Repetition and its Avoidance: The Case of Javanese

Moira Yip

University of California, Irvine

It is argued that echo-words result from the tension between a requirement that *penalizes* a sequence of two identical stems, *REPEAT(Stem), and one that *requires* two identical stems, REPEAT(Stem). Based primarily on data from Javanese, I make three main points. First, at least some inputs to the Optimality Grammar must be abstract morphological specifications like PLURAL. They are phonologically incomplete outputs of the morpho-syntax. Second, morpheme realization results from an attempt to meet output targets in the form of constraints: REPEAT, σ_2 =a; PL=s, and so on. Such morphemes do not have underlying forms in the familiar sense (cf Hammond 1995, Russell 1995). Third, the target constraints may be out-ranked by phonological constraints of various kinds, particularly constraints against the repetition of elements, here called *REPEAT. The elements may be phonological (feature, segment) or morphological (affix, stem). These findings support the view of Pierrehumbert (1993a) that identity has broad cognitive roots. The primary data comes from Javanese, but the paper also touches on English and Turkish.[1] Section 1 gives some background on the handling of morphological data in OT. Section 2 discusses identity avoidance in morphology, sets out the basic proposal, and gives sketches of English and Turkish. Section 3 is an extended discussion of Javanese. Section 4 looks at secret languages, and section 5 sums up.

1. Blurring of morphology/phonology boundaries in Optimality Theory:

Optimality Theory (Prince and Smolensky 1993, McCarthy and Prince 1993a, and a host of others) as currently conceived is a theory of not only phonology but also many aspects of morphology. It includes mechanisms for controlling the size and content of reduplicative morphemes, constraints responsible for the precise placement of affixes, constraints that explain the choice among allomorphs, and constraints that pick the right member of a suppletive set. The essence of Optimality Theory is that it is an output-based grammar in which all possible outputs for some input are assessed by a universal set of ranked and violable output constraints. The optimal candidate is evaluated as follows. All outputs which violate the highest ranked constraint are thrown out, and those remaining are evaluated by the next highest ranked constraint. This procedure continues until only a single candidate survives. In the event of a tie at any point in the procedure, the tying candidates are passed on down to the next constraint, which decides matters. Let us see how a selection of morphological phenomena is handled in Optimality Theory.

---

[1] This paper has benefited greatly from comments from audiences at this workshop, at the Conference on Morphology and its relationship to Syntax and Phonology at UC Davis, from the participants in the seminar on Constraints in OT at UC Irvine, and the workshop on Theoretical East Asian Linguistics, UC Irvine, and from a number of individuals, including Diane Brentari, John McCarthy, Orhan Orgun, and David Perlmutter. Special thanks to Diana Archangeli, Dirk Elzinga, Keiichiro Suzuki, and the rest of the Arizona phonology community for making the workshop and this volume possible.

1.1 Reduplication:
McCarthy and Prince 1993a, 1994 lay out an approach to reduplication in which an abstract input morpheme, RED, passes through GEN and results in a set of output candidates in which RED is realized as a full or partial copy of the base. The choice among these candidates is governed by a set of constraints that determine the size of the reduplicant (such as RED=σ), and its content, controlled by a set of constraints that enforce identity between base and RED, and prefer maximal copying. If RED=σ is ranked above MAX(imality), the reduplicant will be monosyllabic (e.g. Ilokano *bas-basa*, *da-da.it*). If MAX is the higher-ranked of the two, reduplication will be total. (e.g. Yoruba *agba-agba*, *oru-oru*) These constraints can also interact with syllable structure constraints. For example, if NOCODA dominates MAX, the reduplicant will be coda-less (e.g. Balangao: *tagta-tagtag, tayna-taynan*).

1.2 Affix placement:
A family of Alignment constraints (McCarthy and Prince 1993b) aligns the edges of prosodic and morphological categories with themselves and with each other. A purely phonological alignment phenomenon would be the placement of feet at the ends of prosodic words: ALIGN-LEFT: (PrWd, Foot) (e.g. English *(Táta)ma(góuchi)*, * *Ta(táma)(góuchi)*). A purely morphological example would be the placement of an affix at the beginning of a stem: ALIGN-LEFT(Affix, Stem) (Tagalog prefix *ag-*). A morphology/phonology interface example would place a foot at the end of a root, ALIGN-RIGHT (Root, Foot) as in Indonesian *bi(cará)-kan* (Cohn and McCarthy 1994)

Particularly striking results come in the treatment of affixes that vacillate between prefixation and infixation as a consequence of the interaction between these alignment constraints and syllable structure constraints. For example, in Tagalog the prefix *um-* of *um-aral* is placed in position by a constraint ALIGN-LEFT (*um-*, stem), but this constraint is dominated by NO-CODA. The result is to force infixation of *um-* before C-initial roots, such as *gr-um-adwet*, since the prefixed form **um-gradwet* would have an extra coda.[2] Zoll (1994) has used a similar approach to explain the behavior of morphemes that surface as floating features at varying positions in the root, like Chaha imperative palatalization; in some cases these moveable affixes may surface as independent segments as well, like Yawelmani suffixal glottalization.

1.3 Allomorphy:
Mester (1994) proposes that a prosodic selection process in the lexicon can pick one allomorph from a set of alternatives by looking at which would form the optimal output with respect to a set of constraints. He studies Latin perfect stems, which can be formed by attachment of either *-u-* or *-s-*. The default choice is attachment of *-u*, e.g. *mon-u-i:*, but in stems with final heavy syllables, *-s* is used instead: e.g. *auk-s-i:* **aug-u-i:*. He suggests that this can be understood as the avoidance of an output in which a single light syllable, *.u.*, cannot be incorporated into a foot because it is 'trapped' between two heavy syllables (one from the root, and one from the final suffix); here I mark foot boundaries with [ ]:

(1)	*aug u i:	cf	auk si:
	[_] _ [_]		[_] [_]

The prosodic selection process thus picks *auk-si:* from a set of alternative outputs {*aug-u-i:, auk-s-*

[2] The root codas survive because PARSE dominates NOCODA. Note that ALIGN-ag >> NOCODA >> ALIGN-um, since *ag-* is always prefixed even if it results in NO-CODA violations.

i:} for the input /auk-PERF-i:/.

1.4 Suppletion:

Tranel (1994) examines French determiners, where the feminine 1-sg-poss *ma* is replaced by the masculine *mon* before vowel-initial feminine nouns, and the masculine *ce* 'this' is replaced by the feminine *cet* before vowel-initial masculine nouns.[3] Standard accounts simply stipulate this distribution, but Tranel's insight is that both suppletions supply an onset for the following syllable, and are thus phonologically driven. He suggests that a suppletion set is judged against constraints that require gender agreement, and onset satisfaction, and that no one form is basic. Gender agreement can be over-ridden by the need for an onset: ONSET >> GENDER. The result will be, correctly, that the grammar will pick the C-final candidate before a V-initial noun, irrespective of gender.

(2)

	ONSET	GENDER
a. ☞ mon_M $abbé_M$		
b. ma_F $abbé_M$	*!	*
c. ☞ mon_M $arme_F$		*
d. ma_F $arme_F$	*!	
e. ce_M $abbé_M$	*!	
f. ☞ cet_F $abbé_M$		*
g. ce_M $arme_F$	*!	*
h. ☞ cet_F $arme_F$		

These results make it hard to identify a clear dividing line between morphology and phonology. What is more, they go much further to blur the distinction than does the interleaving of phonology and morphology found in lexical phonology. In lexical phonology, each component has its own character: the entities are different, and the rules are different. In Optimality Theory, this is not necessarily the case. Alignment is the most striking example. Alignment appears to play a role in pure morphology, in pure phonology, and at the interface.

In this paper, I want to focus on another area in which phonology and morphology appear to overlap, the area of identity avoidance. It is a commonplace in phonology that sequences of adjacent identical elements are avoided, and this is enshrined as the Obligatory Contour Principle, or OCP (Leben 1973, McCarthy 1986, Yip 1988, Odden 1988, Myers 1993, Pierrehumbert 1993a, and others). What has received less attention in OT are superficially similar cases in morphology, although the generative literature includes many such cases. See for example Stemberger 1981, Menn and MacWhinney 1984, Hyman and Mchombo 1992.

[3] *cet* and *cette* are orthographically distinct, but both are phonetically [sɛt]. I follow Tranel in assuming that both are feminine, and the orthography is irrelevant.

2. Identity Avoidance in Morphology

Avoidance of identity in morphology takes several forms. I will divide them into four categories.

(3) a. The same morpheme cannot appear twice in the same word
b. Different but homophonous morphemes cannot appear in the same word, or otherwise adjacent in the sentence
c. Homophonous morphemes cannot appear on adjacent words
d. The output of reduplication cannot be total identity

The first type is rare, perhaps non-existent, but it is not clear that the morpho-phonology underlies this: in general it seems likely that syntactic and morpho-syntactic principles will achieve this end without identity avoidance being involved at all.

The second type is quite common; the references cited above include numerous examples. A familiar and typical example is the English possessive plural: **cats's*, *cats'*. Further examples include Mandarin perfective *le* and Currently Relevant State *le (*Chao 1968, Li and Thompson 1981), Classical Greek determiners (Golston 1994) and Mandarin third person pronoun *ta (*Yeh 1994). A common response in these cases is omission of one morpheme, with the remaining one carrying the semantics of both. This phenomenon is called haplology.

The third type involves identical morphemes attached to adjacent words, but where the morphemes themselves are not string adjacent. Since the presence of a morpheme on one word does not satisfy the requirements of the second word, omission of a morpheme is rarely the preferred strategy for resolving the situation; instead we are more likely to see syntactic movement, replacement by an alternative morpheme, or simple blocking. Cases of this kind include English *-ing* (see Ross 1972, Milsark 1988, and, for a different view, Pullum and Zwicky (1991)), and Hindi *-ko* Dative and Accusative markers (Mohanan 1992). In the Hindi example, sequences of two NP's, each marked by the suffix *-ko,* are avoided. For a discussion of cases of types two and three, see Yip (forthcoming).

The fourth type is the primary focus of this paper. I will address the phenomenon of so-called "echo words": reduplication accompanied by a small change such that the two halves are not quite identical. English *table-shmable* is an example of an echo-word. I will propose that these result from a tension between two constraints, one requiring repetition (reduplication) and one banning repetition (identity-avoidance). I will begin with an overview, then I will use English and Turkish as brief illustrations of aspects of my proposal. Finally, a complex example of identity avoidance in Javanese echo-words will be discussed at some length.

2.1 A Summary of the Proposal:

The central theme of this paper is the avoidance of complete identity. In phonology the OCP has been the usual way of addressing such issues, but the term OCP becomes less useful when one looks at morphology. Firstly, carrying the term over into morphology tends to imply that morphology is a sub-branch of phonology, but this is of course not true; it is more that both phonology and morphology are subject to a single general principle that avoids repetition. Secondly, talking of "contours" in the domain of morphemes is inappropriate. In the rest of the paper I will use constraints of a family I will call *REPEAT, as defined sweepingly below.[4]

---

[4] Earlier versions of this paper used either the OCP, or *ECHO. Neither of these terms was entirely satisfactory. I owe a debt for the name *REPEAT to Menn and MacWhinney's Repeated Morph Constraint (1984).

(4) *REPEAT: Output must not contain two identical elements

Like many constraints, this is subject to adjacency effects, and it may also be judged gradiently at a featural level. The consequence is that violations will be more serious the nearer two things are, and the more similar they are. In most of this paper these subtleties will play no role.

The model I am proposing has two main parts. I outline the proposal below; further details will become clear during the body of the paper. First, there is a set of UG constraints:

(5) REPEAT: Output must contain two identical elements
*REPEAT: Output must not contain two identical elements
*REPEAT(Input): Output must not contain elements identical to input
MORPHDIS: "Distinct instances of morphemes have distinct contents, tokenwise" (McCarthy and Prince 1995:67)

The REPEAT constraint forces reduplication by self-compounding. Instead of supposing that there is an affix, RED, which must be filled, it assumes that the input has only a morphological annotation such as "PLURAL", and the grammar includes a constraint $\text{REPEAT}_{\text{Plural}}$ which must be satisfied for all inputs. This can be combined if necessary with constraints governing the size of the reduplicant: I will have nothing to say about this latter point.

The *REPEAT constraint blocks complete repetition.[5] If *REPEAT >> REPEAT, we derive the echo-word pattern: reduplication that falls just short of complete identity. This proposal explain why echo-words seem to be most common in the case of word reduplication, where the reduplication would otherwise be total: in the case of partial reduplication, *REPEAT is satisfied anyway by the failure to copy the entire base. The ranking REPEAT >> *REPEAT will mask the effects of *REPEAT completely, giving total reduplication.

The *REPEAT(Input) constraint is a very particular kind of *REPEAT constraint in that it compares input and output, not two portions of the output. It penalizes any output that fully realizes the input. As such, it selects highly opaque outputs. This constraint is necessary to explain language game data, and it is presumably low-ranked, perhaps even absent, from normal grammars. Like ordinary *REPEAT it is held in check by lower ranked REPEAT, which encourages reduplication, so we find that each half of the output is made minimally different from the input.

Summarizing, the tension between these first three constraints gives rise to the following partial typology:

(6)	REPEAT >> *REPEAT	true reduplication
	*REPEAT >> REPEAT	echo words: change in one half
	*REPEAT(Input) >> REPEAT	secret languages: change in both halves

The fourth constraint, MORPHDIS:"Distinct instances of morphemes have distinct contents, tokenwise", is drawn from McCarthy and Prince (1995). They find the need for a constraint that is violated any time a segment does double duty to fulfil more than one morphological role. If this constraint dominates *REPEAT, sequences of homophonous morphemes will be acceptable. If the

[5] I will assume that *REPEAT assesses complete morphological entities, such as stems, but an alternative is to assess all identity as the aggregate of individual identities between pairs of segments.

ranking is reversed, however, we will observe haplology: to avoid repetition, one set of segments is recruited to do the work of two morphemes:

(7)	*Repeat >> MorphDis	haplology
	MorphDis >> *Repeat	no haplology

Repeat bears obvious similarities to various constraints proposed in the OT literature on reduplication, particularly McCarthy and Prince (1993, 1994, 1995). It does the work of two constraints in their 1995 paper, Ident-BR, and Max-BR. For the purposes of this paper, it is sufficient to merge these two into the single Repeat.

The second part of the proposal is given below:

(8) a. Inputs consist of morphologically annotated roots, rather than roots with phonologically specified affixes: /kæt$_{PL}$ /, not /kæt-s/, and /udan$_{HAB\text{-}REP}$/ not /udan-RED/.
b. These are realized in order to satisfy specific output constraints

This is very similar to proposals of Hammond (1995) and Russell (1995). It is also what seems to be assumed by Mester (1994), although not in an OT framework. The primary advantage of this proposal in the present context is that it allows for the absence of an affix precisely when some other affix or the root itself is able to satisfy the output constraint in question. The discussion of English in the next section will illustrate this point, and it will play an important role in Javanese.

2.2 English *'s*:

The best known case of haplology comes from English. The plural /s/ and the possessive /s/ cannot co-occur, although adding possessive /s/ to an irregular plural is fine, and so is adding it to a singular ending in /s/, or even a singular ending in /sɪs/..

(9)	*Singular*	*Plural*	*Possessive Sg.*	*Possessive Pl.*	
	child	children	child's	children's	
	mouse	mice	mouse's	?mice's	
	cat	cats	cat's	cats'	*cats's
	Katz	Katzes	Katz's	Katzes'	*Katzes's
	coreopsis		coreopsis's		

Compare especially *Katz's* vs **cats's*; *coreopsis's* vs. **Katzes's*. Two strategies are used to avoid /s-s/. One strategy is haplology: the omission of a morpheme, as in the possessive plural *cats'*. The other is insertion of a buffer vowel, as in the simple possessive *Katz's*, *Kat*[sɪz] (and between all stridents and suffixal -*s*).

Optimality Theory, as an output-based grammar, is well-suited to capturing Stemberger's (1981) insight that this and other cases of haplology do not appear to involve deletion so much as a failure to insert a superfluous morpheme if a homophonous morpheme is already in the right position. Thus if the plural /s/ is present, a plural possessive can satisfy the need to end all possessives in /s/ without adding a second /s/. This explanation, though, does not extend to the vowel-insertion between a root /s/ and a suffix /s/, and thus no unified explanation is possible. Within Optimality

Theory, we can provide a single straightforward account.

I will now offer an explicit Optimality Theory analysis of the core aspects of identity avoidance, using this as my first example. Suppose, following Myers (1993) , that *REPEAT (ie his OCP) is a constraint that can be ranked with respect to the other constraints of the grammar. Further suppose that *REPEAT is a sort of meta-constraint (Pierrehumbert 1993b) which can be instantiated with different arguments, and includes at least the following family;

(10) *REPEAT (feature) *REPEAT (segment)
*REPEAT (affix) *REPEAT (stem)

Consider a case in which insertion is the preferred remediation strategy. *REPEAT must then dominate some sort of constraint against epenthesis which, following Prince and Smolensky 1993, I will call FILL. Also high-ranked will be the output constraints that require some morphological category to be phonologically instantiated in a particular way:

(11) English *'s*:
a. PLURAL: Plurals must consist of a stem plus an *-s* affix.
b. POSS: Possessives must consist of a phrase plus an *-s* affix.
c. *REPEAT (s): *REPEAT (feature), where feature=[strident]
d. FILL: Don't insert

PLURAL, POSS, *REPEAT(s) >> FILL (Epenthesis as last resort)

In the tableau below, the possessive plural *cats'* with only one *s* wins because the candidate with two *s*'s violates *REPEAT(s), and the candidate with epenthesis violates FILL. Crucially, the single *s* satisfies the PLURAL and POSS constraints.

(12)

cat_{PLPOSS}	PLURAL=s	POSS=s	*REPEAT(s)	FILL
cat_{PLPOSS}-s-s			*!	
☞ cat_{PLPOSS}-s				
cat_{PLPOSS}-s-I-s				*!

In the possessive of *Katz's*, the affix must be retained to satisfy POSS. FILL is thus violated in order to satisfy the higher-ranked *REPEAT(s).

(13)

$Katz_{POSS}$	PLURAL=s	POSS=s	*REPEAT(s)	FILL
$Katz_{POSS}$-s			*!	
$Katz_{POSS}$		*!		
☞ $Katz_{POSS}$-I-s				*

These tableaux demonstrate that the omission of one affix after the possessive plural of *cat* versus the epenthesis into the simple possessive of *Katz* follow from the dominance of *REPEAT(s), and of the output requirement that the plural and the possessive must end in an *'s* morpheme. This output requirement blocks deletion of a lone plural or possessive morpheme, and *REPEAT(s) forces use of the fall-back strategy, epenthesis. Two *'s* affixes will never be optimal, because they will always violate either *REPEAT(s), if adjacent, or FILL, if separated by an epenthetic vowel, and there is always available a candidate with only one affix that violates neither. This analysis thus allows us to link the morphological "haplology" of the plural and possessive morphemes with the phonological epenthesis of the English Plural Rule by assuming that *REPEAT(s) plays a role in both "components".[6]

English demonstrates the advantages of assuming that affixes are not present underlyingly, but are a response to satisfying an output constraint. In the next section we will see the role of *REPEAT when it interacts with reduplication.

2.3 Echo-word Formation: Turkish

Many languages have reduplicative processes that replace one portion of the reduplicant with fixed segmental material. English *table-shmable* is an example of such a process: see Yip (1992) for a range of cases. The segmental material is sometimes arguably the default segment of the language, as argued by McCarthy and Prince for Akan, and Yip for Chaoyang (1993). In other cases, however this is not so: no-one has argued that /ʃ, m / are the default consonants of English. A striking characteristic of many such word formation processes is that if the input contains segments identical to the fixed replacement ones, so that the expected output would mimic total reduplication, the process either does not apply at all, or a different set of replacement segments is used. For example, the Tengxian dialect of Chinese (Deng 1995) reduplicates adjectives, replacing the rhyme of the first half by [ɐŋ]:

(14)	dun	dɐŋ dun	'short
	lɛŋ	lɐŋ lɛŋ	'cold'
	kou	kɐŋ kou	'tall'

This system is very productive, applying to more than 200 adjectives. Systematically, adjectives whose rhyme is [ɐŋ] or [aŋ] fail to undergo this process, instead using one of several alternatives available in the language: /nɐŋ/ does not yield *nɐŋ nɐŋ, but rather [nɐŋ hɐŋ tʃɛŋ].

A second example is drawn from Turkish, which reduplicates the first CV of the adjective to form an emphatic form. This CV addition is followed by a coda consonant from the set /p,s,m,r/, subject to the constraint that this consonant cannot be identical or too similar to any consonant of the base. For details, see Dobrovolsky (1987), Demircan (1987).[7]

(15) a.	kap-kara	'jet black'	ap-aĉik	'wide open'
	cep-cevre	'very much around'	sap-sari	'fully yellow'
b.	sim-siki	'extremely tight'	bem-beyaz	'snow white'

---

[6] As several members of the workshop have pointed out, the epenthesis depends crucially on *REPEAT, but the haplology does not. The haplology follows instead from some notion of economy or faithfulness that penalizes insertion of [s], and from viewing the plural and possessive as output constraints that can be jointly satisfied by a single *'s*. However, the haplologized forms are certainly *consistent* with the *REPEAT ranking necessary for the epenthetic cases.

[7] Thanks to Orhan Orgun for help with this section.

göm-gök 'sky-blue' bum-burusuk

c. kas-kati 'extremely hard' bes-belli 'unmistakably obvious'

d. ter-temiz 'spotless' sir-siklam 'wet through'

tor-top 'fully round'

The precise choice of consonant depends on a number of factors, and there is some degree of freedom, but the avoidance of repetition is a major consideration. Closer consonants, and coda consonants, exert more influence than do more distant ones, in line with the view of identity avoidance put forward in Pierrehumbert (1993a).

This echo-word type of reduplication accompanied by melody replacement shows a clear tension between a desire for repetition, which can be seen as the need to satisfy a constraint REPEAT, and avoidance of repetition, or a satisfaction of a constraint *REPEAT. *REPEAT is higher ranked, ruling out total reduplication, but REPEAT plays a central role in ensuring that the overall system is still one of reduplication, with only a minimal difference between base and reduplicant. In the next section I will discuss one complex case of this type, Javanese, in some detail, showing how this tension is played out. I will offer arguments in favor of viewing this type of reduplication as the satisfaction of a REPEAT constraint, rather than the addition of an underlying RED affix.

First, however, the Turkish data provide a good example of the advantages of dispensing with a set of underlying forms in favor of an abstract underlying marker, here "Intensive", and allowing GEN to generate a set of options from which the grammar chooses.[8] It is clear that the inserted consonants should not be individually listed in the lexicon word-by-word, because then we would expect that any consonant of Turkish could appear in the prefix in some word. Deriving all of *p,m,r,s* from a single underlying phoneme is unappealing given the constraints that would be needed to select, say, [r] as an optimal output for /p/ (Dobrovolsky 1987). So this leaves two alternatives: (a) a set of *p,m,r,s* inputs, each submitted to GEN and judged by a set of output constraints or (b) a single abstract input, which has among its outputs the *p,m,r,s* forms. In each case the set of *p,m,r,s* constraints would still be needed, so the multiple inputs of option (a) become superfluous. Alternative (a) seems close to what Tranel (1994) has in mind, whereas (b) is more abstract, along the lines of Mester (1994), Hammond (1995) and Russell (1995).

Choosing option (b), then, let us assume that GEN produces multiple options for each Intensive word. A set of preference constraints for the allomorph picks *p* first (since this is always used before vowel-initial roots), then *m, s,* or *r,* then any other consonant. In the following tableau, "p" is short for "The Intensive prefix should end in [p]", and so on.

(16)

aji_{INT}	p	m	s	r	Other C
☞ap-aji		*	*	*	*
am-aji	*!		*	*	*
as-aji	*!	*		*	*
ar-aji	*!	*	*		*

[8] Similar conclusions have been reached by Russell (1995) and Hammond (1995).

A set of *REPEAT constraints over-rides the preference for [p]; using 3 plausible and self-explanatory constraints, based roughly on Demircan (1987), we have the following tableau; note that *REPEAT here is violated even by similar but non-identical consonants like *p* and *b*.

(17)

$burusuk_{INT}$	*REPEAT-C_1	*REPEAT-C_2	*REPEAT-[-son] codas	p
bup-burusuk	*!			
☞ bum-burusuk				*
bus-burusuk			*!	*
bur-burusuk		*!		*

No clear ranking arguments are available for any of the first three constraints.

The Turkish data show *REPEAT in a particularly forceful way, and also reinforce the advantages of an abstract underlying representation. We now turn to the most detailed case, Javanese. Javanese avoids repetition of two kinds: repetition of the entire stem, and repetition of the vowel [a]. It also has two output constraints, REPEAT, and a requirement that the second syllable have the vowel [a]. These interact in interesting ways, as we see in the next section.

3 Javanese:

3.1 Overview of Basic Habitual Repetitive Formation

Javanese has a pattern of reduplication that is usually referred to as Habitual Repetitive, shortened to Hab-rep. It applies to verbs, adjectives, and even nouns. The whole stem is reduplicated, and then the vowel in the last syllable of the first half is replaced by [a]. Most roots are bi-syllabic, so usually the second syllable has the vowel [a]. However, if the stem is longer or shorter it becomes clear that the locus of [a] is consistently the final syllable of the first half. Some typical data is given below; all examples are given in phonemic transcription unless allophonic details become relevant. Javanese has six vowel phonemes, /i,u,e,o,a,ə/. For a full treatment of the phonology of Javanese vowels see Dudas (1968), Yallup (1982). The data here is drawn from Dudas (1968), Kenstowicz (1986), and Horne (1964).

(18) **Normal pattern of Habitual-Repetitive (Hab-reps) Reduplication:**

eliŋ	elaŋ-eliŋ	'remember'
tuku	tuka-tuku	'buy'
eleʔ	elaʔ-eleʔ	'bad'
bul	bal-bul 'puff'	

In most cases, it is not possible to tell whether reduplication here is prefixing or suffixing in nature, a point made independently by McCarthy and Prince (1995). This suggests that it is not a type of affixation at all, but rather compounding of the stem with itself, with both halves of equal status. The reduplication is accomplished in response to the constraint REPEAT(Stem), which rules out any output without a reduplicated stem.

I will now motivate the claim that the reduplicative unit, and thus the unit referred to in the

*REPEAT constraint, is the stem. In most cases root, stem, or word are all plausible candidates for analysis, but Hab-reps of doubled causatives, which are themselves bi-morphemic, are informative in this regard. The following data from Dudas show that the reduplicant includes nasal prefixes, and the initial consonant of suffixes after vowel-final roots.

(19)		*root*	*hab-rep*	*doubled causative*	*hab-rep causative*
	'job, task'	gawe	gowa-gawe	ŋgaweʔ-ŋgawe-ʔake	ŋgowaʔ-ŋgawe-ʔake
	'mistaken'	salah	salah-seleh	njalah-njalah-ake	njalah-njeleh-ake
			solah-salah		njolah-njalah-ake

At first glance, it might seem that reduplication maximally satisfies a σσ template, but the following data show that this is not correct:

(20)			
	bul	bal-bul	'puff'
	melaku	meloka-melaku	'walk'

The generalizations are as follows: (i) the reduplicant always includes the entire root and (ii) affixal material may be included up to σσ, especially to close the final syllable. A striking fact about Javanese and related languages is that the canonical root has the form (C)VC(C)VC: two syllables, the second of which is closed. The reduplicant apparently aspires to that shape, suggesting that the relevant unit is root or stem. Clearly the Hab-rep causative reduplicants are not roots, but they could be stems. Stem, after all, is a derived morphological notion, unlike root. Suppose, then, that the speaker imposes a canonical stem analysis on these words, as shown below: (- marks affixal boundaries, | marks stem boundaries)

(21)		
	Actual morphological structure	ŋ-gawe-ʔake
	Morphological "reanalysis"	|ŋ-gawe-ʔ|ake

Compounding of the stem with itself then gives |ŋ-gawe-ʔ||ŋ-gawe-ʔ|ake. For the causative, this is the surface form. For the Hab-rep, the σ_2=a constraint forces vowel changes.

One possible formal treatment of the stem re-analysis would involve a set of constraints like the following:

(22) ALIGN-R (Stem, C) (Takes in affixal C)
ALIGN-R (Root, Stem) (Keeps root as near end as possible thus taking in only minimal C material)
ALIGN-L (MWd, Stem) (Takes in prefix)

ALIGN-R (Stem, C) >> Align-R (Root, Stem)

For reasons of space I do not give the relevant tableau here.

These causatives also provide an argument in favor of treating this type of reduplication as the satisfaction of an output constraint,, rather than the attachment of a RED affix, which is then filled by base material (McCarthy and Prince 1993a). These particular causatives (though not all causatives) are themselves reduplicated forms, and Dudas implies they are the base for the Hab-reps. If they served as a base for attachment of a Hab-rep RED morpheme, we would expect additional reduplication, giving something like either **njalah-njeleh-njeleh-ake* or **njalah-njalah-njeleh-njeleh-ake*. Instead, there is no extra reduplication: the single (causative) reduplication is enough to satisfy the Hab-rep's need for a reduplicated output. I formulate the constraint below.[9]

(23) REPEAT(Stem) : Hab-Reps must consist of two identical stems.

This situation - one reduplication satisfying both causative and Hab-rep expectations - is strongly reminiscent of the way [s] can satisfy the needs of both the plural and the possessive in English. See Russell (1995) on Nisgha for many similar cases.

3.2 The Introduced Vowel [a]

The introduced [a] would traditionally be analyzed as an affix that forms part of the Hab-rep morphology. I will argue that its appearance is instead the result of an output constraint requiring the vowel of the appropriate syllable to be [a]. I formulate this constraint below:

(24) **σ_2=a**: The final syllable of the first half of Hab-reps must have an [a] nucleus

The interest of the Javanese Hab-reps lies in their diverse mechanisms for avoiding identity of various kinds. First, the output may never have both halves completely identical to each other. The constraint in (24) achieves this immediately if the input ends in any vowel other than /a/, but what if it ends in /a/? The data are given below:

(25)	udan	udan-uden	'rain'	*udan-udan
	kumat	kumat-kumet	'have a relapse'	*kumat-kumat
	edan	edan-eden	'crazy'	*edan-edan
	tak	tak-tek	'tap'	*tak-tak

Simple satisfaction of **σ_2=a** would result in perfect total reduplication. Instead, the vowel of the second half dissimilates to [e]. The following constraint embodies the avoidance of total identity typical of Hab-reps; I should emphasize that other forms of reduplication in the language do allow complete reduplication , such as *abat-abat* 'century, PL'.

(26) *REPEAT (Stem): Hab-reps must not consist of two identical stems.

These two constraints are both surface true and undominated.

The first argument in favor of treating [a] as the response to an output constraint, rather than as an affix, is based on the fact that identity violations can never be resolved by changing this introduced [a]. The dissimilation site is always the *other* /a/. If /a/ were an affix, it would be necessarily to somehow stipulate the choice of target, but the output-based analysis immediately explains the immunity of the introduced /a/ to change. The following tableau demonstrates this point; here and throughout I simplify the output forms for ease of exposition. For instance, *udan-uden* is the surface realization of the output form /udan-ude<[low, back]>n/, where the [low, back] features

[9] I put aside until later the issue of how to deal with morpheme-specific constraints in OT.

of the input vowel /a/ have been left unparsed. And *tuka-tuku* is the surface realization of the output form /tuk<u>𝔸-tuku/, where one root vowel has been left unparsed, and the vowel [a] has been inserted by GEN, as shown by the outline capital letter 𝔸.

(27)

/udan/	σ_2=a	*REPEAT(Stem)
☞ a. udan-uden		
b. uden-udan	*!	
c. udan-udan		*!

If reduplication is compounding, a change in either half is a Faithfulness violation, a failure of Input-Output correspondence, not of Base-Reduplicant correspondence. So it follows that the change of /a/ to [e] violates Faithfulness (more specifically, PARSE-Feature, see section 3.3.2), which must thus be ranked below *REPEAT. The ranking can be validated by the following tableau:

(28) *REPEAT(Stem) >> FAITHFULNESS

/udan/	*REPEAT(Stem)	FAITHFULNESS
☞ a. udan-uden		*
b. udan-udan	*!	

Roots like *udan* with /a/ in the second syllable could satisfy σ_2=a in one of two ways. Either they could introduce a special /a/ (like /tuku/ does to form *tuka-tuku*), or they could use the underlying /a/ itself. Faithfulness requires that using the underlying /a/ will be optimal. From these data there is no way to tell if this is correct, but we shall see later that the underlying /a/ can be used when available. This will provide the second and clinching argument that /a/ is not an ordinary affix, but an output requirement that can be met by either an available underlying /a/, or one inserted by GEN.

A different kind of identity avoidance is found if the input has /a/ in the first syllable. From what we have seen so far, we would expect to find outputs in which the first half has /a/ in both syllables, so that /lali/ would have a Hab-rep *lala-lali*, but instead we observe dissimilation of the root /a/ to [o]:

(29)	lali	lola-lali	'forget'	*lala-lali
	adus	odas-adus	'bathe'	*adas-adus
	melaku	meloka-melaku	'walk'	*melaka-melaku

Following the same analytical approach used above, I formulate the following constraint:

(30) *REPEAT (a): Sequences of /a/ are not allowed.

The domain of this constraint is the stem, with the introduced [a] analyzed as part of the stem, confirming our earlier claim that it is not an affix. Note that the constraint does not apply across stem boundaries, since *lol**a**-l**a**li* is well-formed. Following precisely parallel arguments to those we used for a>e, we may understand why it is the root /a/, not the introduced [a], that changes to [o]. The following tableau demonstrates this point: candidate (b), in which the introduced /a/ has changed to [o], violates σ_2=a, and is thus eliminated.

(31)

/lali/	σ_2=a	*REPEAT(a)
☞ a. lola-lali		
b. lalo-lali	*!	
c. lala-lali		*!

The ranking of *REPEAT(a) >> FAITHFULNESS is validated by the following tableau:

(32)

/lali-lali/	*REPEAT(a)	FAITHFULNESS
☞ a. lola-lali		*
b. lala-lali	*!	

At this point, then, we have seen two types of dissimilation that conspire to remove identity violations. I leave for further research the question of why one chooses [e] while the other chooses [o] as their output vowel.

Before proceeding to the most interesting cases of roots with /a/ in both syllables, I need to provide a little additional background on the Javanese vowel system.

3.3 Identical Vowel Roots:

3.3.1 Relevant Background on Javanese phonology:

Kenstowicz (1986) has shown that roots with two identical vowels must have a single vowel melody occupying two nuclear slots. The argument rests on the fact that allophonic rules show their effects on both root vowels, even if only one of the vowels is in the proper context. There are two relevant rules. First, mid-vowels lax in closed syllables:[10]

(33)	suwe	'long time'	idjɛn	'alone'
	tjuwo	'bowl'	katɔn	'support'

Second, low vowels round and raise in word-final position:

(34)	medja-ku	'my table'	medjɔ	'table'
	djiwa-mu	'your soul'	djiwɔ	'soul'

[10] In general, mid-vowels will be shown by their tense symbols, unless the laxing process plays a role in the discussion.

The data below show that in roots with two identical vowels, these changes affect both vowels even if the first one does not meet the context for the rule

(35)	a. gɔdɔg	'to boil'	lɛrɛn	'to rest'
	b. basa-mu	'your language'	bɔsɔ	'language'

Kenstowicz (1986) argues that this behavior is to be expected if these roots have one melody linked to both nuclei. This representation in turn follows if Javanese obeys the OCP as an MSC on the vowel tier.[11] The analysis has two consequences when we return to Hab-reps. First, the *REPEAT(a) constraint can now be seen as a specific instance of a more general constraint *REPEAT(seg) (the OCP) found throughout Javanese vowel phonology. Second, we expect that dissimilation will affect all reflexes of a single stem vowel, even if only one is in the relevant context. This prediction is confirmed in the next section.

3.3.2 Habitual Repetitives Formed from Identical Vowel Roots

The following data show roots where both vowels are /a/. Such roots have three possible Hab-reps, one less common than the other two, and used only when one of the other variants has already been used.

(36)	salah	solah-salah	'make a mistake'
		salah-seleh	
		solah-seleh	(less common)
	lawas	lowas-lawas	'old'
		lawas-lewes	
		lowas-lewes	(less common)

Consider first *salah-seleh*; the dissimilation of the second /a/ to [e] is expected (cf *udan-uden*), but the dissimilation of the first is not (contrast *lola-lali*). Under Kenstowicz's analysis, however, it is precisely as predicted: the single /a/ melody changes to [e], with across-the-board consequences. Indeed, Kenstowicz mentions these data as support for his analysis.

The more interesting aspect of these data has to do with their bearing on the question of whether [a] is an affix. Suppose it were an affix; we have seen from forms like /lali/ > *lola-lali* that roots with initial /a/, and an introduced /a/ in the second syllable, undergo dissimilation of the root /a/ to [o]. We can thus use this as a diagnostic for the presence of an introduced /a/. If /a/ is introduced, we will get *solah*...., but if it is underlying we will get simply *salah*..... , since there is only one melodic /a/ and no violation of *REPEAT(a). Both forms are found; crucially, the *salah*... form shows that not all Hab-reps have an introduced /a/, and thus that /a/ is not an affix but the result of satisfying an output constraint in one of two ways: use an underlying /a/, as in *salah-seleh* or introduce an [a], as in *lola-lali*.

It remains to explain why there are three options, *salah-seleh*, *solah-salah* and *solah-seleh*. Given the resources of the system (GEN's ability to introduce [a], plus two dissimilation mechanisms), these are all and only the ways to satisfy the three undominated constraints σ_2=a, *REPEAT(Stem), and *REPEAT(seg). Let us see why this is so. First, note that ...*salah*... does not

[11] The high vowels behave somewhat differently, but nothing in this paper hinges on our understanding of the high vowels.

violate *REPEAT(a) because it is a single melody filling two nuclei. Second, note that all three options clearly satisfy σ_2=a, and *REPEAT(Stem). Further, the latter constraint correctly rules out **salah-salah*.

But why are all three acceptable? After all, each output also violates Faithfulness in some way, so we might expect that one output would be singled out as optimal because it violates Faithfulness minimally. To answer this question we must look at the examples in more detail. Let me assume for the sake of concreteness that a>e and a>o both involve a loss of features ([low], [back]), and thus violate PARSE-Feature. The insertion of [a] by GEN must involve instead FILL-Seg (Prince and Smolensky 1993). Given our earlier diagnostic for the presence of inserted [a] (a>o dissimilation), the two outputs with *solah...* must have an inserted [a]. Using outline capitals $ for the inserted [a], we have the following partial tableau (unviolated constraints are ignored here):

(37) Predicted but incorrect tableau

/salah-salah/	PARSE-Feature	FILL-Seg
a. salah-seleh	*(e)	
b. sol$ h-salah	*(o)	*!
c. sol$ h-seleh	*(e)*(o)!	*

Under any ranking, we expect that candidate (a) should be optimal, and yet in fact (a-b) are equally good, and (c) is acceptable, though much rarer. What these facts suggest is that the status of the introduced [a] is variable, and its appearance does not violate any constraint, contra the assumption embodied in tableau (37). Suppose that speakers cannot decide whether they are dealing with an affix or not. If they represent it as an affix, they get candidate (b), and there is no FILL-seg violation because the segment is an underlying morpheme, not the inserted $ shown in the tableau. If they do not represent it as an affix, because the output constraint can be satisfied without it, then they get candidate (a). Candidate (c) is clearly inferior, as desired, since it has two PARSE-F violations. Of course, we still need an understanding of the circumstances under which sub-optimal candidates can be used. See Hayes (1993), Kiparsky (1993) Pater(1994) for discussion of optionality in OT.[12]

The picture of /a/ that emerges is of an output-based system that can be satisfied in one of two ways. Firstly, any available stem /a/ may be recruited to the cause. Secondly, a cost-free affix can be used when necessary (*tuka-tuku*), and even sometimes when not strictly necessary (*solah-salah*). Speakers, in other words, are uncertain about the status of /a/, and vacillate when free to do so.

3.4 Interaction of Hab-reps and open syllable a>ɔ

Javanese has one last way of satisfying *REPEAT(Stem) constraint. Recall that /a/ becomes [ɔ] in final open syllables. This process can be used to destroy total identity, eliminating the need for a>e dissimilation. The data below shows the process in action outside the reduplicative system.

---

[12] Since *solah-seleh* is apparently only used *after* one of the other variants, one possibility is that it is a Hab-Rep formed on a Hab-Rep base. If /solah-salah/ were the *input*, we would expect a>e dissimilation in the second half to *solah-seleh*, giving the desired output. Of course, such cyclic derivation is potentially problematic in OT.

(38) medja-ku 'my table' medjɔ 'table'
djiwa-mu 'your soul' djiwɔ 'soul'
basa-mu 'your language' bɔsɔ 'language'

I formulate a descriptive working constraint, leaving its real nature for future research:

(39) *a]$_{PrWd}$: Avoid [+low, -round] at the end of the prosodic word.

I assume that a > ɔ is the result of GEN adding [+rd], and thus that it violates FILL-F. It does not violate PARSE, since all the features of /a/ are parsed.

The Hab-rep data below contrasts the familiar closed-syllable cases we saw earlier with cases of final open-syllable /a/.

(41) a. *Final Closed Syllable:*

udan	udan-uden	'rain'
kumat	kumat-kumet	'have a relapse'
edan	edan-eden	'crazy'

b. *Final Open Syllable:*

medjɔ	(/medja/)	medja-medjɔ	'table'	*medja-medje
tjobɔ	(/tjoba/)	tjoba-tjobɔ	'have a go at'	
sidɔ	(/sida/)	sida-sidɔ	'succeed'	

For these data to fall out right, it must be the case that violations of PARSE-F (i.e. a>e changes) are more serious than violations of FILL-F (i.e. a>ɔ changes). So PARSE-F >> FILL-F. This is shown in the following tableau:

(42)

/sida#/	*a]$_{PrWd}$	*REPEAT (stem)	*REPEAT (a)	PARSE-F	FILL-F
☞ sida-sidɔ# +rd					*
sida-side #				*!	
sida-sida#	*!	*!			

The ranking of PARSE-F >> FILL-F seems very natural, because it captures the observation that feature loss is more expensive because it can lead to neutralization of underlying contrasts (as it does here: /a/, /e/, and /o/ are all phonemes of Javanese), but feature addition is cheaper because it typically does not (/a/ is the only source for [ɔ] in open syllables). To put it another way, the a>e change leads to recoverability problems, whereas the a>ɔ change does not. The allophonic rounding of final /a/ is thus a cheaper way to keep the two halves of the reduplicant non-distinct than the neutralizing a>e change would be.[13,14]

3.5 Summary of Javanese

Let us review the analysis so far. There is a set of constraints, summarized here:

(43) σ_2=a, *REPEAT(Stem), *REPEAT(a), *a]$_{PrWd}$

|

REPEAT(Stem), FILL-Seg, PARSE-F

|

FILL-F

I have proposed that there is a tension between the requirement that *penalizes* a sequence of two identical stems, *REPEAT(Stem), and the one that *requires* two identical stems, REPEAT(Stem). The two constraints are repeated here for comparison.

(44) REPEAT (Stem) : Hab-Reps must consist of two identical stems.
(45) *REPEAT (Stem): Hab-reps must not consist of two identical stems.

It is the former that produces the effect of total reduplication, but this is then minimally destroyed by the latter. The constraint REPEAT(Stem) plays a crucial role in ensuring that the dissimilations we observe are indeed minimal: a single vowel changes, just enough to satisfy *REPEAT(Stem). [15]

The final issue that needs discussion here is the issue of how to handle language-specific and morpheme-specific constraints in OT. Prince and Smolensky (1993) state that all constraints are universal, and that all variation comes from the re-ranking of constraints. For purely phonological constraints, such as those controlling syllable structure, this idea has proved extremely powerful, but for morphological constraints it is much less clear that the strongest claims of universality make sense. It is certainly possible to claim that every language contains a constraint like Tagalog's ALIGN-L (*um*, Stem), and that it is invisible in any language that doesn't have an affix -*um*-, and low-ranked in languages that do have -*um*-, but where it is a suffix, not a prefix. But the conclusion to which one is forced is unpalatable: every grammar includes all the specific morphological constraints found in the set of all human languages, and indeed all possible human languages, even those not yet discovered. A more reasonable approach seems to be to assume that UG contains not tokens of

[13] There are some complications surrounding roots like /lara/ with two /a/ vowels ending in an open syllable. The only acceptable Hab-rep is apparently [lora-lɔrɔ], whereas I would expect [lara-lɔrɔ] to be preferred.

[14] An odd problem arises: why can't a>ɔ be used even in closed syllables, if FILL-F is so low ranked? That is, why is /eden/ >edan-edɔn bad? It may be possible to exclude this by banning instances of [ɔ] derived from /a/ in closed syllables. For example, we do not find [ɔ] in the first syllable of [atmɔ] (cf *atma-ne* 'his soul'). One possible reason for this might be an avoidance of neutralization; [ɔ] is the reflex of /o/ in closed syllables, so hypothetical [ɔtmɔ] could come from either /atma/ or /otma/, if [ɔ] were permitted in closed syllables. I leave this for further research.

[15] One unexplained issue is why only the vowel /a/ ever dissimilates. /udan/ could surface as *udan-idan* and satisfy *REPEAT(Stem), and yet such changes are never found.

constraints, but schemas like the Generalized Alignment schema of McCarthy and Prince (1993b), or *REPEAT schema, like the OCP, that can be instantiated for specific phonological, prosodic, and morphological entities. Individual grammars contain tokens of these constraints, perhaps multiple tokens differently ranked, but UG does not. For English then, all alignment constraints that control prefix placement outrank all syllable structure constraints, so we observe pure prefixing. In Tagalog, the same is true for the prefix *ag-*, which is prefixed to both C- and V- initial stems, but the alignment constraint for *-um-* must be ranked below NOCODA to explain *gr-um-adwet*, so we derive the ranking ALIGN-*ag* >> NOCODA >> ALIGN-*um*.

If this approach is on the right lines, then we can refine the account given here a little as follows. Reduplication is common in Javanese, and most types of reduplication are total. REPEAT(Stem) is thus true for a large class of constructions in Javanese, including the plural, and doubled causatives and locatives. *REPEAT(Stem) is only true for Hab-Reps, so the grammar contains a set of REPEAT(Stem) constraints, and one *REPEAT(Stem) constraint, ranked above its counterpart REPEAT(Stem) constraint.

I have argued that two aspects of Hab-Reps are best handled by output constraints. Firstly, the introduced [a] is present in response to the constraint σ_2=a. Secondly, the reduplication itself, and the dissimilations that accompany it, are the result of a set of contradictory output constraints *REPEAT(Stem), *REPEAT(a), vs. REPEAT(Stem). The Javanese data argues for two conclusions. First, at least some morphology must be handled by output constraints, and the inputs (if any: see Russell 1995) have morphological features, such as the information that something is a Hab-Rep, but not necessarily concrete morphemes, like /a/, or even abstract ones, like RED. Second, conditions that enforce identity, and avoid identity, can make reference to phonological objects, like the vowel /a/, or morphological entities, like Stem.

4. Secret Languages: Input-Output Identity Avoidance

A rather different type of avoidance of repetition can be found in secret languages. Up till now the *REPEAT constraint has penalized identity between two output entities, but secret languages penalize identity between input and output. Their function, after all, is to obfuscate, so that the uninitiated listener cannot understand. I will call this constraint *REPEAT(Input), and its seems likely that it is high ranked only in secret languages, given its communicative disadvantages.

Chinese language games usually involve reduplication, and replacement of some base segments by fixed segments. Data from Cantonese is given below; after reduplication, the first onset is replaced by /l/ and the second vowel is replaced by /i/. See Chao (1931), Yip (1982), Bao (1990) for more details.

(46) Cantonese La-Mi: CVC > lVC CiC

kɔ:ŋ	lɔ:ŋ kiŋ
sat	lat sit

If the input contains the segments /l/ or /i/, the output would be expected to contain syllables identical to the input syllable. However, in these circumstances /k'/ and /u/ are used instead of /l/ and /i/, successfully ensuring that neither syllable is identical with the input, as demanded by the constraint *REPEAT(Input):

(47) *output σ = input σ

t'in	lin t'un	*lin *t'in*	i>u

lat k'at lit *_lat_ lit l>k'
lin k'in lun *_lin lin_ l>k' and i>u

The following tableau demonstrates the analysis, showing how *REPEAT(Input) controls the choice of affix:

(48)

/lin/	*REPEAT(Input)	AFFIX = /l/ AFFIX = /i/	AFFIX = /k'/ AFFIX = /u/
a. lin lin	**!		**
b. kin lin	*!	*	*
c. lin lun	*!	*	*
☞ d. k'in lun		**	

There is a general constraint active in the language which blocks sequences of two Labials, Yip (1988), and which causes dissimilation of a reduplicated Labial coda in the secret language, as shown below:

(49) *REPEAT (Lab)
sap lap sit *la_p_ si_p_

The dissimilation caused by this constraint removes the need to use /u/ instead of /i/, because after the labial coda has dissimilated to coronal, neither syllable is identical to the input.

(50) t'im lim t'in *lim t'un

The tableau below shows how the two *REPEAT constraints can both be satisfied with the regular /l/ and /i/ affixes.

(51)

/t'im/	*REPEAT(Lab)	*REPEAT(Input)	AFFIX = /l/ AFFIX = /i/	AFFIX = /k'/ AFFIX = /u/
a. lim t'im	*!	*		*
☞ b. lim t'in				*
c. lim t'um	*!		*	
d. lim t'un			*!	

In a rule-based analysis, the data in (50) is hard to explain. The replacement of /i/ by /u/ has to be made conditional on the absence of a labial in the first syllable. Here it follows straightforwardly: the *need* for replacement is removed by the coda change.

Although the constraint relating input and output is problematic for earlier versions of OT, McCarthy and Prince 1995 have discovered a need for such constraints in the closely related area of reduplicative identity, where some languages apparently require reduplicant-input identity (I-R Faithfulness) in addition to the more usual base-input identity (B-R identity). What secret languages seem to require is the inverse of both of these: non-identity between base and input, and between reduplicant and input, or *REPEAT(Input). Overall considerations of Faithfulness (in any form), and the reduplicative nature of the construction, which like Javanese compounding reduplication must try to satisfy REPEAT(Stem), keep deviations from the input to the minimum necessary to avoid *REPEAT(Input) violations. In this tableau, for *REPEAT(Input) one star is assigned for each syllable that is identical to the input; for REPEAT(Stem), one star is assigned for each segment that is a deviation from the stem. Candidate (b) has changed an unnecessary number of segments, so it loses to candidate (a).

(52)

/ma/	*REPEAT(Input)	REPEAT(Stem)
☞ a. la mi		**
b. lu mi		***!
c. ma ma	**!	
d. la ma	*!	
e. ma mi	*!	

5. Conclusions

I have argued that avoidance of identity is found in many areas of morphology, not just in phonology, and that these should be given a unified treatment in terms of a family of *REPEAT constraints. The most familiar identity avoidance principle is the OCP, but this would appear to be part of a more general cognitive pattern, as argued by Pierrehumbert (1993a). If *REPEAT has both phonological and morphological manifestations, one might wonder if REPEAT does also. Its morphological role has been discussed in this paper: what would its phonological effects look like? Two suggestions spring to mind: vowel copying might be one such example, and even phenomena like harmony could perhaps be thought of in this way.

Echo-words show an interplay between two diametrically opposed constraints, REPEAT and *REPEAT. In a domain whose most striking characteristic is the repetitive nature of the output we find evidence of the inverse: a dislike of repetition. OT, with its violable constraints, is perfectly suited to capturing this balance.

References

Bao, Z-M 1990 Fan-qie languages and reduplication. *Linguistic Inquiry* 21.

Chao, Y-R 1931 Fan-qie Yu Ba Zhong. *Bulletin of the Institute of History and Philology*, Academia Sinica, II, 320-354.

Chao, Y-R (1968). A Grammar of Spoken Chinese. Berkeley: University of California Press.

Cohn, A. and J. McCarthy 1994 Alignment and Parallelism in Indonesian Phonology. Ms, Cornell and UMass.

Demircan, Ömer 1987 Emphatic reduplications in Turkish. H. Boeschoten and L. Verhoeven, ed, *Studies in Modern Turkish: Proceedings of the 3rd Conference on Turkish.* Tilburg University Press. pp 24-41

Deng, Yurong (1995) The reduplication of monosyllabic adjectives in Tengxian dialect. (In Chinese). *Fangyan* pp33-46.

Dobrovolsky, M. 1987 Why CVC in Turkish Reduplication. In P. Liliw and M. Saari, eds., *The Nordic Languages and Modern Linguistics 6* Helsinki: 131-146.

Dudas, K. M. 1968 The Phonology and Morphology of Modern Javanese Ph D Dissertation, U. of Illinois.

Golston, C. 1994 The Syntax-Phonology Interface in Optimality Theory. Paper presented at *WCCFL* XIII, Berkeley

Hammond, M. 1995 There is no lexicon. Ms, U of Arizona, Tucson.

Hayes, B. 1993 Metrics as an optimization problem: the case of English folk songs. Handout of talk presented at ROW-1, Rutgers, NJ.

Horne, E. 1964 Javanese-English Dictionary. Yale University Press.

Hyman, L. and S. Mchombo 1992 Morphotactic constraints in the Chichewa verb stem. In L. Buszard-Welcher, L.Wee and W. Weigel, eds., *Proceedings of the 18th Annual Meeting of the Berkeley Linguistics Society* : 350-364 Berkeley Linguistic Society, Berkeley.

Kenstowicz, Michael 1986 Multiple Linking in Javanese. *NELS 16*

Kiparsky, P. 1993 Variable rules. Handout of talk presented at ROW-1, Rutgers, NJ.

Leben, Will (1973). Suprasegmental phonology. Doctoral dissertation, MIT.

Li, Sandra and C. Thompson 1981 Mandarin Chinese: A Functional Reference Grammar. Berkeley: U. of California Press.

McCarthy, John (1986b). OCP Effects: Gemination and antigemination. LI 17: 207_263.

McCarthy, John & Alan Prince. 1993a. Prosodic Morphology I: Constraint Interaction and Satisfaction. To appear, MIT Press. Technical Report #3, Rutgers University Center for Cognitive Science. Pp. 184.

McCarthy, John & Alan Prince. 1993b. Generalized Alignment. In Geert Booij & Jaap van Marle (eds.), Yearbook of Morphology 1993. Dordrecht: Kluwer. 79-153. Technical Report #7, Rutgers University Center for Cognitive Science, 1993. ROA-7.

McCarthy, John & Alan Prince. 1994. The Emergence of the Unmarked: Optimality in Prosodic Morphology. In Merce` Gonza`lez (ed.), Proceedings of the North-East Linguistics Society 24. Amherst, MA: Graduate Linguistic Student Association. 333-379. ROA-13.

McCarthy, J. and A. Prince 1995 Faithfulness and Reduplicative Identity. Ms., U Mass Amherst and Rutgers University.

Menn. L. and B. MacWhinney 1984 The Repeated morph constraint. *Language*: 60: 519-541

Mester, R-A 1994 The Quantitative Trochee in Latin. *Natural Language and Linguistic Theory*.12.1: 1-62

Milsark, G. 1988 Singl-ing. LI 19.4: 611-634

Mohanan, Tara 1992 Case OCP: A Constraint on Word Order in Hindi. Ms, National University of Singapore.

Myers, Scott. 1993. OCP Effects in Optimality Theory. Ms., University of Texas, Austin. Pp. 31 ROA-6.

Odden, David (1988). AntiAntigemination and the OCP. LI 19: 451_75.

Pierrehumbert, Janet 1993a Dissimilarity in the Arabic Verbal Roots. In Amy Schafer, ed.,*Proceedings of NELS 23* Amherst, MA: Graduate Linguistics Student Association.pp 367-381

Pater, J. 1994 Optionality and optimality: syllable weight effects in English secondary stress. Ms McGill.

Pierrehumbert, Janet 1993b Alignment and Prosodic Heads. ESCOL 1993.

Prince, Alan and Paul Smolensky. 1993. Optimality theory: constraint interaction in generative grammar. To appear, MIT Press. TR-2, Rutgers University Cognitive Science Center.Pp. 234.

Pullum. G. and A. Zwicky 1991 A misconceived approach to morphology. *Proceedings of WCCFL 10.*

Ross, J. R. 1972 Doubl-ing . LI 3: 61-86

Russell, K. 1995 Morphemes and Candidates in Optimality Theory. Ms, U. of Manitoba.

Stemberger, J. 1981 Morphological Haplology. *Language*. 57.4:791-817.

Tranel, Bernard 1994 French Liaison and Elision Revisited: a unified account with Optimality Theory. Ms, UC Irvine.

Yallup, C. 1982 Phonology of Javanese Vowels. In A. Halim, L. Carrington, and S. Wurm, eds., *Papers from the Third International Conference on Austronesian Linguistics.* Canberra, Australian National University.

Yeh, Ling-hsia 1994 Haplology in Mandarin. Ms., UMass Amherst

Yip, M. 1982 Reduplication and CV Skeleta in Chinese Secret Languages. LI 13.

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

Yip, M. 1992. Reduplication with fixed melodic material. In K. Broderick, ed., *Proceedings of NELS 22* Amherst, Ma: Graduate Linguistics Association.

Yip, M. 1993 The interaction of ALIGN, PARSE-PLACE, and *ECHO in reduplication. Talk presented at Rutgers Optimality Workshop I, Rutgers University, New Brunswick, NJ.

Yip. M . (Forthcoming) Identity avoidance in phonology and morphology. To appear in the Proceedings of the Conference on Morphology and its relation to Syntax and Phinology. UC Davis, May 1995.

Zoll, Cheryl. 1994. Anchors away: a unified treatment of latent segments and floating features. Ms., University of California, Berkeley. Pp. 55.