that only candidates that violate both conjuncts incur a violation of the conjoined constraint. Local conjunction has been usefully invoked to account for derived-environment effects (Lubowicz, 2002) and opacity effects (see Kager, 1999 for discussion). A conjoined constraint does not replace its conjuncts, but instead it is universally ranked above the component constraints.

(32) Universal ranking schema in local constraint conjunction

$[C_1 \;\&\; C_2]_{□}$ » C_1, C_2

The ranking for the constraint conjunction in Chamorro is given in (33).

(33) Local conjunction of LINEARITY and *COMPLEXONSET

[LINEARITY & *COMPLEXONSET]$_{Wd}$ » LINEARITY » *COMPLEXONSET

Given the domain of the word, the effect of the conjunction is to exclude candidates that have complex onsets and metathesis, as shown in (34).

(34) Infixation with initial consonant clusters

	/um, brabu/	LIN & *COMPONS	NOCODA	LIN	OCP	*COMPONS
	a. (bum)rabu		*!			
	b. mu(brabu)	*!		*		*
☞	c. (bruma)bu				*	*

Metathesis and complex onsets are out via the violation of the locally conjoined constraint, as shown in (34) (b). The splitting up of the consonant cluster through the infix results in a fatal violation of the NOCODA constraint in Chamorro, as shown in (34) (a). Consequently, infixation after a consonant cluster is the only option, even though more than one sonorant is found in the domain of the OCP. This result is shown in (34) (c). Note that I did not test the behavior of the infixes with consonant clusters in my fieldwork. It is a question for future empirical work on Chamorro to explore the behavior of the infixes in this context in more detail and, in particular, to investigate if there is age-grading similar to the data presented further below.

Given that hyperinfixation is not allowed with IN or UM, the alignment constraints ALIGN-BY-□(IN, Wd, L) and ALIGN-BY-□ (UM, Wd, L) are ranked together. Furthermore, M-PARSE must be ranked high since the null output is not an option in Chamorro. Analogous to Tagalog, constraints forbidding insertion of consonants or vowels must be ranked high. For the other aspects of the pattern at issue, the constraint sub-hierarchy is (35).

(35) Infixation and optional metathesized prefixation

[LINEARITY & *COMPLEXONSET]$_{Wd}$ » ONSET, ALIGN-BY-□(UM, Wd, L), ALIGN-BY-□(IN, Wd, L) » NOCODA » OCP-AFF ([SON]$_{(□□)1\ BASE}$), LINEARITY » *COMPLEXONSET

It has been demonstrated in the two preceding sections that infixation in Chamorro may be blocked because of an OCP-type constraint which penalizes consecutive sonorant consonants in the first bimoraic foot of the base. The metathesized prefixation alternative is available through the high ranking of M-Parse and the low ranking of Linearity. The high ranking of the Onset constraint drives the infixation and the metathesis option in that both processes create syllables with onsets. Metathesis with initial clusters is disallowed because of the local conjunction of the constraints against metathesis and against complex onsets.

3.3 Infixation without metathesis

Recall that metathesized prefixation is not attested for a middle-aged group of speakers, roughly between 35 to 55 years of age. This infixation-only pattern may be accounted for straightforwardly through constraint re-ranking. Consider the data in (36).

(36) Infixation only

Obstruent-initial bases

(a)	tungo'	tiningo'	tumungo'	'to know'
(b)	faisen	finaisen	fumaisen	'to ask'
(c)	konne'	kinenne'	kumonne'	'to take'
(d)	'ayek	'inayek	'umayek	'to choose'
(e)	'espiha	'inespiha	'umespiha	'to seek'

Sonorant-initial bases

(f)	li'e'	lini'e'	lumi'e'	'to see'
(g)	risibi	rinisibi	rumisibi	'to receive'
(h)	na'i	nina'i	numa'i	'to give'
(i)	nginge'	ngininge'	nguminge'	'to smell'

This group of speakers places UM and IN before the first vowel without concern for the creation of consecutive sonorants. We only need to rerank the constraint Linearity to a high position in the hierarchy to account for this. For illustrative purposes, IN instead of UM has been chosen in the following tableau. Foot and syllable structure has been omitted to improve readability.

(37) Infixation only through re-ranking of Linearity

	Input: /in, li'e'/	Onset	Linearity	Align-by-□	OCP-Aff ([SON]$_{(□□)1\ BASE}$)
	a. inli'e'	*!			
☞	b. lini'e'				*
	c. nili'e'		*!		
	d. li'ine'			*!	

The constraint LINEARITY in this pattern is ranked with the ONSET constraint, analogous to Tagalog. Consequently, metathesis as in candidate (37) (c) is not available to create CV syllables, regardless of which segment begins the root. Infixation as in (37) (b) is correctly the only option. The constraint sub-hierarchy for the infixation-only pattern is given in (38).

(38) Infixation only

[LINEARITY & *COMPLEXONSET]$_{Wd}$ » LINEARITY, ONSET, ALIGN-BY-□(UM, Wd, L), ALIGN-BY-□(IN, Wd, L) » NOCODA » OCP-AFF ([SON]$_{(□□)1\ BASE}$) » *COMPLEXONSET

The essential change in comparison to the hierarchy in (35) is the higher ranking of the linearity constraint.

3.4 Infixation and metathesis

Chamorro speakers under the age of 35 have been found to use infixation and metathesized prefixation with UM regardless of the base-initial segment. Note that IN exhibits metathesis only with base-initial sonorants, as with the other groups of speakers. This distinction in the behavior of UM is shown to support McCarthy's (2002a) idea that alignment constraints may refer to syllables or segments in calculating the distance to a given edge.

(39) Infixation and optional metathesis: UM versus IN

Obstruent-initial bases

(a)	tungo'		tiningo'	(i)	tumungo'	'to know'
				(ii)	mutungo'	
(b)	faisen		finaisen	(i)	fumaisen	'to ask'
				(ii)	mufaisen	
(c)	konne'		kinenne'	(i)	kumonne'	'to take'
				(ii)	mukonne'	
(d)	'ayek		'inayek	(i)	'umayek	'to choose'
				(ii)	mu'ayek	
(e)	'espiha		'inespiha	(i)	'umespiha	'to seek'
				(ii)	mu'espiha	

Sonorant-initial bases

(f)	li'e'	(i)	lini'e'	(i)	lumi'e'	'to see'
		(ii)	nili'e'	(ii)	muli'e'	
(g)	risibi	(i)	rinisibi	(i)	rumisibi	'to receive'
		(ii)	nirisibi	(ii)	murisibi	
(h)	na'i		nina'i	(i)	numa'i	'to give'
				(ii)	muna'i	

(i)	nginge'	(i)	ngininge'	(i)	nguminge'	'to smell'
		(ii)	ninginge'	(ii)	munginge'	

We observe in this pattern that IN appears as the infix with all initial segments and optionally as the metathesized prefix with initial sonorants, just as in the pattern in §3.1. UM, however, may appear as the infix or the metathesized prefix regardless of the initial segment.

Note that the ALIGN-BY-□ constraints invoked so far only exclude hyper-infixation of a syllable or more inside the left edge of the word. The data in (39) may be understood as showing that metathesized prefixation competes with a much tighter measure, namely alignment by segment. I propose to account for this pattern by invoking the quantized, categorical ALIGN-BY-SEG constraint given in (40).

(40) ALIGN-BY-SEG(UM, Wd, L)
'No segment stands between the left edge of UM and the left edge of Wd.'

Metathesized prefixation passes the ALIGN-BY-SEG constraint because the affix is placed right at the beginning of the word. Infixation, however, violates this constraint. Consequently, the non-crucial ranking of the ALIGN-BY-SEG(UM, Wd, L) and the LINEARITY constraint may optimize metathesized prefixation as well as infixation with obstruent-initial bases. No ALIGN-BY-SEG constraint is invoked for IN, given that this affix does not exhibit the wider availability of metathesis. Consider now the tableau in (41).

(41) UM-infixation with obstruent-initial roots

	Input: /um, tungo'/	ONSET	ALIGN-BY-□	LINEARITY	ALIGN-BY-SEG (UM, Wd, L)
	a. umtungo'	*!			
☞	b. tumungo'				*
	c. mutungo'			*!	
	d. tungumo'		*!		

When the constraint LINEARITY outranks ALIGN-BY-SEG (UM, Wd, L), the metathesized candidate is out and the infixed candidate wins, as the comparison of (41) (c) with (41) (b) shows. Consider the reverse ranking of these two constraints in (42).

(42) CV-prefixation with obstruent-initial roots

	Input: /um, tungo'/	ONSET	ALIGN-BY-□	ALIGN-BY-SEG (UM, Wd, L)	LINEARITY
	a. umtungo'	*!			
	b. tumungo'			*!	
☞	c. mutungo'				*
	d. tungumo'		*!		

The ranking of the alignment constraint ALIGN-BY-SEG above LINEARITY favors metathesized prefixation as in (42) (c) as a variant.

Forms with initial obstruents are of course never in violation of the OCP-type constraint familiar from the earlier discussion. However, the effects of this constraint must be considered for forms with initial sonorants. The tableau in (43) treats the options for sonorant-initial bases with UM.

(43) Infixation and metathesis of UM with sonorant-initial bases

	Input: /um, na'i/	ONSET	ALIGN-BY-□	ALIGN-BY-SEG (UM, Wd, L)	LIN	OCP-AFF ([SON]$_{(□□)1\ BASE}$)
	a. umna'i	*!				
☺	b. numa'i			*		*
☺	c. muna'i				*	
	d. na'umi		*!			

Infixation and metathesis are correctly possible by freely permuting the ALIGN-BY-SEG and the linearity constraint, as indicated by the smiling faces. The OCP-type constraint is too low in the hierarchy to have an effect on the availability of infixation in this pattern. Note that IN does not behave distinctly for this group. This is predicted correctly, given that the ALIGN-BY-SEG constraint invoked has scope only over UM, but not IN.

The constraint sub-hierarchy in (44) accounts for the pattern which shows metathesis regardless of the base-initial segment with UM.

(44) Infixation and metathesis with UM

[LINEARITY & *COMPLEXONSET]$_{Wd}$ » ONSET, ALIGN-BY-□(UM, Wd, L),
ALIGN-BY-□(IN, Wd, L) » ALIGN-BY-SEG(UM, Wd, L), LINEARITY » NOCODA »
OCP-AFF ([SON]$_{(□□)1\ BASE}$) » *COMPLEXONSET

Note that the only difference of this hierarchy to the one for the older speakers in section 3.1 is that the ALIGN-BY-SEG constraint is ranked high enough so that it has a visible effect and that the linearity constraint is ranked with it to optimize the observed variants.

Topping (1973, 1980) has reported optional metathesized prefixation before bases beginning with nasals, but not with liquids. This pattern may easily be accounted for by invoking the OCP-type constraint OCP-AFF ([NAS]) utilized in the analysis of Toba Batak in §4 and putting it in the place of OCP-AFF ([SON]) in the constraint sub-hierarchy in (35). Thus, the unified approach to segmental effects in infixation pursued in this paper uncovers a connection between Toba Batak and the variety of Chamorro reported in Topping's works. This is a desirable result, given that these languages are genetically and typologically related. Note that Topping (1973: 170) reports a single item of prefixation of mu- with a base beginning with glottal stop. Given that this option has not been reported anywhere else, I interpret this environment for metathesis as less than robust and, hence, I do not analyze it further in this paper.

To sum up this section, I have described and analyzed new and previously known variation patterns with regard to infixation and metathesized prefixation in Chamorro. The patterned variation among Chamorro speakers of different generations has been analyzed as an outcome of constraint ranking options. As pointed out by an anonymous reviewer, this begs the question why a given pattern is chosen by a particular generation. However, this interesting issue is beyond the scope of the present paper; hence, it must await the results of future research. The theoretical account presented makes crucial use of the split between ALIGN-BY-□ versus ALIGN-BY-SEG constraints. Given that this split is not possible in the gradient alignment framework of McCarthy & Prince (1993a), a strong argument for quantized, categorical alignment in the sense of McCarthy (2002a) is obtained through the present analysis of Chamorro. The positionally sensitive OCP-type constraints make possible a unified approach to the segmental conditioning of infixation in Chamorro and Tagalog. The availability of the metathesis and infixation options in Chamorro is accounted for as a phonological conspiracy driven by the need to parse the affixes UM and IN while creating syllables with onsets at the same time. These results are utilized in the analysis of Toba Batak presented in the next section.

4. Infixation and assimilated prefixation in Toba Batak

Crowhurst (2001; 1998) has concluded that the linear placement of -um- in the Austronesian language Toba Batak is in part determined by segmental factors. Consider the following evidence reproduced from Crowhurst (1998: 591 ff.). The data are primarily from Nababan (1981), but also from Percival (1981) and Crowhurst's work with a Toba Batak consultant.

(45)	Infixation		
(a)	deŋgan	dumeŋgan	'good/better'
(b)	tibbo	tumibbo	'tall/taller'
(c)	taŋis	tumaŋis	'to cry'
(d)	dʒou	dʒumou	'call/ call repeatedly'
(e)	sɔlsɔl	sumɔlsɔl	'regret/has regretted'
(f)	rahis	rumahis	'steep/steeper'
(g)	las	lumas	'hot/hotter'

The data in (45) show that Toba Batak -um- appears before the first vowel of the base when it begins with consonants other than labials and nasals. Initial labials and nasals, however, show a distinct pattern.

(46)	Prefixation with labial assimilation		
(a)	pir	uppir	'hard/harder'
(b)	bege	ubbege	'hear/have heard'
(c)	maɔl	ummaɔl	'difficult/more difficult'

(47) Prefixation with nasal assimilation

(d)	nɛaŋ	unnɛaŋ	'light/lighter'
(e)	ŋali	uŋŋali	'cold/colder'

The data in (46) and (47) show that -um- in Toba Batak surfaces as a prefix when the base begins with a labial or nasal consonant, respectively; um- assimilates to the following consonant in this case.

A closer look at the data in (45) also shows that the segmental conditioning of –um- is not observed when the nasal or labial is in base-medial position. Thus, d+um+eŋgan, t+um+aŋis, and t+um+ibbo in (45) (a) through (45) (c) show that UM in Toba Batak may precede nasals and labials without resulting in prefixation. This is analogous to Tagalog and Chamorro and is interpreted as further evidence for the sensitivity of the segmental effects to the domain of the base-initial bimoraic foot.

In contrast to Tagalog and Chamorro, Toba Batak admits vowel-initial bases freely (Nababan 1981). As the data in (48) show, there is gemination before such bases.

(48) Prefixation with gemination

(a)	uli	ummuli	'beautiful/more beautiful'
(b)	arga	ummarga	'expensive/more expensive'

Given the occurrence of vowel-initial bases, the ONSET constraint cannot be responsible for infixation in Toba Batak. Hence, Crowhurst (2001) invokes the CODACOND constraint provided in (49).

(49) CODACOND: Consonants are not licensed in syllable-final position.

Coda consonants can only be licensed in accordance with the CODACOND constraint if they are assimilated or linked to a following onset.

Assimilation is controlled through the IDENTITY constraints in (50).

(50) IDLAB: Output correspondents of an input labial segment are labial.
IDNAS: Output correspondents of an input nasal segment are nasal.

Crowhurst (2001) invokes the constraint in (51) to account for gemination with vowel-initial bases.

(51) UM≈□: um is the exhaustive content of a syllable.

An anonymous reviewer points out that the UM≈□ constraint may be understood as an instance of the 'emergence of the unmarked' (cf. McCarthy & Prince, 1995), given that morphological boundaries like to line up with syllable boundaries.

Gemination is controlled by the constraint in (52).

(52) *GEMINATE: Consonantal root nodes are uniquely dominated.

Crowhurst (2001) also invokes the gradient alignment constraint from McCarthy & Prince's (1993a) analysis of Tagalog, given for convenience in (53).

(53) Align-um
Align([um]$_{Af}$, L, stem, L)

Note that it is difficult to see how this constraint captures the phonologically conditioned allomorphs of UM in Toba Batak, given that it refers only to a single phonological surface form of the affix. For reasons of space, I abstract away from this issue in the recapitulation of Crowhurst's (2001) analysis to follow. In my own analysis, however, I invoke the archiphonemic notation UM.

Crowhurst's argument for crucial non-ranking rests on the fact that the analysis of infixation with consonant-initial bases in Toba Batak encounters a ranking paradox under the conventional conception of ranking. Consider the following tableaux reproduced from Crowhurst (2001).

(54) Correct optimization of prefixation with initial labials/nasals

	/um, bege/	CodaCond	Um≈□	IdLab	IdNas	Align-um
☞	a. ub.be.ge.				*	
	b. bu.me.ge.		*!			b
	c. um.be.ge.	*!				

The unassimilated candidate violates the coda condition, and the infixed candidate violates the constraint that demands that the morpheme –um- must be contained within a syllable. Consequently, assimilated prefixation as in (54) (a) emerges as optimal. However, the constraint sub-hierarchy displayed in (54) cannot account for the occurrence of infixation with bases that do not begin with labials or nasals, as indicated in (55) by the sad face symbol.

(55) Incorrect optimization of prefixation with other initial consonants

		CodaCond	Um≈□	IdLab	IdNas	Align-um
☹	a. ug.go.kan.	*		*	*	
	b. gu.mo.kan.	*	*!			g
	c. um.go.kan.	**!				

Note that the word-final coda in (55) incurs a violation of the coda condition in every candidate. Crucially, the two top-ranked constraints in (54) and (55) incur the same type of violations regardless of the initial segment of the base so that assimilated prefixation is always optimal. This result is incorrect, however, given that infixation occurs with bases not beginning with a labial or a nasal. The ranking paradox is that the tableau in (54) suggests that Um≈□ outranks the identity constraints, but the tableau in (55) would require the opposite ranking to achieve the correct result. Crowhurst approaches this ranking paradox by proposing that Um≈□ and the identity constraints are crucially non-ranked or co-ranked, as in tableau (56). For better readability, the constraint CODACOND and the candidates fatally violating it are omitted in this tableau.

(56) Crucial non-ranking

		IDLAB IDNAS UM≈□	ALIGN-um
☞	a. ub.be.ge.	* IDNAS	
	b. bu.me.ge.	* UM≈□	b!
	c. ug.go.kan.	* IDLAB *! IDNAS	
☞	d. gu.mo.kan.	* UM≈□	g

Under the co-ranking idea, violation marks from crucially non-ranked constraints are equally important and, hence, may accumulate to exclude candidates. As seen in (56) (a) versus (56) (b), infixation and assimilated prefixation incur one violation each of a co-ranked constraint with bases beginning with a labial or a nasal. The decision is passed on to the alignment constraint, which correctly favors the prefixed candidate. The co-ranked constraints force the decision with bases not beginning with a labial or nasal. Assimilation violates both identity constraints in these cases so that the infixed candidate correctly wins, as seen in (56) (c) versus (56) (d).

Although quite ingenious, the co-ranking idea requires reconsideration of the entire constraint component CON of the OT grammar, given that every constraint may now be ranked or not ranked with respect to any number of other constraints. This predicts a vastly inflated factorial typology that would appear to be too rich to be empirically or observationally adequate. I demonstrate in the remainder of this section that this powerful reorganization of CON is not necessary. Instead, it seems that Crowhurst's initial analysis fails because there is no dissimilatory constraint, even though the suspension of infixation in favor of assimilated prefixation in the environment of consecutive nasals or labials is clearly due to a dissimilatory effect. Consequently, I propose to account for the segmental effects in Toba Batak through OCP-type constraints that are analogous to the ones invoked for Tagalog and Chamorro. These constraints are given in (57) and (58).

(57) OCP-AFF ([LAB]$_{(□□)1\ BASE}$)
No consecutive labial consonants in the initial bimoraic foot in the base of affixation.

(58) OCP-AFF ([NAS]$_{(□□)1\ BASE}$)
No consecutive nasal consonants in the initial bimoraic foot in the base of affixation.

Given that it is not certain to which extent dissimilatory effects are found with other affixes in Toba Batak, the constraints in (57) and (58) are given a general formulation in that they do not refer to a specific affix. Note that the domain of the base-initial bimoraic foot is essential. Infixation is ungrammatical only when these segments are base-initial, but is allowed when they are inside the base, as the discussion of the data in (45) has shown. The presence of these OCP-type constraints in the hierarchy makes it possible for Crowhurst's UM≈□ constraint to be ranked low and, thus, to resolve the ranking paradox. For the present analysis, I invoke the quantized ALIGN-BY-□ constraint that is also at work in Tagalog and Chamorro, repeated for convenience in (59).

(59) ALIGN-BY-□(UM, Wd, L)
'No syllable stands between the left edge of UM and the left edge of Wd.'

With these constraints, prefixation-cum-assimilation is optimized as shown in (60). Note that the foot and syllable structures have been given according to the conventions utilized above.

(60) Assimilated prefixation

	Input: /um, bege/	CODA COND	OCP-AFF ([LAB]$_{(□□)1\ BASE}$)	IDNAS	ALIGN-BY-□	UM≈□
☞	a. ub.(be.ge).			*		
	b. (bu.me).ge.		*!			*
	c. um.(be.ge).	*!				

As before, the coda condition excludes candidates in which unassimilated UM has been prefixed, as shown in (60) (c). OCP-type constraints like the constraint OCP-AFF ([LAB]$_{(□□)1\ BASE}$) are now responsible for the exclusion of the infixed candidate with bases with initial labials or nasals, as seen in (60) (b). As a consequence, UM≈□ may be ranked low, thus resolving the ranking paradox of Crowhurst's analysis. Assimilated prefixation violates only the low-ranking identity constraints and is, hence, correctly optimal. The same constraint sub-hierarchy optimizes infixation with bases not beginning with labials or nasals, as shown in the tableau in (61).

(61) Infixation

	Input: /um, gokan/	CODA COND	OCP-AFF ([LAB]$_{(□□)1\ BASE}$)	IDNAS	ALIGN-BY-□	UM≈□
	a. ug.(go.ka)n.	*		*!		
☞	b. (gu.mo).kan.	*				*
	c. um.(go.ka)n.	**!				

Infixation after non-labials or non-nasals passes the OCP-type constraints, as seen in (61) (b). Thus, infixation in such cases only violates the low-ranking UM≈□ constraint and is, hence, correctly optimal.

The same approach accounts for gemination with vowel-initial bases.

(62) Gemination with vowel-initial bases

	Input: /um, uli/	CODA COND	OCP-AFF ([LAB]$_{(□□)1\ \mathrm{BASE}}$)	ALIGN-BY-□	UM≈□	*GEMINATE
☞	a. um.(mu.li).					*
	b. (u.mu).li.				*!	
	c. (u.lu).mi.			*!		
	d. um.(u.li).	*!				
	e. (u.u)m.li.			*!		

The correctly geminated candidate in (62) (a) only violates the low-ranking antigemination constraint. This type of candidate passes the OCP-type constraints because the affix is outside the base. All other candidates violate the coda condition or the alignment constraint against hyperinfixation. Analogous to Crowhurst's account, illicit gemination with infixation as in *g+um.+mo.kan. loses in comparison to simple infixation as in g+u.m+o.kan. because gemination in the infixation case does not improve on the UM≈□ constraint.

The constraint sub-hierarchy for Toba Batak resulting from the present analysis is given in (63).

(63) Infixation and assimilated prefixation

CODACOND, OCP-AFF ([LAB]$_{(□□)1\ \mathrm{BASE}}$), OCP-AFF ([NAS]$_{(□□)1\ \mathrm{BASE}}$) » IDLAB, IDNAS, ALIGN-BY-□(UM, Wd, L) » UM≈□ » *GEMINATE

The theoretical point of this section is that the positionally sensitive OCP-type constraints invoked in the analysis obviate the need for crucial non-ranking or co-ranking of constraints. This is a welcome result, given that crucial non-ranking increases the power of the OT grammar dramatically. Positionally sensitive OCP-type constraints, on the other hand, flow naturally from the idea that the OCP may hold only over certain domains and that markedness constraints may be sensitive to position.

5. Infixation as rule-based metathesis?

Halle (2001) has argued for a rule-based alternative to the OT approach to infixation, in particular Prince & Smolensky's (1993) account of Tagalog. In this section, I critique Halle's proposal in some detail and I show that his approach to infixation is disadvantageous in comparison to the OT framework laid out in this paper.

Halle proposes that the effect of infixation is achieved because the affixes UM and IN are subject to Onset Metathesis, that is, the rule-based "process that permutes the phonetic features linked to the onsets of two consecutive morphemes" (Halle, 2001: 154). Thus, Tagalog forms such as s+um+ulat 'one who writes' would result from the Onset Metathesis rule having applied to the input /mu+sulat/. For Halle's proposal to go through, the underlying form of UM and IN must be understood as /mu/ and /ni/, respectively, instead of the traditional understanding as /um/ and /in/ adopted in this paper. However, the CV underlying form seems problematic given that the surface form is [um] and [in] in Tagalog and Chamorro in the vast majority of environments.

One principled empirical problem with Halle's proposal is that it is impossible under the Onset Metathesis approach to account for the variation observed in Tagalog. That is, it is not possible to map /mu+preno/ 'to brake' and /mu+gradwet/ 'to graduate' onto [p+um+reno] and [g+um+radwet], respectively, because the intercalating element from the base morpheme is not the whole syllable onset, but only one consonant from it. The output-driven OT approach, however, does not encounter this problem. This is because the demand for output onsets posited in the OT approach may be satisfied by a single consonant or a consonant cluster.

For his argument against infixation in Chamorro, Halle (2001) draws on alternatives to the thrust of the description in Topping (1973) considered by Topping himself. Halle focuses on the following passage:

> It is interesting that the canonical form of both these affixes is just the opposite of what one would expect in a language that is basically CVCV. It is tempting to posit underlying prefixes mu- and ni- for these infixes with appropriate rules of metathesis. (Topping 1973: 170)

In response to this passage, it should be pointed out that Chamorro is not 'basically a CV language'. Topping demonstrates elsewhere in his book (p. 39) that Chamorro is a CCVC language. In addition to consonant clusters in the onset, particularly at the beginning of roots, nasals and voiceless obstruents are routinely admitted in coda position. This holds true in the Spanish-derived elements of the Chamorro lexicon even more so than in the indigenous lexemes. It is easy to find indigenous lexemes and affixes with syllable codas, as in the examples in (64) (a) and (64) (b), respectively.

(64) Syllable codas in Chamorro (data from Topping, 1973; Topping et al., 1975)

(a) niyok	'coconut'	fatfat	'four'	hotkon	'pillar'
(b) ‘an-	‘left over’	chat-	'not very'	fa'-	'change to'

Given data such as (64) and an abundance of CC onsets, it seems clear that Chamorro is not merely a CV language. Instead, onsets are obligatory which is why /um/ and /in/ prefer to be infixed.

Another argument made by Topping (1973) and Halle (2001) for the alleged underlying CV structure of UM and IN is the interaction with Nasal Substitution (see Piggott, 2002; Hyman, 2001; Pater, 1999; Archangeli et al., 1998 for recent discussion of this pattern in other Austronesian languages). In Chamorro Nasal Substitution, the final nasal of certain affixes such as the plural markers fan- and man- acquires the place of articulation of the voiceless obstruent of the following base via coalescence, as in (65).

(65) Nasal substitution with *man-/fan-* at the prefix-base boundary

(a)	chagi	mañagi	'to try'
(b)	peska	fameska	'to fish'

Nasal substitution does not take place, however, with nominalizing IN, as shown in (66).

(66) Absence of Nasal Substitution with nominalizing IN (Topping, 1973: 171 ff.)

(a)	fan-	'plural marker'
	chomma'	'forbid'
	IN	'nominalizing infix'
(b)	fanch<u>in</u>emma'	'forbidden things'

The question is why Nasal Substitution does not occur to result in *<u>fañ-in-emma'</u>. Halle (2001) argues that IN is underlyingly /ni/ and that Nasal Substitution is ordered before Onset Metathesis. Given the input /fan-ni-chemma'/, Nasal substitution does not apply, and Onset Metathesis produces the surface form [fanchinemma']. This account is problematic, however, given that Nasal Substitution in Chamorro is generally absent with multiple prefixation. Consider the following quote from Topping (1973).

> The regular morphophonemic changes caused by <u>man-</u> will be observed except where another type of affixation, e.g., reduplication has already taken place, as in the case of the derived nouns. (Topping, 1973: 235)

Thus, we do not find Nasal Substitution at the prefix-prefix boundary in general, as shown in (67).

(67) Absence of Nasal Substitution with multiple prefixation (Topping, 1973: 235)

paki	'to shoot'	papaki	'gunner'	<u>man</u>papaki	'gunners'
fahan	'to buy'	fafahan	'buyer'	<u>man</u>fafahan	'buyers'

The data in (67) show that the absence of Nasal Substitution is not conditioned by rule ordering or the underlying phonological form of the affix closest to the base, as claimed by Halle, but instead it is a function of morphological complexity. This phenomenon occurs in other Austronesian languages as well. Thus, in the celebrated case of Indonesian Nasal Substitution, nasals merge with voiceless obstruents, as exemplified in (68) (a). However, as pointed out in Piggott (2002), Nasal Substitution does not take place at the prefix-prefix boundary, as shown in (68) (b).

(68) Indonesian nasal substitution (data from Lapoliwa 1981)

(a)	məŋ-pilih	[məmilih]	'to choose'
(b)	məŋ-pər-bəsar	[məŋpərbəsar]	'to change'

In short, given the dependency of Nasal Substitution on morphological complexity, no argument for the supposed CV input form of IN can be derived from the lack of Nasal Substitution in morphologically complex items such as <u>fan+ch+in+emma'</u> in Chamorro.

According to Halle (2001), the underlying form /mu/ appears as [mu] if the following stem begins with a nasal (cf. the variants in Topping's 1973 data referred to above). Elsewhere, /mu/ would undergo Onset Metathesis to appear as [um] on the surface. Note that it seems impossible in Halle's account to connect the alternative forms in Chamorro to dissimilation. Halle's rule of Onset Metathesis would have to be specified to apply to [-nasal] onsets in Chamorro only (given the full set of data considered in this paper, the rule would apply to [-son] onsets). This approach would make it accidental that the mu- and ni- alternatives occur at the surface only in proximity to another nasal or sonorant. Instead, the analysis laid out in this paper offers the explanation that the appearance of the CV variants is due to a dissimilatory effect. This explanation is not available in Halle's approach.

The OT analysis invoking OCP-type constraints presented in this paper has shown that an elegant analysis of the facts from Chamorro, Tagalog and Toba Batak invoking the concept of infixation is possible. Halle, on the other hand, has claimed that his Onset Metathesis account avoids extra complications that arise in the infixation alternatives. In particular, he has claimed that the infixation analysis must stipulate that these affixes are subject to the alternation between CV and VC. However, the OT analysis in this paper has shown that no such stipulation is required. Instead, the change in the phonological form of the affixes falls out naturally from the ONSET constraint. Furthermore, Halle's Onset Metathesis rule must be specified to apply to certain CV prefixes only, given that undisputed CV prefixes surface as such in the languages at hand. This stipulation is not necessary in the OT account. Instead, VC prefixes/infixes undergo the processes under discussion without stipulation, and the surface phonological form and positioning of a given affix fall out from constraint ranking, the core determinant of the grammar of any language in the OT framework.

In short, I have demonstrated in this section that the Onset Metathesis rule-based analysis of UM and IN in Tagalog, Chamorro and Toba Batak is empirically problematic and theoretically inferior to the OT account. Importantly, the present analysis continues to make available the functionalist explanation that infixation after word-initial consonants results in a more satisfactory syllable structure than prefixation. This maintains the significant advantage of OT that such insights are a keystone in the theory of edge-oriented infixation.

6. Conclusion

The results of the comparative approach taken in this paper are of interest to general phonological and morphological typology and to descriptive and theoretical Austronesian linguistics. The OT analysis in this paper contributes to the ongoing debate concerning constraints and their ranking and evaluation. More specifically, the present analysis illuminates the form of alignment constraints, supports the growing research on positional markedness constraints, and deepens the understanding of segmental phonology in Optimality Theory. As far as research in the interface of phonology and morphology is concerned, the paper has shown that infixation is not exclusively Prosodic Morphology. Whereas prosodic constraints play an important role in the positioning of infixal morphemes, the placement and shape of such morphemes is to a significant degree determined by the segmental phonology as well. The original OT analysis in McCarthy & Prince (1993a, b) and Prince & Smolensky (1993) claimed that infixation is driven by the no-coda requirement. The research in this paper shows that the prosodic constraints driving infixation are not uniform. Whereas Toba Batak employs a coda condition, the central prosodic force in Tagalog and Chamorro infixation is the preference for onsets.

Dissimilation is the central segmental process affecting infixation in the languages studied. It is evident that the grammars of languages may show different responses if infixation would create a dissimilatory environment. Note that the feature structure of the sounds making up the infixes remains stable in the languages at hand, unlike other segmental dissimilation effects where alternations in phonological features are observed. Three responses to avoid dissimilatory environments have been observed and analyzed in this paper: Ineffability in Tagalog, prefixation with metathesis in Chamorro, and prefixation with assimilation in Toba Batak. Importantly, the dissimilatory effects are asymmetrical in that offending sequences involve segments at the left periphery of the morphological base, but not further inside it. In addition, the potential infixes are outside of the scope of dissimilation when they appear as prefixes. This array of facts has been accounted for by the proposal that OCP-type constraints may invoke positional domains such as the base-initial bimoraic foot. Conceived in this way, segments further inside the base do not offend because they are outside the initial bimoraic foot. Prefixes, on the other hand, are acceptable alternatives because they are attached to the morphological base, but are not contained in it. OCP-type constraints are markedness constraints in that they prohibit a given structure. Thus, the proposal of positional domains for OCP-type constraints has been able to draw on recent research in positional faithfulness and markedness. The constraints that flow from this proposal unify the account of segmental blocking effects on infixation and make disparate innovations such as the CONTROL component and crucial non-ranking of constraints obsolete. The analysis of the Chamorro variation data from the author's fieldwork has provided an important argument against gradient alignment and for quantized, categorical alignment. This result brings out an important methodological point for researchers wishing to advance phonological and morphological theory, namely, to look closely into subtle variation patterns within an individual language.

References

Anderson, S. R., 1992. A-Morphous Morphology. Cambridge University Press, Cambridge.
Anderson, S. R., 1993. Wackernagel's revenge: Clitics, morphology, and the syntax of second position. Language 69, 68-98.
Alderete, J., 1997. Dissimilation as local conjunction. In: Kusumoto, K. (Ed.), Proceedings of the North East Linguistic Society, 27. GLSA, Amherst, MA, pp. 17-32.
Anttila, A., 1997. Deriving variation from grammar. In: Hinskens, F., van Hout, R., Wetzels, W. L. (Eds.), Variation, Change, and Phonological Theory. John Benjamins, Amsterdam, pp. 35-68.
Archangeli, D., Moll, L., Ohno, K., 1998. Why not *NC̥. In: CLS 34: The main session. Chicago Linguistics Society, Chicago, pp. 1-26.
Bakovic, E. 2000. Harmony, Dominance and Control. Rutgers University Ph.D. dissertation. [Rutgers Optimality Archive #360]
Barsel, L. A., 1994. The verb morphology of Mori, Sulawesi. Pacific Linguistics Series B-111. Australian National University, Canberra.
Beckman, J., 1997. Positional faithfulness, positional neutralisation and Shona vowel harmony. Phonology 14, 1-46.
Blevins, J., 1999. Untangling Leti infixation. Oceanic Linguistics 38, 383-403.
Blevins, J., Garrett, A., 1998. The origins of consonant-vowel metathesis. Language 74, 508-556.
Bloomfield, L., 1917. Tagalog Texts with Grammatical Analysis. University of Illinois studies in language and literature. University of Illinois, Urbana.
Boersma, P., 1998. The OCP in functional phonology. Ms., University of Amsterdam. [Rutgers Optimality Archive #283]
Boersma, P., 2000. The OCP in the perception grammar. Ms., University of Amsterdam. [Rutgers Optimality Archive #435]
Carstairs-McCarthy, A., 1995. Phonological constraints on morphological rules. In: Spencer, A., Zwicky, A. M. (Eds.), The Handbook of Morphology. Blackwell, Oxford, pp. 144-148.
Casali, R. F., 1997. Vowel elision in hiatus contexts: Which vowel goes? Language 73, 493-533.
Chung, S., 1983. Transderivational relationships in Chamorro phonology. Language 59, 35-64.
Chung, S., 1998. The Design of Agreement. Evidence from Chamorro. The University of Chicago Press, Chicago.
Cooreman, A. M., 1987. Transitivity and Discourse Continuity in Chamorro Narratives. Mouton de Gruyter, Berlin.
Crowhurst, M., 1998. Um infixation and prefixation in Toba Batak. Language 74, 590-604.
Crowhurst, M., 2001. Coda conditions and um infixation in Toba Batak. Lingua 111, 561-590.
De Lacy, P. 2002. The Formal Expression of Markedness. University of Massachusetts, Amherst Ph.D. dissertation. [Rutgers Optimality Archive #542]

Gibson, J. D., 1992. Clause Union in Chamorro and in Universal Grammar. Garland, New York.
Goldsmith, J., 1976. Autosegmental Phonology. MIT Ph.D. dissertation. Distributed by Indiana University Linguistics Club, Bloomington.
Golston, C., 1996. Direct Optimality Theory: Representation as pure markedness. Language 72, 713-748.
Halle, M., 2001. Infixation versus onset metathesis in Tagalog, Chamorro, and Toba Batak. In: Kenstowicz, M., (Ed.): Ken Hale: A Life in Language. MIT Press, Cambridge, MA, pp. 153-168.
Halle, M., Vergnaud, J.-R., 1987. An Essay on Stress. MIT Press, Cambridge, MA.
Hayes, B. 1989. Compensatory lengthening in moraic phonology. Linguistic Inquiry 20, 253-306.
Hayes, B. 1995. Metrical Stress Theory. University of Chicago Press, Chicago.
Hayes, B., Abad, M., 1989. Reduplication and syllabification in Ilokano. Lingua 77, 331-374.
Horwood, G., 2002. Precedence faithfulness governs morpheme position. Ms., Rutgers University. [Rutgers Optimality Archive #527]
Hume, E., 1998. Metathesis in phonological theory: The case of Leti. Lingua 104, 147-186.
Hume, E., 1999. The role of perceptibility in consonant/consonant metathesis. In: Shahin, K. *et al.* (Eds.), pp. 293-307.
Hyman, L., 2001. The limits of phonetic determinism in phonology: *NC revisited. In: Hume, E., Johnson, K. (Eds.), The Role of Speech Perception in Phonology. Academic Press, San Diego pp. 141-185.
Ito, J., Mester, A., 1986. The phonology of voicing in Japanese. Linguistic Inquiry 17, 49-73.
Ito, J., Mester, A., 1998. Markedness and word structure: OCP effects in Japanese. Ms., UC Santa Cruz. [Rutgers Optimality Archive #255]
Kager, R., 1999. Optimality Theory. Cambridge University Press, Cambridge.
Kiparsky, P., 1993. Variable rules. Talk presented at Rutgers Optimality Workshop I, New Brunswick, New Jersey.
Kisseberth, C., 1970. On the functional unity of phonological rules. Linguistic Inquiry 1, 291-306.
Klein, T. B., 2000. Umlaut in Optimality Theory. A Comparative Analysis of German and Chamorro. Niemeyer, Tübingen.
Lapoliwa, H., 1981. A Generative Approach to the Phonology of Bahasa Indonesian. Pacific Linguistics Series D-34. Australian National University, Canberra.
Lombardi, L., 2001. Why place and voice are different: Constraint-specific alternations in Optimality Theory. In: Lombardi, L. (ed.), Segmental Phonology in Optimality Theory. Cambridge University Press, Cambridge, pp. 13-45.
Lubowicz, A., 2002. Derived environment effects in Optimality Theory. Lingua 112, 243-280.
McCarthy, J., 1986. OCP effects: Gemination and antigemination. Linguistic Inquiry 17, 207-264.
McCarthy, J., 2000. The prosody of phase in Rotuman. Natural Language and Linguistic Theory 18, 147-197.

McCarthy, J., 2002a. Against gradience. Ms., University of Massachusetts, Amherst. [Rutgers Optimality Archive #510]; also forthcoming in Phonology.

McCarthy, J., 2002b. A Thematic Guide to Optimality Theory. Cambridge University Press, Cambridge.

McCarthy, J., Prince, A., 1986. Prosodic morphology. Ms., University of Massachusetts, Amherst and Brandeis University.

McCarthy, J., Prince, A., 1993a. Generalized alignment. In: Booij, G., van Marle, J. (Eds.): Yearbook of Morphology 1993. Kluwer, Dordrecht, pp. 79-153.

McCarthy, J., Prince, A., 1993b. Prosodic morphology I: Constraint interaction and satisfaction. Ms., University of Massachusetts, Amherst and Rutgers University.

Myers, S., 1997. OCP effects in Optimality Theory. Natural Language and Linguistic Theory 15, 847-892.

Myers, S., Carleton, T., 1997. Tonal transfer in Chichewa. Phonology 13, 39-72.

Nababan, P.W.J., 1981. A Grammar of Toba Batak. Pacific Linguistics Series D-37. Australian National University, Canberra.

Orgun, C. O., Sprouse, R. L., 1999. From MPARSE to CONTROL: Deriving ungrammaticality. Phonology 16, 191-224.

Pater, J., 1999. Austronesian nasal substitution and other *NC̥ effects. In: Kager, R., van der Hulst, H., Zonnefeld, W. (Eds.): The Prosody-Morphology Interface. Cambridge University Press, Cambridge, pp. 310-343.

Percival, W. K., 1981. A Grammar of the Urbanized Toba Batak of Medan. Pacific Linguistics Series B-76. Australian National University, Canberra.

Piggott, G., 2002. Modes of affixation and exceptional phonology: Getting the constraints right. Talk presented at the 10th Manchester Phonology Meeting.

Pittman, R., 1966. Tagalog -um- and mag-: An interim report. Papers in Philippine Linguistics No. 1. Pacific Linguistics Series A-8. Australian National University, Canberra, pp. 9-20.

Plag, I., 1998. Morphological haplology in a constraint-based morpho-phonology. In: Kehrein, W., Wiese, R. (Eds.), Phonology and Morphology of the Germanic Languages. Niemeyer, Tübingen, pp. 199-215.

Prince, A., Smolensky, P., 1993. Optimality Theory: Constraint interaction in generative grammar. Ms., Rutgers University and University of Colorado, Boulder.

Schachter, P., Otanes, F., 1972. Tagalog Reference Grammar. University of California Press, Berkeley, CA.

Shahin, K., Blake, S., Kim, E-S. (Eds.), 1999. The Proceedings of the Seventeenth West Coast Conference on Formal Linguistics. CSLI Publications, Stanford, CA.

Smith, J., 2003. Towards a compositional treatment of positional constraints: The case of positional augmentation. In: Carpenter, A., Coetzee, A., de Lacy, P., (Eds.), Papers In Optimality Theory II. GLSA, Amherst, MA, pp. 337-370. [Rutgers Optimality Archive #550]

Smolensky, P., 1993. Harmony, markedness and phonological activity. Talk presented at Rutgers Optimality Workshop I, New Brunswick, New Jersey.

Stemberger, J. P., Handford Bernhardt, B., 1999. Contiguity, metathesis, and infixation. In: Shahin, K. *et al*. (Eds.), pp. 610-624.

Suzuki, K., 1998. A Typlological Investigation of Dissimilation. The University of Arizona Ph.D. dissertation. [Rutgers Optimality Archive #281]

Tesar, B., Smolensky, P., 1998. Learnability in Optimality Theory. Linguistic Inquiry 29, 229-268.

Topping, D., 1973. Chamorro Reference Grammar. University of Hawai'i Press, Honolulu.

Topping, D., 1980. Spoken Chamorro. University of Hawai'i Press, Honolulu.

Topping, D., Dungca, B., Ogo, P., 1975. Chamorro-English Dictionary. University of Hawai'i Press, Honolulu.

Tryon, D. T. (Ed.), 1995a. Comparative Austronesian Dictionary. An Introduction to Austronesian Studies. Part 1. Mouton de Gruyter, Berlin.

Tryon, D. T., 1995b. The Austronesian languages. In: Tryon, D. T. (Ed.), pp. 5-44.

Yip, M., 1988. The Obligatory Contour Principle and phonological rules: A loss of identity. Linguistic Inquiry 19, 65-100.

Yip, M., 1998. Identity avoidance in phonology and morphology. In: LaPointe, S. G., Brentari, D. K., Farrell, P. M. (Eds.), Morphology and its Relation to Phonology and Syntax. CSLI Publications, Stanford, CA, pp. 216-246.

Yu, A., 2002. Understanding infixes as infixes. Ms., University of Chicago. [Rutgers Optimality Archive #523]

Zorc, R. D., 1995. Tagalog. In: Tryon, D. T. (Ed.), pp. 335-341.