Maximal Words and the Maori Passive*

Paul de Lacy
University of Massachusetts, Amherst

1. Introduction

Many languages exhibit Minimal Word conditions: restrictions on the smallest possible Prosodic Word (PrWd) in a language (McCarthy & Prince 1986). In this paper, I propose that there are also restrictions on the maximum size of PrWds – i.e. 'maximal word' effects. Since very little has been written about maximal word conditions, one of the two aims of this paper is to show that such restrictions exist. The other aim is to show that upper bounds on PrWd size reduce to general constraints on prosodic structure.

The empirical focus of this paper is the Polynesian language Māori [máːoɾi], spoken in New Zealand. I show that PrWds in this language are allowed to contain at most one trochaic foot and no other footable sequences. This restriction allows bimoraic and trimoraic ωs – {(húka)}, {(táŋa)ta}, {ku(ɾíː)} – and four-mora PrWds with a medial foot – {ta(mái)ti} – or an initial uneven trochee – {(kóːɾe)ɾo}. However, it bans all other PrWd types, such as those with four light syllables or two light syllables and a heavy.[1]

Since there is pressure for PrWd and root edges to coincide, the restriction on PrWds translates into severe limitations on root and word size in Māori. I argue that the PrWd limits are also responsible for the passive suffix's many realizations; representative examples are given in the following table:

(1)

	Active (root)	*Passive*	*Gloss*
a.	hoɾo	hoɾo-a	'fall in fragments'
b.	kopou	kopou-a	'appoint'
c.	hoka	hoka-ia	'run out'
d.	mahue	mahue-tia	'put off'
e.	aɾihi	aɾihi-tia	'chop'

I propose that the short forms of the passive *-a* and *-ia* appear only when they can form an acceptable PrWd with the root. For example, {(hóɾo)a},

* I am indebted to my consultants Wharepapa Savage, Te Puhi Kapa, and a consultant who wishes to remain anonymous. My thanks to the following people for comments on various versions of this paper: Ross Clark, Catherine Kitto, Jane Grimshaw, John McCarthy, Alan Prince, the audience at the Rutgers Optimality Research Group, the audience at AFLA VIII, and the audience at Auckland University.

[1] Throughout this paper I use . to mark syllable boundaries, ( ) for foot boundaries, { } for ω boundaries, and – to indicate morpheme breaks.

MIT Working Papers in Linguistics 44, 20–39
Proceedings of AFLA VIII

{ko(póu)a}, and {ho(kái)a} are within the maximal word limit: they contain only one foot and no other footable sequences. In contrast, neither of the short forms could appear with a longer root in the same PrWd: both *{(ári)hi-a} and *{(ári)(hì-a)} are unacceptable since the former contains an unfooted footable sequence and the latter has too many feet. In such cases, the passive is forced to form a PrWd on its own, with attendant consonant epenthesis: {(ári)hi}{(tía)}.

I will show that maximal word limit is the primary factor in controlling the passive's realizations. I will argue that other forms of the passive – not shown in the data above – follow from general conditions on Māori phonology. I will also present a number of new observations about the data.

The theoretical aim of this paper is to show that the maximal word limit observed in Māori follow from general constraints on prosodic structure. This proposal ties in with McCarthy & Prince's (1994a) Generalized Template Theory – a reductionist approach to templatic restrictions.

Section 2 presents an outline of the theoretical approach and identifies the primary constraints used to effect the maximal word limits. The theory is applied to Māori's PrWd size restrictions in section 3. I discuss typological predictions of the theory in section 4, and conclusions in section 5.

2. Theory

The theory of maximal word effects proposed in this paper is a reductionist one: no special devices effect maximal word limits. Instead, I argue that maximal word limits are due to general prosodic constraints outranking faithfulness constraints. This approach is closely related to proposals made by McCarthy (1993) and McCarthy & Prince (1986 et seq.), called Generalized Template Theory (see McCarthy 2001 for comprehensive references).

McCarthy & Prince (1986) argue that Minimal Word effects are produced by general constraints on prosodic structure. Since Selkirk's (1984) Strict Layer Hypothesis requires every ω to contain foot and feet are minimally binary, it follows that every PrWd must contain a minimum of two moras (or syllables). McCarthy & Prince (1990, 1994a) extend this approach to templatic restrictions: they show that general prosodic constraints obviate the need for independent templatic devices.

The work that is most relevant for present concerns is found in McCarthy & Prince (1994b). The authors argue that there is no need for constraints that baldly state reduplicant form, such as 'RED=CVCV'. Instead, the emergent effect of prosodic constraints determines reduplicant shape. They illustrate with an analysis of Diyari reduplication, showing that the constraints ALLFTL and PARSE-σ ensure that reduplicants are maximally bimoraic in this language.

(2) PARSE-σ "Every syllable belongs to a foot"
ALLFTL"Every foot appears at the left edge of a PrWd." (M&P 1993b)

These constraints outrank MAX-BR, a constraint that requires reduplicants to contain their base's material. The reduplicant is underlined in the candidates

below; foot boundaries are not marked because they have no bearing on the result.[2]

(3)

	/RED+ŋandawalka/	ALLFT L	PARSE-σ	MAX-BR
	(a) {(ŋanda)(walka)}{ŋandawalka}	x x!		
	(b) {(ŋanda)wa}{ŋandawalka}		x!	x x x
☞	(c) {(ŋanda)}{ŋandawalka}			x x x x x

Despite the fact that (c)'s reduplicant is the least faithful copy of the base, it forms the most unmarked PrWd. The other candidates are ruled out because they are relatively more marked – (a) contains too many feet while (b) has a stray syllable.

Such size limits not only apply to reduplicants: Ito & Mester (1992, 1994) point out that Japanese loanword truncations place upper limits on their size, and Mester (1994) argues that cretic shortening in Latin may be seen as an imposition of a word size limit.

Since Optimality Theory allows free ranking permutation, analyses like the one above imply that prosodic constraints can limit the size of other morphological elements, not just reduplicants. If MAX-IO replaces MAX-BR, all morphemes will be maximally bimoraic or split between PrWds. In such a language, the root /takapa/ would surface as {taka}, {kapa}, or be split into two PrWds – the faithful {takapa} would fatally violate one of the foot-related constraints. In the following sections I argue that this prediction is borne out in Māori; its maximal word limit is produced by ranking constraints on footing about faithfulness constraints.

Apart from the constraints already mentioned, I will employ the following footing constraints:

(4) LAPSE_{FT} "Adjacent unstressed moras must be separated by a foot boundary." (Green & Kenstowicz 1995, Prince 1983, Selkirk 1984).

*FT- "Incur a violation for every non-head foot."

LAPSE_{FT} is violated when a footable sequence is not parsed into feet. *FT- bans every foot except for the head foot, effectively preventing PrWds with more than one foot. I will show that the constraints in (4) together favour small PrWds – ones that contain a single foot and no footable sequences.

3. The Māori Passive

In this section I show that there is an active restriction on the maximum size of PrWds in the Polynesian language Māori. While the main focus of this section

[2] See McCarthy & Prince for details of the analysis and data. McCarthy and Prince's example *ŋandawalka* is deduced from the source material.

is the passive suffix, I will begin by presenting some background to Māori phonology, starting with the phonemes:

(5) *Consonants*

	p	t	k	
	f / ɸ			h
	m	n	ŋ	
		ɾ		
	w			

(6) *Vowels*

i	i:	u	u:
e	e:	o	o:
	a	a:	

Syllable structure is (C)V(V). Syllable rimes may contain either a long vowel or a diphthong. In all diphthongs the second vowel is equally or less sonorous than the first: i.e. [ai ae ao au eo ei oe ou iu ui].[3] Other vowel sequences (e.g. [oa io]) form separate syllables.

Bimoraic syllables contain either a long vowel or a diphthong.[4] For present purposes, it is enough to say that stress falls on a bimoraic syllable, otherwise the initial. Stress will be marked in all data presented below.[5]

(7) [táma] 'boy'
[táŋata] 'man'
[maɾáe] 'meeting area'
[kuɾíː] 'dog'

Content words – nouns, verbs, and adjectives – are minimally bimoraic. In foot terms, Māori employs trochaic feet and aims to have them at the left edge of the PrWd if possible (i.e. {(táŋa)ta}, *{ta(ŋáta)}.

Every bimoraic root or affix is contained inside its own PrWd. Standard diagnostics for PrWd boundaries are syllabification and stress. For example, the compound /taka afe/ 'circuitous' ('go'+ 'encircle') is stressed as [tákaáfe], not *[taká:fe], indicating that there is a PrWd break – and therefore a syllable break – between the medial [a]'s: [{táka}{áfe}]. Similarly, the prefix *taki-* {numeral modifier} forms its own PrWd: [{táki}{íwa}] 'by nine', *[{takíːwa}].

Monomoraic affixes appear inside the PrWd of their root. For example, the monomoraic prefixal reduplicant clearly falls inside the same

---

[3] The status of [iu] and [ui] as diphthongs varies across dialects; in any case, they are rare and do not prove to be significant in the following discussion.

[4] Evidence that diphthongs are bimoraic comes from stress (see below) and Minimal Word restrictions.

[5] For a fuller description of Maori stress, see Biggs (1961), Hohepa (1967) and Bauer (1993). Some dialects do not stress word-final diphthongs (i.e. [márae]), but non-final diphthongs are stressed in every dialect: e.g. [tamáiti] 'child', *[támaiti]. This variation will have no bearing on the proposals below, so I will not discuss it further.

PrWd as its root since it bears the stress: e.g. {híhiŋa} ‘fall in a large amount’ (<híŋa ‘fall’).

The affixes that are of most interest in this paper are those that have both bimoraic and monomoraic realizations – i.e. the passive and gerund. Below I will show such realizations to be conditioned by the maximal word limit. The first step, though, is to identify the size restrictions on Māori PrWds.

3.1 Maximal Words

Māori roots may contain two, three, or four moras. Four mora roots only come in two types: with an initial heavy syllable followed by two light syllables (HLL) and with a medial heavy syllable followed by two light syllables (LHL).

(8) (i) Bimoraic Roots

[tá.ma] ‘boy’	[héu] ‘eaves’
[ú.a] ‘be rained upon’	[úː] ‘bite, gnaw’
[kái] ‘food’	[páː] ‘fortified village’

(ii) Trimoraic Roots

[ká.ra.ŋa] ‘sing’	[káu.ɾi] ‘type of tree’
[á.mi.o] ‘roam’	[te.káu] ‘ten’
[áː.mi] ‘gather’	[ku.ɾíː] ‘dog’

(iii) Four Mora Roots

(i) HLL	(ii) LHL
[máː.o.ɾi] ‘Maori, normal’	[ta.mái.ti] ‘child’
[kóː.ɾe.ɾo] ‘tie’	[ta.ɾái.wa] ‘driver’
[páː.ke.ha] ‘Caucasian’	[ma.náː.ki] ‘show kindness’

Many of the LHL roots are historically derived from bimorphemic forms (e.g. *tamaiti* ‘child’ < *tama* ‘boy’ + *iti* ‘small’). Nevertheless, they are now single roots, with meanings that are often unrelated to their (historically) component morphemes. It is clear that four mora roots are a recent addition to Māori, and it is likely that they were once prohibited. In fact, I will show that this earlier ban persists in modern Māori – while four mora PrWds are tolerated, they are avoided when possible.[6]

I propose that the limits on root shape are due to restrictions on the size of PrWds: PrWds may contain only one foot and no footable sequences. This requirement affects roots because each root is required to be contained inside a single PrWd.

As mentioned above, feet are trochaic; they may consist of one heavy syllable ($\acute{\sigma}_{\mu\mu}$), two light syllables ($\acute{\sigma}_{\mu}\sigma_{\mu}$), or a heavy-light sequence ($\acute{\sigma}_{\mu\mu}\sigma_{\mu}$). To require only one foot per PrWd, I employ the constraints introduced in section 2: *FT- and LAPSE_{FT}. Either one or both of these constraints are violated by

[6] Many names are exceptions to the generalizations made above. These are either morphologically complex or onomatopoeic. For example, the name {tú:i:} ‘parson bird’ consists of two heavy syllables, and derives from the sound of its call. This fact is unsurprising: even English names exhibit prosodic structures not found in other words (Liberman & Prince 1977).

PrWds that contain more than one foot or have unfooted sequences of moras. For example, the four-mora PrWd {karaŋata} cannot help but contain a non-head foot {(kára)(ŋàta)} – violating *FT- – or an unfooted sequence {(kára)ŋata}, so violating LAPSE_{Ft}.[7]

The two foot-related constraints conflict with the requirement that root material be preserved: MAX_{Root} (McCarthy & Prince 1994b). With this ranking, roots are forced to truncate if they get too large.

(9) LAPSE_{FT}, *FT- » MAX_{Root}

/karaŋata/	*FT-	LAPSE_{FT}	MAX_{Root}
(a) {(kára)ŋata}		x!	
(b) {(kára)(ŋàta)}	x!		
☞ (c) {(kára)ŋa}			x

The faithful candidates (a) and (b) fail because they violate one of the footing constraints. Candidate (a) manages to avoid violating *FT- by having one foot, but in doing so it ends up with two unfooted syllables, fatally violating LAPSE_{FT}. Candidate (b) satisfies LAPSE_{FT}, but can do so only by fatally violating *FT-. In short, four-mora roots of this type will inevitably violate a footing constraint, dooming them to loser status.

The same ranking rules out almost all other roots with four or more moras. The two exceptions are PrWds with a medial foot, as in {ta(mái)ti}, and those with an initial uneven trochee {(kó:re)ro}. Neither of these forms violate *FT- or LAPSE_{FT} since both contain a single foot and no unfooted footable sequences. The following tableau illustrated this situation with *ko:rero*. As shown the winning form (a) must contain an uneven trochee; those with even trochees – (b), (c) – violate one of the foot constraints.

(10)

/ko:rero/	*FT-	LAPSE_{FT}	MAX_{Root}
☞ (a) {(ko:re)ro}			
(b) {(kó:)rero}		x!	
(c) {(kó:)(rèro)}	x!		
(d) {(kó:)re}			x x!

In the following section I show that the PrWd size restrictions are not just a historical accident, but are active in the phonology of Māori.

3.2 The Passive: Introduction

The Māori passive has received a great deal of attention (Williams 1971[1844]; Biggs 1961; Hohepa 1967; Hale 1968, 1973, 1991; Kiparsky 1971; Kenstowicz & Kisseberth 1979:171-174; McCarthy 1981; Sanders 1990, 1991; Harlow 1991; Bauer 1993; Blevins 1994; Kibre 1998). Most previous work has focused on the issue of learnability; the key data for this issue is the following:

[7] The candidate {ka(ráŋa)ta} is ruled out by constraints requiring initial stress, also responsible for initial stress in {(táŋa)ta}, *{ta(ŋáta)}.

(11)

Active	*Passive*	*Gloss*
afi	afitia	embrace
hopu	hopukia	catch
aɾu	aɾumia	follow
mau	mauɾia	carry
kite	kitea	see, find
hoka	hokaia	run out
tahu	tahuna	light

Hale (1968) pointed out that the data allow for two types of analysis. A purely morphological analysis would have a large number of passive allomorphs: *-kia, -mia, -ria, -tia*, and so on; each root would specify which passive allomorph it took. In contrast, a phonological analysis would have consonants be part of the root: i.e. *hopu* is /hopuk/, and the passive is /ia/. The final consonant would delete when the root appears on its own because codas are banned (i.e. /hopuk/ → [hopu]), but the consonant can appear when it is an onset – i.e. with the passive: /hopuk+ia/ → [hopukia].

In his analysis, McCarthy (1981) pointed out that the phonological analysis is not as straightforward as one might think. It does not account for the realization *-a* in *kitea*, nor does it account for its realization as *ia* in *hokaia* and *-na* in *tahuna*. McCarthy showed that several extra phonological rules are needed to deal with these forms, but pointed out that only the passive provides evidence for their existence. With this extra complexity, McCarthy argued, the learner faces a far from trivial challenge in constructing a phonological account of the passive.

To complicate matters, there is also a third option – intermediate between the Phonological and Morphological approach. In this theory, roots contain final consonants, but the passive does consist of several suppletive allomorphs: i.e. *-a*, *-ia*, etc. Allomorph choice is determined by phonological conditions.

While my primary aim is not to discuss learnability, the analysis I present in the following pages does have bearing on this issue. I aim to show that the passive's realization is determined by the PrWd size limit. Certainly, some other conditions do influence the output form of the passive, but I show that these reduce to general prosodic restrictions that are visibly active in other processes in Māori. In short, I will argue that the Phonological approach is viable and does not require any devices that only find support in the passive's alternations.

3.3 Generalizations

In this section I present a brief statement of the passive's various realizations. The following sections will focus on separate parts of the description. I must note that the following generalizations made about the data do not entirely agree with previous descriptions. The generalizations presented below were based on

an exhaustive search of two Māori dictionaries – Williams (1971) and Ngata (1993). I also checked the forms with my consultants.[8]

The following table summarizes the generalizations. One example form is given for each root shape; further examples are given in the following sections when appropriate. I give the number of examples that support each generalization after the gloss.

Since the prosodic form of the output proves significant , I have marked stress, syllable and PrWd boundaries in the examples. Evidence that forms such as {ínu}{mía} and {tápuhi}{tía} form separate PrWds comes from stress placement and intonation; specifically, the H* of the declarative tune falls on the rightmost ω's stressed syllable, and the pitch rise occurs over the passive suffix in just these words.

(12)

Root shape	*Passive*	*Example*	*Gloss*	*Num.*
C-final				
/…an/	ina	{e.pái.na}	throw	22
other n-final	a	{táo.na}	cook in oven	54
other C-final	ia	{í.nu}{mí.a} {ko.ha.ɾa}{kí.a}	drink split open	137
Bimoraic V-final				
(C)i:	a	{píː.a}	bathe	6
other (C)V:	ia	{púː.ia}	bundle	12
(C)V(C)a	ia	{hi.kái.a}	plant	41
other (C)V(C)V	a	{hó.ɾo.a}	fall down	434
Larger V-final				
HL	a	{táu.te.a}	consider	13
LH	a	{ko.póu.a}	appoint	13
LLL LHL, HLL	tia	{tá.pu.hi}{tí.a} {kóː.ɾe.ɾo}{tí.a}	sort out talk, say	81

One other realization of the passive is found with 26 roots: these lengthen their initial syllable as well as suffixing a passive form (e.g. /kume/ → [kuːmea] 'be angry, fight'). I have discussed this pattern elsewhere, so I will leave it aside here (de Lacy 1999; also see Harlow 1991).

One important issue relates to the claim that trimoraic and four mora vowel-final forms take *-tia*.[9] Evidence for this claim comes from two sources. One is that trimoraic and longer loanwords take *-tia* in the passive (Hale 1968, Blevins 1994:41). The other is that no trimoraic form takes *-ia* or *-a*.[10] Since

[8] My consultants were unfamiliar with a number of the forms from Williams. In those cases, I asked them to comment on the naturalness of the passive termination.

[9] Blevins (1994), citing Ray Harlow, reports that some dialects have *-hia* and some *-ŋia* as the default passive form.

[10] There are about 21 apparent exceptions to this claim: e.g. *kohuki~kohukia* 'impel', *tapahi~tapahia* 'stamp (foot), disobey'. However, Williams points out that most of these forms end in an fossilized suffix, *-i*. I suggest that this suffix is still recognized as distinct from the root, so *tapahi* is underlyingly /tapah-i/. In passivisation, the fossilized suffix is eliminated and the passive concatenates as expected: /tapah-ia/ → {tapa}{hia}.

the majority of trimoraic roots take *-tia*, it is therefore difficult to imagine that V-final trimoraic forms take anything but this form of the passive.

The following analysis takes the underlying form of the passive to be /ia/; I will provide arguments for this proposal in section 3.7.

3.4 Maximal words and the passive

I propose that the realizations of the passive are primarily controlled by the maximal word limit. If it is possible to create the least marked PrWd – a trimoraic one – the passive will truncate from /ia/ to [a] to do so: /horo+ia/ → {hóroa}. Failing that, the aim will be to create an admissible four-mora PrWd: e.g. /kopou+ia/ → {ko(póu)a}. When truncation cannot produce the right result, the passive is placed in its own PrWd, with attendant epenthesis: /mahue+ia/ → {máhue}{tía}.

I will start by showing that the passive truncates when necessary, and that the PrWd size limit determines when this truncation takes place. Evidence comes from trimoraic roots that contain a heavy syllable; such roots take *-a* in the passive:

(13) *Trimoraic roots with a heavy syllable*

(i) $\sigma_{\mu}\sigma_{\mu\mu}$ roots

horoi	{ho(ɾói)a}	'clean'
kopou	{ko(póu)a}	'appoint'
tapae	{ta(páe)a}	'present'
tapi:	{ta(pí:)a}	'mend'

(ii) $\sigma_{\mu\mu}\sigma_{\mu}$ roots

keue	{(kéue)a}	'move'
haere	{(háere)a}	'go over/for'
a:mi	{(á:mi)a}	'gather'
hu:hi	{(hú:hi)a}	'cover'

The reason that these roots take *-a* is because the more faithful alternative [ia] violates the maximal word limit: e.g. *{ko(póu)ia}, *{(háere)ia}. So, the passive will truncate if doing so is the only way to form an admissible PrWd:

(14) Truncation under duress

	/kopou+ia/	*FT-	$LAPSE_{FT}$	MAX
☞	(a) {ko(póu)a}			x
	(b) {ko(póu)ia}		x!	
	(c) {ko(póu)(ia)}	x!		

The alternative realization – *tia* – is ruled out because it contains an epenthetic consonant [t]. The constraint against consonant epenthesis – DEP-C – outranks MAX, so banning the form *{ko(póu)}{(tía)}. The *-tia* realization and epenthesis will be discussed in more detail in section 3.6.

Support for this approach comes from the gerund: /tapahi + aŋa/ appears as [tapahaŋa], not *[tapahiŋa] (cf. /hoki+aŋa/ → [hokiŋa] 'return').

3.5 The emergent maximal word

The maximal word limit that affects affix form is more stringent than the one imposed on roots. While four mora roots are tolerated in Māori, they are avoided in affixed forms. Specifically, the passive will truncate to form a three mora PrWd in order to avoid a four mora one. Such truncation takes place with bimoraic roots that end in a non-low vowel.

(15) *Bimoraic V^{low} roots*

hoɾi	{hóɾia}	'cut'
huke	{húkea}	'excavate'
moe	{móea}	'marry, beget'
miɾo	{míɾoa}	'twist'
ehu	{éhua}	'bail'
hau	{háua}	'strike'

The alternative to the trimoraic output forms above is a four mora PrWd: e.g. *{hu(ké-i)a}. The avoidance of such PrWds is due to the constraint ALLFTL (see (2)). With ALLFtL outranking MAX, the passive will truncate to avoid four-mora PrWds with the form $\{\sigma_\mu(\sigma_{\mu\mu})\sigma_\mu\}$.

(16) Avoidance of {LHL}

	/huke/	ALLFTL	MAX
☞	(a) {(húke)a}		x
	(b) {hu(kéi)a}	x!	

Importantly, the constraint MAX_{Root} outranks ALLFTL; since MAX_{Root} specifically preserves root material, it keeps roots like *tamáiti* from being eliminated:

(17) LHL roots

	/tamaiti/	MAX_{Root}	ALLFTL
☞	(a) {ta(mái)ti}		x
	(b) {(táma)}	x x x!	

With the ranking || MAX_{Root} » ALLFTL » MAX ||, the ban on four mora PrWds only emerges in affixation. In other words, the most desirable PrWd is bi- or trimoraic, with four mora PrWds only possible under duress.

3.6 The last resort

So far I have argued that the maximal word limit forces the passive to truncate when necessary. However, there is one situation where truncation does not achieve the right result. With trimoraic roots that consist entirely of light syllables, neither *-a* nor *-ia* will form an acceptable PrWd: e.g. /mahue+ia/ → *{mahu(é-i)a}, *{(máhu)ea}. In this situation, there is only one remaining option: the passive must appear inside its own PrWd, resulting in {(máhu)e}{(tía)}.

The reader will no doubt have noticed that a [t] appears in the output form; the passive is not *{(máhu)e}{(ía)}. The appearance of the [t] relates to an independent restriction in Māori: if an affix starts a PrWd, that PrWd must begin with a consonant. The effect of this restriction is seen in two facts: (i) all prefixes begin with consonants and (ii) prefixal cannot reduplicate vowel-initial words (Keegan 1996:36).[11] So, *{máhue}{ía} is ruled out by the affix restriction; the only way for an affix to appear in its own PrWd is for a consonant to be epenthesized, hence {máhue}{tía}.

Consonant epenthesis violates the constraint DEP-C. Since avoidance of a maximal word violation is clearly preferable to avoiding epenthesis, *FT- and LAPSE_{FT} must both outrank DEP-C:

(18) Epenthesis as a last resort

	/mahue+ia/	*FT-	LAPSE_{FT}	DEP-C
☞	(a) {(máhu)e}{(tía)}			x
	(b) {mahu(éi)a}		x!	
	(c) {(máhu)(èa)}	x!		

In short, placing the passive in its own PrWd with attendant epenthesis is a last resort; it only takes place when truncation cannot satisfy the maximal word limit.

The fact that /kopou+ia/ surfaces as {kopóua} and not *{kopóu}{tía} shows that DEP-C outranks MAX. With this ranking, even admissible four mora PrWds are more harmonic than epenthesis.

(19) $\{\sigma_{\mu}\sigma_{\mu\mu}\sigma_{\mu}\}$ is preferable to epenthesis

	/kopou+ia/	DEP-C	MAX
☞	(a) {ko(póu)a}		x
	(b) {ko(póu)}{(tía)}	x!	

While this concludes the discussion of the primary effects of the maximal word limit, there are still several realizations of the passive that require further comment. These include the forms found with C-final roots, [a]-final roots, and long vowels.

3.7 Epenthesis and the last resort

Epenthesis forms an important part of the analysis presented above: it is the constraint against epenthesis – DEP-C – that renders the *-tia* realization least harmonic. So, the form *{kopóu}{tía} is not rejected because it has two PrWds, but rather because it has an epenthetic consonant. In fact, epenthesis is the only reason that the *-tia* realization is ruled out in this situation; apart from

[11] For the sake of brevity, I will not provide an in-depth discussion of the formal mechanism that produces this restriction. As a brief explanation, this restriction is no doubt due to the emergent effect of ONSET, outranked by (i) a condition on root contiguity – ruling out root-medial epenthesis – and (ii) PrWd-root alignment, ruling out epenthesis at the left edge of roots.

epenthesis, output forms with *-tia* satisfy all the other constraints, obeying the maximal word limit.

This approach makes an important prediction: if *-ia* can appear in its own PrWd without epenthesis, no constraint will prevent it from surfacing faithfully. This prediction is borne out in C-final roots: these all end up with *-ia* in a separate PrWd:

(20) *C-Final Roots*

Underlying Form	*Passive*	*Gloss*
inum	{ínu}{mía}	'drink'
monok	{móno}{kía}	'prepare food'
fauf	{fáu}{fía}	'tie'
nekeh	{néke}{hía}	'move'
inum	{ínu}{mía}	'drink'
kuːŋ	{kúː}{ŋía}	'nip'
apuɾ	{ápu}{ɾía}	'heap upon'
kohaɾak	{kóhaɾa}{kía}	'split open'
manakoh	{mánako}{hía}	'accept'
matakuɾ	{mátaku}{ɾía}	'be feared'

The competing form is one with a single PrWd and the realization *-a* (e.g. *{(móno)ka}). This form is ruled out by MAX since the passive's [i] is deleted. In comparison, forms with the passive in a separate PrWd do not violate any of the constraints identified so far: {(móno)}{(kía)} does not violate LAPSE_{FT} or *FT-, and does not violate DEP-C. The latter fact is crucial – it shows that the {tia} realization is not avoided because it appears in a separate PrWd, but because it has an epenthetic consonant.

The form {(móno)}{(kía)} gives some insight into requirements on root-PrWd containment. Some constraint must require roots to appear inside a single PrWd – I dub this WRAP(Root, PrWd), after Truckenbrodt 1995. This constraint requires every vocalic element of a root to be contained inside the same PrWd, preventing roots from forming two separate PrWds to satisfy the maximal word limit: i.e. /kaɾaŋata/ → *{(káɾa)}{(ŋáta)}. WRAP(Root, PrWd) must at least outrank MAX_{Root}.

(21) WRAP » MAX_{Root}

/kaɾaŋata/	WRAP(Root, PrWd)	MAX_{Root}
☞ (a) {(kára)ŋa}		x
(b) {(kára)}{(ŋáta)}	x!	

In the present analysis, WRAP(Root, PrWd) must only apply to the vocalic members of roots, otherwise [{mono}{kia}] would be banned.

The C-final forms are important in determining the input's form. I have assumed that the passive is underlyingly /ia/. The alternative – that it is /a/ – incorrectly predicts that /inum+a/ should surface as *{(ínu)ma}. This form does not violate the maximal word conditions, so it is difficult to see what would rule it out in favour of {ínu}{mía}, a form that (at least) violates DEP-V.

3.8 Violability

Throughout this section I have argued that the maximal word limit is imposed by constraints on footing, and not some independent templatic requirement. If the limit is truly imposed by constraints, though, one would expect these to be violable. The violability of the maximal word limit is shown in three realizations of the passive. In these cases, the form of the passive that is most harmonic in terms of the maximal word limit is ruled out by higher ranking constraints.

One case involves the OCP (Goldsmith 1976). The OCP bans adjacent identical elements within the same PrWd in Māori, ruling out [V_i:V_i] sequences (e.g. *[a:a]). The OCP also influences the passive's realization with bimoraic long-vowel roots. Although most roots of this shape take the passive form [ia], those with an [i:] take -*a*.

(22) *Roots with long vowels*

(i) [(C){e,a,o,u}:]		
ko:	{kó:.i.a}	'dig with k⃝'
a:	{á:.i.a}	'drive, urge'
pu:	{pú:.i.a}	'make into bundle'
(ii) [(C)i:]		
hi:	{hí:.a}	'raise'
ki:	{kí:.a}	'mention'
pi:	{pí:.a}	'bathe'

The reason that non-[i:] long vowels take -*ia* follows from the ranking presented so far. The candidate {(kó:.i).a} does not exceed the maximal word limit, nor does it violate ALLFTL. Its competitor {(kó:)a} fatally violates MAX, and {(kó:)}{(tía)} violates DEP-C:

(23) The HLL Output

	/ko:+ia/	DEP-C	MAX
☞	(a) {(kó:i)a}		
	(b) {(kó:)a}		x!
	(c) {(kó:)}{(tía)}	x!	

However, -*ia* is not the most harmonic form for [(C)i:] roots: i.e. *{(pí:i)a}. The reason is that the output clearly violates the OCP.

(24)

	/pi:+ia/	OCP	MAX
	(a) {(pí:i)a}	x!	
☞	(b) {(pí:)a}		x

The OCP also affects roots that end in [a]. While most bimoraic roots take -*a* as their passive form (see (15)), those that end in [a] take -*ia*.

(25) [a]-final bimoraic roots

hika	{hikáia}	'plant'
pona	{ponáia}	'tie'
tia	{tiáia}	'paddle vigorously'

The alternative would have adjacent [a]'s: *[hi.ka.a]. Again, the OCP can be used to rule this form out:

(26)

	/hika+ia/	OCP	ALLFTL
☞	(a) {hi(kái)a}		x
	(b) {(híka)a}	x!	

One other candidate deserves further comment: the form *{hi(ká:)}, with a long vowel, does not violate the OCP and so should be the winner.[12] I suggest that such forms are ruled out by a constraint that requires the passive to have some unique output exponent. Such a constraint is proposed by McCarthy & Prince (1995) – MORPHDIS prevents all the passive's melodic elements from coalescing, as is the case in **hiká:*.

In any case, a similar constraint is needed to explain why the passive does not delete entirely with trimoraic forms. Given the input /mahue+ia/, the output form {máhue} – where the passive has deleted entirely – wins out over {máhue}{tía} since the latter violates DEP-C while the former only violates the lower-ranked MAX. Some constraint that requires the passive to have an output exponent must therefore outrank DEP-C.

Another case involves /n/-final roots. Such roots do not behave like other C-final roots; roots that end in /an/ metathesize the /n/ with the passive's [i], while other /n/-final roots take *-a*, not *-ia*.

(27) /an/-final roots [22 forms]

/epan/	{e(pái)na}	'throw'
/huan/	{hu(ái)na}	'determine'
/weɾan/	{we(ɾái)na}	'burn'

/n/-final roots [54 forms]

/akon/	{(áko)na}	'learn'
/takin/	{(táki)na}	'stick in'
/wao/	{(wáo)na}	'part combatants'

To deal with [n]'s behavior, I suggest that there is a constraint against [ni] sequences, which I will call **ni* here. Admittedly, this constraint is ad hoc; I expect that the real reason behind avoidance of [ni] can be related to the plethora of cooccurrence conditions found in Māori (Kawasaki 1990, de Lacy

[12] In fact, I found 12 [a]-final forms that take *-a* (e.g. ɾaŋa~ɾaŋa: 'charge', taɾa~taɾa: 'gossip'). However, all the forms are from William's dictionary alone, and he offers alternative forms with *-ia* or *-ina* for six of them. For the seven forms for which Ngata provided the passive, all were recorded as taking *-ia* or *-ina*, not *-a*. So, the *-a* termination for [a]-final roots is very marginal.

1998). However, developing this line of research would go beyond the scope of this paper. For present purposes, it is enough that some constraint that bans [ni] sequences at least outranks MAX:

(28)

	/takon+ia/	*ni	ALLFTL	MAX
	(a) {(táko)}{(nía)	x!		
☞	(b) {(táko)na}			x
	(c) {ta(kói)na)}		x!	

As for the metathesized candidate, it is as if the [n] did not exist: the OCP sees right through it, effectively banning [a(n)a] sequences:

(29)

	/epan+ia/	*ni	OCP	ALLFTL
	(a) {(épa)}{(nía)	x!		
☞	(b) {e(pái)na}			x
	(c) {(épa)na)}		x!	

Of course, more must be said about these roots. For example, why does only /n/ undergo metathesis? An adequate response would need to invoke separate constraints on order preservation (i.e. LINEARITY), a step that is unfortunately beyond the scope of the present work.

Despite the remaining questions about the /n/-final forms, it is clear that the maximal word condition still determines the outcome of passivization of these forms. In fact, the *ni constraint acts much like the OCP: it bans the most harmonic form in terms of the maximal word limit. Even so, the footing constraints are still crucial in picking the winning form.

3.9 Alternatives and Learnability

In the introduction to this section, I suggested that the maximal word effects throws some light on the learnability problem. McCarthy's (1981) concern with a phonological approach was that it required devices that only found justification in the passive's realizations. In the preceding sections, I showed that the primary conditioning effect on the passive is a restriction whose influence is seen in almost every part of Māori phonology and morphology – the maximal word limit. Evidence for other conditioning effects – such as the OCP and the requirement on C-initial PrWds – is found not only in passive's alternations, but in other aspects of the phonological system. In short, there is very little the learner would have to postulate working from the passive alone.

With a viable Phonological approach, there is little to recommend a purely morphological analysis. As Hale (1968 et seq.) pointed out, a morphological analysis must treat as coincidental the fact that the gerund -(a)ŋa surfaces with exactly the same consonants as found in the passive: e.g. inu~inumia~inumaŋa; hopu~hopukia~hopukaŋa; waru~waruhia~waruhaŋa. Hale (1968, 1991) also presents a number of arguments for *-tia* as the default realization of the passive. For example, *-tia* is used with verbalized nouns;

Blevins (1994) adds the observation that *-tia* is used for trimoraic loanwords. Again, these facts do not obviously follow from a Morphological approach.

I also consider a 'compromise' approach problematic. Such an approach adopts the Phonological Analysis' idea that roots can end in consonants, but maintains that the passive has several separate lexically listed allomorphs, at least *-ia*, *-tia*. In this approach, the choice of the allomorph is determined by constraints; the allomorph that appears in the most harmonic form surfaces successfully.

There are two reasons why a compromise approach is undesirable; both relate to the [t] in *-tia*. In the present theory, the [t] in the passive's *-tia* is epenthetic. So, one would expect [t] to show up as the epenthetic consonant with the gerund as well. It does: loanwords take *-tia* in the passive and *-taŋa* in the gerund. This fact does not obviously follow from an approach where /tia/ and /taŋa/ are underlying forms: it does not explain why both have the consonant /t/, as opposed to some other more marked consonant, or even why both have the same initial consonant.

The other reason relates to the implementation of such a theory. The idea behind the compromise approach is that forms with both the *-(i)a* and *-tia* allomorphs compete and whichever is in the most harmonic form wins. Despite its reliance on constraints in determining the passive's realization, the proposals I have made above cannot be easily adapted to work for the compromise theory. As I argued in section 3.7, the reason that the *tia* realization is avoided is because of its epenthetic consonant. Since the [t] is not epenthetic in the compromise theory, some other reason is needed to explain why the *tia* allomorph is avoided.

I suggest that finding a prosodic reason for the avoidance of *tia* will be extremely difficult. The obvious approach is that *tia* is avoided because it requires its own PrWd. With a constraint against proliferation of PrWds, the *tia* allomorph will only be chosen when all candidates with a single PrWd exceed the maximal word limit: i.e. *{mahuea}, *{mahueia}, ✓{mahue}{tia}. However, this approach faces problems in dealing with C-final roots such as /inum/. The most harmonic form with such roots is {ínu}{m-ía}, with the passive in a separate PrWd. Crucially, the candidate *{ínum-a} is ruled out. But what rules it out? If the aim is to avoid proliferation of PrWds, *{ínuma} is incorrectly predicted to be more harmonic than {ínu}{mía}.

So, placing the passive inside a separate PrWd is not a problem – in fact, it is the most harmonic option for C-final roots. The reason that *tia* is avoided, then, must reduce to something else. In the present theory it is because the [t] is epenthetic, but this is an option not open to the compromise theory.

To summarize, I maintain that only a fully phonological analysis provides a workable solution to the passive's alternations. The morphological analysis fails to account for the consistency of consonants found in the passive's and gerund's alternations. I do not see that the compromise approach can provide a workable solution either; in this approach the occurrence of [t] in both the default passive and gerund forms is coincidental and some non-obvious reason must be invoked to explain why the *-tia* realization is avoided.

4. Typology

In this section I will not enumerate all possible maximal word limits since such a discussion depends on details of the constraints assumed. Instead, I will discuss some general properties of prosodic constraints that have a bearing on maximal words. One is minimality: markedness constraints generally militate against proliferation of structure. The other is binarity: prosodic constraints often promote binary structures, as for feet (e.g. FTBIN) and often in stressed syllables.

Apart from constraints requiring binarity – discussed below – markedness constraints prefer less structure over more. For example, the constraint ALLFTL can only be satisfied if a PrWd contains a single foot. Since a PrWd size limit comes about when a prosodic constraint outranks a faithfulness constraint and prosodic constraints promote the minimum structure, size restrictions must therefore promote some prosodically minimal structure.

For example, the Māori maximal word limit requires the minimum number of feet in a PrWd. Similarly, it minimizes the number of unfooted sequences: none are allowed. Other maximal word limits support the idea that maximal word limits must promote prosodic minimality. For example, Ura – spoken in Vanuatu – allows only two and three-mora roots (Crowley 1999:116-7). Ura's restriction can also be seen as minimizing the number of feet in a word, but to a slightly stricter extent than in Māori. The maximal word limits that emerge in many reduplications impose an even stronger minimality requirement: root reduplicants are maximally bimoraic, minimizing both feet and stray syllables.

The minimization effect of prosodic constraint rules out many type of PrWd size restriction. For example, no maximal word limit can allow PrWds with three feet but prohibit larger ones since there is no imaginable sense in which a tri-podal PrWd is prosodically minimal or unmarked.

On the other hand, some prosodic constraints promote increased structure, though in a very restricted way. FTBIN, for example, favours bimoraic feet over monomoraic ones. Similarly, ONSET prefers a bisegmental $[CV]_\sigma$ syllable over a monosegmental $[V]_\sigma$ one.

Binarity requirements could also produce maximal word effects. For example, Ito & Mester (1992, 1994) argue that there is a size restriction on the output of Japanese loanword truncation. The condition allows binary branching PrWds, but no larger; they can contain a single foot and an unfooted syllable or two feet, but no other structure. In terms of binarity, this structure is unmarked and so can potentially be an upper bound on PrWds. Similarly, Selkirk & Tateishi (1988) have argued that MaPs in Japanese are maximally binary.

So, there are two general properties of prosodic constraints that can affect maximal word limits: minimality and binarity. Together, these predict that maximal word limits will place a rather strict upper bound on PrWd size. In effect, maximal word conditions will require PrWds to be prosodically minimal or maximally binary.

The final typological issue I will address relates to how maximal word limits emerge in a language. Māori is an interesting case because it exhibits three different maximal word conditions. The least strict one relates to roots: they are maximally trimoraic, with the exception of four mora roots of the shape

LHL or HLL (e.g. *tamáiti, kó:rero*). A more restrictive condition emerges in allomorphy: here four mora PrWds are undesirable, as shown by the avoidance of {ho(ró-i)a} in favour of the trimoraic {(hóro)-a}. An even more restrictive version emerges in reduplication: the maximum size of a reduplicant is bimoraic: e.g. pirapirau, *piraupirau. The difference between these restrictions comes down to faithfulness. Root-faithfulness is highest ranked, and so avoids the severest form of the maximal word condition. Constraints that preserve affixes are ranked lower, allowing other effects of the maximal word constraints to show. Faithfulness between bases and reduplicants is ranked lowest, allowing the maximal word constraints to have full force, as illustrated in section 2.

As in Māori, we may expect maximal word limits in languages to vary in their domain of application. For example, root-faithfulness may be ranked so highly that roots are unaffected; in that case, maximal word effects may only emerge in affixation or reduplication. In fact, it seems that often the only effect of PrWd size limits is seen in reduplication and truncation.

5. Conclusions

The aims of this paper were to show that maximal word limits exist and that such conditions reduce to general prosodic constraints. To that end I showed that PrWd size limits control the Māori passive suffix's realizations.

The analysis of the passive also showed that the maximal word limit is effected by several separate, violable constraints. The most important part of the analysis, though, was that the constraints make no mention of PrWd size; they are general prosodic constraints, requiring footing (LAPSE_{FT}, *FT-) and foot alignment (ALLFTL). In short, maximal word limits can be subsumed under the general enterprise to reduce all size-related restrictions to general properties of CON, providing a single theory to cover reduplication, truncation, templatic morphology.

References

Bauer, W. (1993). *Maori.* London & New York: Routledge.

Biggs, B. (1961). The structure of New Zealand Maaori. *Anthropological Linguistics* 3: 1–53.

Blevins, J. (1994). A phonological and morphological reanalysis of the Maori passive. *Te Reo* 37: 29–53.

Crowley, T. (1999). *Ura: a disappearing language of southern Vanuatu.* Canberra: Australian National University.

de Lacy, P. (1998). A cooccurrence restriction in Maori. *Te Reo* 40:10–44.

de Lacy, P. (1999). Circumscriptive morphemes. in C.Smallwood and C.Kitto (eds). *Proceedings of AFLA VI.* Toronto: University of Toronto.

Goldsmith, J. (1976). *Autosegmental phonology*. Doctoral Dissertation, MIT.

Green, T. & M. Kenstowicz. (1995). The lapse constraint. ms. MIT. Rutgers Optimality Archive #101.

Hale, K. (1968). Review of Hohepa (1967) – 'A Profile Generative Grammar of Maori'. *Journal of the Polynesian Society* 77: 83–99.

Hale, K. (1973). Deep-surface canonical disparities in relation to analysis and change: an Australian example. In T.Sebeok (ed). *Current Trends in Linguistics 11*. The Hague: Mouton.

Hale, K. (1991). Remarks on G.Sanders "Levelling in the history of Polynesian passive formations". *Journal of the Polynesian Society* 100: 99–101.

Harlow, R. (1991). Consonant dissimilation in Maori. In Robert Blust (ed.) *Currents in Pacific linguistics: papers on Austronesian languages and ethnolinguistics in honour of George W. Grace* Pacific Linguistics Series C, no.117. Canberra: The Australian National University, pp.117–128.

Hohepa, P. (1967). *A profile generative grammar of Maori.* Bloomington, Ind.: Indiana University Press.

Ito, J. and A. Mester. (1992). Weak layering and word binarity. University of California at Santa Cruz. Linguistic Research Center Report, pp.92–109.

Ito, J. and A. Mester. (1994). Japanese phonology. In J.Goldsmith (ed.) *The handbook of phonological theory*. Cambridge, MA: Blackwell, pp. 817-838.

Kawasaki, N. (1990). Cooccurrence restrictions on consonants in some Polynesian languages. *University of Massachusetts Occasional Papers (UMOP)* 13: 49–84, Amherst, MA: GLSA Publications.

Keegan, P. (1996). *Reduplication in Maori*. MPhil dissertation, University of Waikato.

Kenstowicz, M. & C. Kisseberth (1979). *Generative phonology: description and theory*. San Diego: Academic Press.

Kibre, N. (1998). Formal property inheritance and consonant/zero alternations in Maori verbs. Rutgers Optimality Archive #285.

Kiparsky, P. (1971). Historical linguistics. In William Dingwall (ed.) *A survey of linguistic science*. University of Maryland, pp.577–649.

Liberman, M. and A. Prince (1977). On stress and linguistic rhythm. *Linguistic Inquiry* **8**: 249–336.

McCarthy, J. (1981). The role of the evaluation metric in the acquisition of morphology. in C. L. Baker and J. McCarthy (eds.) *The logical problem of language acquisition*. Cambridge, MA: MIT Press, pp. 218–248.

McCarthy, J. (1993). Template form in prosodic morphology. In L. S. Stvan and et al. (eds). *Papers from the third annual formal linguistics society of Midamerica conference.* Bloomington, IULC, pp. 187–218.

McCarthy, J. (2001). *A thematic guide to Optimality Theory*. Research Surveys in Linguistics, Cambridge University Press.

McCarthy, J. and A. Prince. (1986). *Prosodic morphology*. ms. University of Massachusetts, Amherst & Rutgers University.

McCarthy, J. and A. Prince (1990). Foot and word in prosodic morphology: The Arabic broken plural. *Natural Language and Linguistic Theory* 8: 209–283.

McCarthy, J. and A. Prince. (1993a). *Prosodic morphology I*. Rutgers Technical Reports TR-3.

McCarthy, J. and A. Prince (1993b). Generalized alignment. In G. Booij and J. van Marle (eds.) *Yearbook of Morphology*. Dordrecht, Kluwer. pp.79–153.

McCarthy, J. and A. Prince (1994a). Two lectures on prosodic morphology. Rutgers Optimality Archive #59.

McCarthy, J. and A. Prince (1994b). The emergence of the unmarked: Optimality in prosodic morphology. In M.Gonzàlez (ed) *Proceedings of NELS 24*. Amherst, MA, GLSA. pp.333–379.

McCarthy, J. and A. Prince. (1995). Faithfulness and reduplicative identity. In J. Beckman, S. Urbanczyk and L. Walsh Dickey. (eds) *University of Massachusetts Occasional Papers in Linguistics (UMOP) 18*. Amherst, MA, GLSA: 249–384.

Mester, A. (1994). The quantitative trochee in Latin. *Natural Language and Linguistic Theory* 12: 1–62.
Ngata, H.M. (1993). *English-Maori dictionary*. Wellington: Learning Media.
Prince, A. (1983). Relating to the grid. *Linguistic Inquiry* 14: 19-100.
Prince, A. and P. Smolensky. (1993). *Optimality Theory: Constraint interaction in generative grammar.* Rutgers Technical Reports TR-2, Rutgers University.
Sanders, Gerald (1990). On the analysis and implications of Maori verb alternations. *Lingua* 80: 149–196.
Sanders, G. (1991). Levelling and reanalysis in the history of Polynesian passive formations. *Journal of the Polynesian Society* 100:71–90.
Selkirk, E. (1984). *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.
Selkirk, E. and K. Tateishi. (1988). *Minor phrase formation in Japanese.* In L. Macleod and et al. (ed) *Papers from CLS 24.* Chicago: CLS, pp. 316–336.
Truckenbrodt, H. (1995). *Phonological phrases: their relation to syntax, focus, and prominence.* Doctoral dissertation, Massachusetts Institute of Technology.
Williams, H.W. (1971) [1844]. *A dictionary of the Maori language*. 7th edition. Wellington, N.Z.: A.R.Shearer, Government Printer.

Paul de Lacy
Linguistics Department
University of Massachusetts
Amherst MA 01003
USA

delacy@linguist.umass.edu