A Concatenative Theory of Possible Affix Types[1]

Justin M. Fitzpatrick
MIT

This paper builds a theory in which Raimy's (2000a) notion of non-linear phonological precedence is exploited to produce a typology of possible affix types. I investigate the range of logically possible non-linear precedence structures and show how it closely mirrors observed morphological phenomena in the world's languages. The result is that disparate types of morphology can be unified under a single formal mechanism, rather than distributing them across methods of concatenation, prosodic circumscription, and copying/repetition. As a way to constrain this system, I argue for a merger of Yu's (2003) work on infixation and phonological landmarks with Raimy's system of non-linear precedence. I show empirically that a somewhat powerful and idiosyncratic approach to morpho-phonological ordering is necessary and that a combination of non-linearity and anchor points provides some of the power that is needed in the modelling of human language affixation. I then discuss some ways in which the resulting system may be both too weak and too powerful.

1. Introduction

The grammar must contain some mechanism by which a particular bound morpheme is positioned with respect to a stem in a fairly consistent way. In this way, a prefix /ɛn/ as in 'enable' would be differentiated from the suffix /ɛn/ as in 'loosen'. A common solution to this problem (e.g., in Halle and Marantz 1993) is to include diacritics like PREFIX and SUFFIX in the lexical entry of a given morpheme, as in (1).

(1) a. /ɛn/, SUFFIX
b. /ɛn/, PREFIX

This approach to the encoding of the placement of bound morphemes begs two questions. First, what, if anything, do these diacritics have in common formally? Are they just names with completely different information

[1]I want to thank the organizers of the 1st Escola de Verão de Lingüística Formal da América do Sul (EVELIN) in January 2004 for making possible a great meeting of strangers and friends in the middle of the beautiful Brazilian summer. Most of the material here (excluding Thao reduplication, in section 4) was presented in the EVELIN course *Morphological Spell-Out*. My thanks to all the participants and organizers at EVELIN.

content? Given the form of entry in (1), it would seem that this is so, and one might find this answer unsatisfying. Second, why do we have exactly these diacritics with exactly these meaning and not others? Is there some reason language should have only two, or three, or four possible types of "positional diacritics" and not some other number?

In this paper I present a way to address both of these questions. I show that there is a way of encoding morpheme placement under which diacritics are replaced with a system that allows transparent connections between bound morpheme types, in response to the first concern raised above. More importantly, perhaps, this system results in a logical typology of affix types that closely mirrors the set of well-known bound morpheme types in the world's languages, and goes a long way toward answering the second question above. If successful, the theory presented here would answer that question as follows: There are exactly these types of (positionally defined) affixes because these are precisely those that are logically possible given the tools of the system.

I start from the assumption that placement of bound morphemes is encoded idiosyncratically in the lexicon.[2] In developing this system, I examine the representational possibilities allowed by Raimy's (2000a) treatment of reduplication and infixation through non-linear phonological precedence. I propose, following Raimy (2000a), that phonological precedence can be non-linear. That is, a single segment may immediately precede or immediately follow more than one segment. Once non-linear precedence is allowed, however, we should ask what sorts of non-linear representations are allowed by universal grammar. I argue for a set of insertion points for precedence relations along the lines of Yu's (to appear; 2003) treatment of infixation.[3] This proposal allows us to develop a full typology of possible non-linear representations, and, more specifically, a typology of possible affix types. This typology will include prefixes, suffixes, infixes, and reduplication, thus placing the latter type of morphology back in the realm of affixation, a departure from the line of thinking initiated by Steriade (1988), in which reduplication is the result of a separate (set of) mechanism(s).

In section 2, I provide a brief introduction to non-linear precedence. Section 3 contains a formal typology of affix types. In section 4 I address remaining concerns regarding the power of the model: (a) does a system of idiosyncratic insertion points capture the predictable aspects of morpheme shape and location?, (b) is non-linear precedence still too weak? I conclude that the idiosyncracy allowed by the current system is necessary given the range of observed patterns of affixation (necessarily a subset of all *possible* patterns), but that there remains a possibility that the resulting system is also too strong.

---

[2] This assumption is the root of one possible weakness of the theory, which I address at the end. Namely, the complete divorce of syntax from the determination of morpheme location may be too strong a move.

[3] Raimy (2004) also embraces Yu's work and proposes an adoption of his system of anchor points. In this way Raimy (2004) closely mirrors the proposals in the present paper.

2. An Explicit Representation of Precedence

Raimy (2000a:12) observes that, "there are non-trivial and non-derivable ordering relationships between segments in phonology." These relationships are non-trivial in at least two ways: (a) there are no palindromic or anagrammatic languages where, e.g., [kæt]=[tæk], and (b) phonological rules, processes, and constraints must have access to ordering information. A rule like `A⇒B/C__D` can only apply if the information is available that `C` immediately precedes `A` and `A` immediately precedes `D`. Crucial Optimality Theoretic (OT) constraints like LINEARITY and CONTIGUITY also make explicit use of immediate precedence information. For example, the definition of LINEARITY in McCarthy and Prince (1995) includes '<' as the 'precedes' relation ($\Re$ is the correspondence relation).

(2) LINEARITY (McCarthy and Prince 1995)
S_1 is consistent with the precedence structure of S_2, and vice versa.
Let $x,y \in S_1$ and $x',y' \in S_2$. If $x \Re x'$ and $y \Re y'$, then $x<y$ iff $\neg(y'<x')$.

Raimy (2000a) proposes that precedence relations be explicitly represented in phonology: A→B is read 'A immediately precedes B.' In this system, (3a) and (3b) both represent what is traditionally represented as [kæt].[4]

(3) a. START → k → æ → t → END
b. END ← t ← æ ← k ← START

The beginning (START) and end (END) junctures are needed here in order to be fully explicit about order. These are the only symbols that are not both immediately preceded and immediately followed by at least one segment. If these symbols were excluded, some other convention would have to be adopted to determine the first and last segment in a precedence representation. START and END are further motivated by the need to refer to the first and last segment of a form (for total reduplication, apocope, etc.), as well as for the computation of *completeness* and *economy* in linearization (discussed below).

Since, in most theories of phonological representation, precedence relations are implicitly present as some sort of ordering convention (usually left-to-right), this enrichment of the symbolic vocabulary is not an enrichment of the theory. That is, the simple addition of explicitly represented immediate precedence relations does not in itself make the theory more powerful. Rather, it allows us to make observations and ask ques-

[4] For the moment, immediate precedence relations only seem motivated between X-slots (or a CV skeleton) on the timing tier. Thus the representations in (3) are shorthand for (i).

(i)
$$\begin{array}{ccccccccc} \text{START} & \rightarrow & X & \rightarrow & X & \rightarrow & X & \rightarrow & \text{END} \\ & & | & & | & & | & & \\ & & k & & æ & & t & & \end{array}$$

tions that would not have been obvious otherwise. For example, it now becomes clear that the total precedence relation for phonological representations (the set of all pairs <A,B> such that A precedes B) has been assumed to be transitive, irreflexive, and asymmetric (that is, linear): at all levels of representation a segment or timing slot immediately precedes one segment and immediately follows one segment. Representations in which a segment both precedes and follows another are generally avoided without much thought to the consequences of such representations.[5] And yet, though this assumption may be justified for the output of phonology at the interface with articulatory-perceptual systems (cf. Chomsky's (1995) *bare output conditions*), Raimy's (2000a) novel move is to ask whether this assumption is necessary for levels of grammar further removed from the surface. That is, though at the surface only linear representations like (4a) (from Ilokano, Kenstowicz 1994:624) are legible to extra-linguistic systems, might it be the case that, within phonology, nonlinear representations like (4b) are allowed?

(4) a. START → k → a → l → d → í → ŋ → END 'goat'
b. START → k → a → l → d → í → ŋ → END
(Mapped to: *kal-kaldíŋ* 'goats')

While precedence in (4a) is linear, the set of precedence relations in (4b) is not. Some segments are immediately followed or immediately preceded by more than one segment: *k* immediately follows both START and *l*, while *l* immediately precedes both *d* and *k*. There is nothing logically problematic in such a representation; relations may have a variety of properties at various levels of representation. Transitive precedence may be reflexive and symmetric at one level (where *k* may transitively precede itself, and *k* and *l* each transitively precede the other) and irreflexive and asymmetric at another level.

Non-linear representations such as (4b) are mapped to linear representations that include repetition of phonological material. Under this mapping, which I will call *linearization*, (4b) is mapped to *kal-kaldíŋ* 'goats'. Some of the empirical motivation for this treatment of reduplication can be found in Raimy (2000a,b), Fitzpatrick (2004), Fitzpatrick and Nevins (to appear, 2004), and Raimy and Idsardi (1997). A computational implementation of the model, as well as a formal learning model, has been developed in current work by the present author with Andrew Nevins, Galen Packard, and Aaron Iba (see Nevins and Iba (2004) for an early version of the learner).

[5] Some important exceptions can be found in the tradition of autosegmental phonology, which uses representations that deviate significantly from simple strings of symbols. For example, Mester (1988) presents a non-linear approach to ordering in reduplication.

2.1. Linearization

By hypothesis, a non-linear representation such as (4b) must be *linearized* if it is to be legible to sensory-motor systems.[6] *Linearization* is simply the reconcatenation of a non-linear representation (i.e., one containing "loops") as a linear representation. That is, though the input to linearization can be non-linear, the output is not.

The linearization procedure is guided by three principles: Completeness, Economy, and Shortest. *Completeness* ensures that the segments and relations in the input are maximally spelled out in the output. It does this in a local manner; at each segment, backward-pointing links are chosen over forward-pointing links. Whether a link points backward or forward is determined by examination of the output tape—that is, that part of the representation that has already been linearized.

Given local completeness, (4b), repeated as (5a), is linearized as (5b). Linearization begins at the START symbol and proceeds forward along immediate precedence relations. At *l*, we find two possible ways forward: l→d and l→k. Completeness ensures that the backward-pointing link to *k* is followed, rather than the link to *l*. From here, linearization proceeds to the end symbol (END).

(5) a. START → k → a → l → d → í → ŋ → END

b. START→k→a→l→k→a→l→d→í→ŋ→END 'goats'

Informally, *Economy* simply says "do no more than is necessary." This condition legislates against gratuitous output that is not necessary for satisfaction of the other conditions by biasing linearization towards as-yet untraversed links. For example, (4a) is not spelled out as (6) since this is the result of unnecessarily spelling out the l→k loop multiple times. One traversal of l→k is enough to satisfy Completeness.

(6) START→k→a→l→k→a→l→k→a→l→d→í→ŋ→END

This interplay between Completeness and Economy can be visualized in a tableau (7). Here hyphens stand for immediate precedence relations. "Violation marks" are listed under Economy for segments and relations that have been spelled out more than once.[7]

(7)

	Input: (4a)	COMPLETENESS	ECONOMY
☞	kal-kaldìŋ		k-a-l
	kaldìŋ	(l-k)!	
	kal-kal-kaldìŋ		k-a-l-k!-a-l

[6] This hypothesis is not logically necessary. Representations like (4b) might be legible to sensory motor systems and be "linearized" at a later stage (e.g., when pronounced).

[7] By hypothesis, these linearization constraints cannot be reranked to produce a factorial typology, a key feature of OT constraints. Presentation in tableau form is purely for the sake of exposition.

Finally, in the case of multiple "loops," as in (8) (from Lushootseed (Bates et al. 1994; Broselow 1983; Urbanczyk 2001; Fitzpatrick 2004)), *Shortest* dictates that, if a given segment has multiple backward-pointing relations, inside loops are spelled out before outside loops. In (8), two backward-pointing relations leave *l*. Linearization could traverse either of these links first, since both satisfy Completeness and Economy: both relations are backward-pointing and neither has been traversed. Shortest ensures that the inside loop l→a is followed first.

(8) START→b→a→l→i→END UR for [balalbali] 'forgetting'

Again, the output tape is used to determine length of a loop; l→a and l→b are compared for the number of segments that intervene between the target of the relation (*a* and *b*, respectively) and the source (*l* for both). While one segment, *a*, intervenes between *b* and *l*, none intervenes between *a* and *l*, so l→a is shorter. The full interplay of the three principles in the linearization of (8) is shown in tableau (9).

(9)

Input: (9)	Shortest	Completeness	Economy
☞ bal-al-bali			a-lb-a-l
bal-bal-ali	*!		b-a-la-l
bali		(l-a)!(l-b)	
bal-ali		(l-b)!	a-l
bal-al-al-bali			a-l-a-lb-!a-l

Given the extensions to the system that I propose below, an addition or change must be made to this linearization scheme. In a representation such as (10) (from Laotian Katu, discussed below), no linearization principle can chose between [katas] and [kartas]. That is, no principle forces the spell-out of [r].

(10)
r
START→k→a→t→a→s→END Linearized as: ?

(11)

	Input: (11)	Shortest	Completeness	Economy
a.	☞ katas			
b.	☞ ka-r-tas			

Intuitively, *katas* does not completely spell out the precedence structure in (10). In prior work I have treated Completeness in a local manner, as suggested above: At each segment, as-yet untraversed backward-pointing links are chosen over forward-pointing links. But if representations like (10) are allowed, several possibilities arise.

First, one might add to Completeness, Economy, and Shortest a principle that legislates in favor of early affix spell-out. Call this principle *Non-Stem.* Given a segment X that precedes multiple segments, if one relation links the segment to affixal material (X→affix), and another relation links the segment stem material (X→stem), the affixal link (X→affix) is followed first. An alternative to the use of morphological information would be the adoption of a global form of Completeness, in which different possible linearizations of a form are compared. This form of completeness would chose (11b) over (11a) since the latter fails to include structure that is included in the former. The local version is computationally more tractable, and is used in computational implementations by the present author with A. Nevins, G. Pickard, and A. Iba. However, I will adopt globel Completeness here since its allows for cleaner exposition. I will leave open the possibility, however, that one could find empirical motivation for one approach over the other.

It should be noted that linearization is not simply a patch necessitated by the treatment of reduplication as the result of looping precedence relations. On the contrary, some linearization procedure is implicit in most treatments of affixation. This is particularly clear for infixation, where a stem and affix must be reconcatenated in such a way that the affix finds its way within the original stem. Therefore, this approach can truly claim to contain no reduplication-specific mechanisms (*pace* Downing 2001).

3. A Typology of Affix Types

Though perhaps best known in the context of a derivational theory of reduplication (Raimy 2000b), this explicit, non-linear approach to precedence is neither limited to reduplication (as is clear from Raimy's (2000a) analysis of infixation), nor inherently derivational (as is clear from, e.g., the analysis of Lushootseed multiple reduplication in Fitzpatrick (2004)). Rather, it simply provides a new, more explicit representational vocabulary that can accommodate many types of morphological combinations. In fact, once precedence is made explicit in this way, a typology of formally possible affixation types immediately arises that closely resembles the types of morphology observed in human languages.

In a realizational theory of morphology such as Distributed Morphology (DM), syntactic objects (heads/features/feature structures) are put together in the syntax to create hierarchical structure. These objects are then *spelled out*; that is, they are given phonological content through *vocabulary insertion.* It is important to note that, though the structural location of a piece of syntax has some effect on the position of the vocabulary item (e.g., how close to the root it appears), there is no necessary relation between linear position (e.g., preceding/following a stem) and hierarchical order, at least under the assumptions I am making here. This is especially clear if syntactic structures themselves are unordered, as they are if one assumes what Embick and Noyer (2001) call the *Late Linearization Hypothesis* (12).

(12) The elements of a phrase marker are linearized at Vocabulary Insertion. (Embick and Noyer 2001:562)

For example, assume the root in (13) is spelled out as X. *A priori*, B, which is the next thing to be given phonological exponence, could be spelled out as a prefix or a suffix (that is, it could be placed to the left or right of the stem), giving bX or Xb. The same is true of A; there are then four possibilities for the spell-out of (13): abX, aXb, bXa, Xba.

(13) [A [root B]]

But languages do not usually allow such variation. Therefore, for any affix, the language-knower must store its position of insertion.[8] Stating this explicitly using the precedence notation developed above, a prefix follows START and precedes the initial segment. Thus a prefix like *un-* attaches to *able* as in (14).

(14)

ʌ → n

START→ ej →b→l→END

Linearized as:
START→ʌ→n→ej→b→l→END

Suffixes arise when an affix follows the final segment and precedes the END symbol. The pre-linearization representation of the suffix *-ful* attached to *joy* would be (15).

(15)

f→ ʊ →l

START→dʒ→o→j→END

Linearized as:
START→dʒ→o→j→f→ʊ→l→END

Infixes, which have generally been a nuisance to morpho-phonologists, are not a blemish in this system. Rather, they are exactly like prefixes and suffixes except that they are inserted with segmental material on both sides, rather than preceding or following the START or END symbol. In Laotian Katu, a verb is nominalized by the insertion of an *-r-* after the first syllable nucleus and before the following segment (e.g., *katas* 'to name', *kartas* 'name'; data from Yu to appear).

(16)

r

START→k→a→t→a→s→END

Linearized as:
START→k→a→r→t→a→s→END

Thus far we have seen material inserted following some element X and preceding some element Y, where X precedes Y. But what if Y precedes X? This is precisely where reduplication arises. If segmental material is included in the affix, so-called 'fixed-segment' reduplication is the result, as in Hindi echo reduplication, where *paanii* 'water' becomes *paanii-vaanii*

[8]This is true even in gradient alignment approaches to infixation (McCarthy and Prince 1995), where an affix's position is encoded in an affix-specific alignment constraint.

'water and such things'. Here *-v-* immediately follows the last segment and immediately precedes the first nucleus.

(17) START→p→a:→n→i:→END Linearized as: pa:ni:va:ni:

v

If no segmental material is specified, or a 'null affix' is involved, more "standard" reduplication, such as we saw in Ilokano *kal-kaldíŋ* arises (4) (repeated as (18)).

(18) START → k → a → l → d → í → ŋ → END

Reduplication arises when a "looping" precedence relation is added to a representation. Under this view, substrings labelled "base" and "reduplicant" play no role in the phonological computation. There is no Red morpheme to be aligned or given phonological content, only the introduction of a new precedence relation to a segment, which, upon linearization, gives the appearance of distinct copies. These affix types are summarized in Table 1.

(19) **Table 1**

	Follows	Precedes	Seg. Content?
Prefix	START	Initial Seg.	Yes
Suffix	Final Seg.	END	Yes
Infix	Seg. X	Seg. Y (X→Y)	Yes
Reduplication (Fixed Seg.)	Seg. X	Seg. Y (Y precedes X)	Yes
Reduplication	Seg. X	Seg. Y (Y precedes X)	No

It is plain to see that, when viewed as segmental material and precedence instructions, prefixes, suffixes, and infixes are all essentially the same. They are all affixes with a specified segmental content that immediately follow an element X in a stem and immediately precede an element Y, such that X→Y. They differ only in the identity of X and Y.[9]

The various known types of reduplication arise when the order of X (the immediately preceding element) and Y (the immediately following element) is reversed: Y precedes X, possibly non-immediately. If, in addition to a new precedence relation, segmental material is specified, fixed segment reduplication arises. Different shapes of reduplicant arise depending on where the back-pointing precedence relation is added. This loop is maximally from the final to the initial segment. After linearization, such a representation leads to total reduplication, as in Indonesian *buku* 'book'

[9]Other logical possibilities, such as prefixes, suffixes, and infixes that lack segmental content are not included here as examples since they are string-vacuous.

⇒ *buku-buku* ‘books’ (20). A shorter loop will lead to partial reduplication (that is, repetition of only part of a base form).

(20) START → b → u → k → u → END

The only remaining logical possibilities given the pieces we can manipulate are *syncope* and *templatic morphology.*

(21) Table 1b

	Follows	Precedes	Seg. Content?
Syncope/ Jump Links	Seg. X	Seg. Y (X precedes Y but not X→Y)	No
Overwriting(?)	Seg. X	Seg. Y (X precedes Y but not X→Y)	Yes

Syncope can result if an immediate precedence relation is inserted between segments X and Y, where X precedes Y, but does not *immediately* precede Y. One can see that one way of representing vowel syncope of *memory* /mɛmoɹi/ to [mɛmɹi] in some dialects is as in (22).

(22) START→m→ɛ→m→o→ɹ→i→END

It is not clear to me that this is how syncope should in fact be treated.[10] However, Halle (2004) presents arguments that what he calls partial reduplication is the result of repetition of a string, with concomitant “skipping” of material at one or the other edge. Using the present notation, such a representation appears as in (23a) (linearized as [BCDABCD])or (23b) (linearized as [ABCABCD]), depending on which edge is skipped. While Halle couches his analysis in a typographically different scheme, the two systems are, I believe, formally equivalent.[11]

(23) a. START→A→B→C→D→END

b. START→A→B→C→D→END

[10] In fact, linearization of these forms could be quite problematic under the assumptions made above, which would produce forms without syncope if no further assumptions are made.

[11] Halle proposes that these “jump links” are limited to the edges of reduplicated material, a proposal that he encodes in his development of repetition and skipping junctures. These assumptions could also be adopted in the present system, though only, as far as I can tell, by fiat, rather than by adopting a more limited symbolic vocabulary.

Finally, certain types of templatic morphology could perhaps be treated as infixation, perhaps mixed with "overwriting" of a substring. The latter could be accommodated if an infix immediately follows X and immediately precedes Y, where X precedes Y but does not *immediately* precede it. I will not discuss either of these options further since I have little to say about them, nor particular reasons for proposing that they should or should not be treated within the precedence scheme developed here.

I have shown that making precedence explicit results in a new look for concatenative and non-concatenative morphology. In effect, non-linear precedence representations allow many apparently disparate morphological phenomena to be viewed as different aspects of the exact same system. In this system, precedence relations are the bread and butter of morpho-phonology. When one investigates the range of logically possible simple precedence structures we find a typology of familiar morpho-phonological processes.

Though this unification of morpheme types is conceptually attractive, one would naturally ask for empirical justification for these precedence representations. The treatment of prefixes and suffixes as in (14) and (15) does not, in the end, differ markedly from standard concatenative approaches to affixation. Thus we can expect little, if any, empirical difference. However, this treatment of infixation and reduplication does offer several advantages over previous accounts. I refer the reader to Raimy (2000a,b); Fitzpatrick (2004); Fitzpatrick and Nevins (to appear, 2004); Raimy and Idsardi (1997) for evidence regarding reduplication, and to Raimy (2000a) for evidence regarding infixation. Readers might also wonder what the limits of this system are. For example, could a morpheme call for the addition of a link between the fifth obstruent and the first high vowel in a stem? In the next section I will present a list of simple phonological landmarks that are used in the insertion of new phonological material by the morphology.

4. Insertion Points and the Power of the Model

The typology of non-linear representations presented above was developed with an "anything goes" approach. That is, we sought an answer to the question *What sorts of representations result from the addition of a forward- or backward-pointing loop, with or without segmental material, to a linear stem?* This question yields a surprisingly small and familiar set of answers. However, many issues remain. In the preceding section, I played fairly fast and loose with the insertion of phonological material. Non-linear representations were discussed above with little attention paid to the mechanism of insertion or the set of possible insertion points. In the following section, I will develop a theory of possible insertion point based on Yu (to appear, 2003). I then discuss the difference between the current approach, which holds that there is a high level of idiosyncrasy in morphology, and approaches with the enviable goal of leaving most of the morphology-phonology interface up to general principles and constraints.

4.1. Phonological Landmarks

As with any theory that removes stipulations on representational possibilities, one can ask where the limits are of the new system. Downing (2001) asks this question while discussing Raimy's (2000a) analysis of Axininca Campa, which includes an explicit reference to the initial consonant of a stem:

> "What is unexplained is why the first consonant is targeted rather than some other position in the base string. And if this position can be stipulated, what prevents writing a looping rule that stipulates that copy begins with the second consonant or any other arbitrary position in the string?"

Later, Downing asks "the question of what this theory *cannot* do" (Downing 2001, emphasis not in original). The worry here is that the exponence of a morpheme would be allowed to be something like: "Insert a precedence relation between the fifth vowel and the second obstruent in the stem." Of course this is the same important question as one would raise for any theory. For example, in alignment-based OT, why can one not write a constraint that says: "do not be to the right of four or more consonants," which would force infixation of a suffix when attached to stems with four or more consonants? The answer here, as with all theories, is that, based on what they think human language allows, linguists place limits on the types of rules and constraints they will consider. For example, it is generally accepted that language does not count in most realms, and when it does (e.g., in stress calculations), not past two or three. These sorts of general expectations are shared by most linguists, and stand in the background in almost all theorizing. This means that, while important, the question of what constrains the present model is not necessarily more or less urgent in this case than in any other.

This paper is couched in an essentially *realizational* theory of morphology. That is, morphemes are put together in the syntax, and during the computation to PF these formal feature structures are exchanged for phonological material. Thus in (24) the root will be spelled out first, giving some phonological material.

(24) [A [root B]]

Next, the affixes will be spelled out, first B, then A, in a way that is sensitive to surrounding morpho-syntactic features and phonology. Precedence relations are part of the phonological information necessary for spelling out morphemes, and one would like to know where additional precedence relations can be added to a stem. As with most other linguists, we assume that arbitrary counting is not within the realm of human language. Rather, we assume a short set of possible insertion points. These are based on the set of points found by Yu (2003) to be used cross-linguistically in infixation. Yu calls these points of infixal insertion "pivots."

(25) a. Edge Pivots
 (i) First Consonant (C_1)
 (ii) First Vowel (V_1)
 (iii) Final Syllable (σ_f)
 (iv) Final Vowel (V_f)
 b. Prominence Pivots
 (i) Stressed Syllable (Str-σ)
 (ii) Stressed Vowel (Str-V)

For Yu, infixes (including what he calls "reduplicative infixes") can appear on either side of these pivots. For example, Atayal animate focus marker *-m-* appears after the first consonant for the stem, regardless of whether it is part of a consonant cluster.

(26) qul ⇒ q**m**ul 'snatch'
 hŋuʔ ⇒ h**m**ŋuʔ

However, Yu's system cannot be assimilated *as-is* to the precedence-based system above because he essentially assumes that statements about affix position are *surface true*. Though precedence, and therefore *before* and *after*, are crucial notions here, Yu makes different use of them. For Yu, an infix appears before or after a pivot in a surface form. In the scheme developed above for vocabulary items, morphemes are specified for precedence relations in the lexicon. These specifications tell the grammar where to insert their phonological material, but there is no guarantee that the lexical information will be transparently present on the surface. This is particularly clear for reduplication. When treated as a looping precedence relation, there is no sense in which a reduplicated form contains a base and a reduplicant on the surface. There is simply no way to point to one or the other. Therefore Yu's list of pivots and the relations *before* and *after* cannot be adopted wholesale as statements about surface forms. However, a slight modification rectifies the problem. I propose the following list of *anchor points.*

(27) a. Initial X
 (i) First Consonant (C_1)
 (ii) First Vowel (V_1)
 (iii) First Segment (Seg_1)
 (iv) First Syllable (σ_1)
 (v) First Foot (Ft_1)
 b. Final X
 (i) Final Consonant (C_f)
 (ii) Final Vowel (V_f)
 (iii) Final Segment (Seg_f)

(iv) Final Syllable (σ_f)

c. Prominence Points

(i) Stressed Syllable (Str-σ)

(ii) Stressed Vowel (Str-V)

(iii) Stressed Foot (Str-Ft)

This list can be broken down into the primitives *initial, final, consonant, vowel, segment, syllable*, and *foot*. Beyond these, I propose that the primitive relations *immediately precedes* (ImPrec) and *immediately follows* (ImFol) can be used in the statement of where an affix is attached. The idea is that human language makes reference to a certain finite set of points, and that these recur in phonological rules and constraints as well as morphological insertion. These are the only positional primitives that can appear in the description of the position of a morpheme in a stem.

I argue below that the phonological exponent of a morpheme needs to be idiosyncratically recorded in the lexicon. For affixes and reduplication (now considered part of the same system), the phonological exponent has the following form:

(28) PREC→AFF→FOL

a. PREC: Segment that immediately precedes the affix

b. AFF: Segmental material (possibly null)

c. FOL: Segment immediately following the affix

Here PREC and FOL are taken from the list of possible anchor points, plus the non-recursive functions ImPrec() and ImFol(), which give the segment that immediately precedes or immediately follows their argument, respectively. These anchor points create a typology of affixes and reduplication, as shown in Table 2 with examples of infixation and Table 3 with examples of reduplication (most data from Yu (2003)).

(29) Table 2

$C_1 \rightarrow$ m $\rightarrow$ ImFol(C_1)	Atayal *animate actor focus*
	qul ⇒ qmul
	hŋuʔ ⇒ hmŋuʔ
ImPrec(V_1) $\rightarrow$ um $\rightarrow$ V_1	Chamorro *verbalizer, actor focus*
	gupu ⇒ gumupu
	tristi ⇒ trumisti
	epanglo ⇒ umepanglo
$V_1 \rightarrow$ i $\rightarrow$ ImFol(V_1)	Miskito *2nd conj 1st pers*
	na.pa ⇒ naipa
	kak.ma ⇒ kaikma

(30) Table 3[12]

$V_1 \rightarrow C_1$	Lushootseed *diminutive*
	bali ⇒ ba-bali
$\mathrm{ImFol}(V_1) \rightarrow C_1$	Lushootseed *distributive*
	bali ⇒ bal-bali
$\mathrm{Seg}_f \rightarrow C_1$	Kinande *pluractional*
	-ohera ⇒ o-hera-hera
$\mathrm{ImFol}(C_1) \rightarrow C_1$	Pangasinan *plural*
	niog ⇒ ni-niog
	amigo ⇒ a-mi-migo
$\mathrm{ImFol}(\text{Str-V}) \rightarrow \text{Str-V}$	Lushootseed *out-of-control*
	bali ⇒ bal-ali
$\mathrm{Seg}_f \rightarrow \mathrm{Seg}_1$	Indonesian *plural*
	buku ⇒ buku-buku
$\mathrm{ImPrec}(V_1) \rightarrow a \rightarrow \mathrm{ImPrec}(V_1)$	Ferhan Tetan *nominalization*
	beik ⇒ babeik
	brakat ⇒ krarakat
$\mathrm{ImPrec}(C_f) \rightarrow (X) \rightarrow V_1$	Takelma *frequentative*
	hemg- ⇒ hememg-
	baxm- ⇒ baxaaxm-
$V_f \rightarrow \mathrm{ImFol}(V_1)$	Amis *plural*
	ɬaŋka ⇒ ɬaŋkaŋka
	ʃamaɬu ⇒ ʃamaɬamaɬu
$V_f \rightarrow \mathrm{ImPrec}(V_f)$	Ineseño Chumash
	taSuSun
	iwawan
	oxyoyon
$\text{Str-}\sigma \rightarrow \text{Str-}\sigma$	Samoan *plural*
	'toa ⇒ to'toa
	a:'vaga ⇒ a:va'vaga

The prosodic morphologist might protest to the treatment of infixation and reduplication in this way. In particular, this approach states that all positional properties of an affix are stored idiosyncratically. I will now provide support for this claim.

4.2. Idiosyncracy

The approach outlined above allows for only a limited range of primitives to be used in the statement of the position of an affix. Yet it assumes that positional information is stored idiosyncratically for each affix. This differs from recent work that seeks to derive the position and shape of certain affixes, especially reduplicative affixes, from more general constraints. I argue in this section that idiosyncratic statements regarding the position of affixes is necessary.

[12] (X) is a timing slot, which is inserted only in certain cases.

In the process of theory falsification, linguistic theories are either accused of over- or under-generation (or both). However, two types of "overgeneration" claims are made. One is the run-of-the-mill verifiable kind: "theory T predicts that X is grammatical in language L, but X is not grammatical in L." The other sort is the following: "theory T predicts that a language of type L is possible, but I don't think L is possible." Though linguists have indeed developed certain intuitions about what types of languages are possible, these intuitions are clearly fallible, and not generally verifiable or falsifiable. That is, no matter how firmly we believe we understand how language works, it is difficult to prove that something is impossible in human language.[13] This is not to say we should not continue to make highly constrained theories. Only highly constrained theories are easily falsifiable, and thus they lead to progress. However, it is clear that an argument from a linguist's intuition about what is possible in language is not a knock-down argument against a theory.

I will discuss here the case of reduplication. Different researchers have claimed that only certain types of reduplication are possible: those that show "emergent unmarkedness". I will argue that these intuitions were wrong, and that reduplication can show very strange patterns indeed. Though they are surely not limitless, I argue that the idiosyncratic shape of each reduplication pattern must be recorded in the lexicon. The limits of this idiosyncrasy can be found in the possible combinations of anchor points and segmental material.

Generality Level 1: TETU

Any theoretician will seek maximum simplicity in a theory while maintaining empirical adequacy. It is not surprising, then, to see efforts to account for the placement, size, and shape of reduplication patterns through constraints or processes that find independent motivation elsewhere and a language.

For example, Spaelti (1997) assumes a Correspondence-Theoretic account of reduplication and claims that, "all cases where reduplication shows deviation from identity [i.e., complete and exact repetition – JMF] are the result of Emergence of the Unmarked" (p. 30).

Correspondence Theory claims that the shape and content of the reduplicant, which is just a phonologically empty RED morpheme in the input, is the result of the ranking of FAITH-BR with respect to phonotactic and (other) alignment constraints. If Spaelti's hypothesis is correct, none of these constraints appeal directly to reduplicants or even specifically to shapes or templates (e.g., foot, syllable, etc.). No constraints says "$\text{RED}_i =$ Foot" or even "X = Foot" (for some X). Rather, shape restrictions "follow from general constraints on prosodic structure" (Spaelti 1997:34).

In support of this contention, Spaelti offers the case of Boumaa

[13] One exciting possibility for showing something is linguistically impossible is raised by work with the language savant Christopher, as reported in Smith and Tsimpli (1994).

Fijian (Dixon 1988). Boumaa Fijian shows right-to-left bimoraic footing. In Spaelti's analysis, this is insured by a constraint:

(31) ALIGN(Foot, R, PrWd, R)[14] (= ALLFEETRIGHT)

ALLFEETRIGHT has the effect that feet stack up as close to the right as possible. Spaelti then adds a single RED-specific constraint:

(32) ALIGN(RED, L, Foot, L)

Though this is a deviation from the complete generality claimed above, let us grant this much morpheme-idiosyncrasy. The combination of these constraints with FAITH-IO and FAITH-BR result in the correct duplicate shape: a single bimoraic foot.[15]

(33)

/RED + talanoa/	Align (32)	ALLFTR	MAX-BR
☞ **(tala)**(tala)(noa)		$\sigma/\sigma\sigma\sigma$	noa
ta(lano)(tala)(noa)	*!	$\sigma/\sigma\sigma\sigma$	a
(tala)(noa)(tala)(noa)		$\sigma/\sigma\sigma\sigma/\sigma!\ \sigma\sigma\sigma$	
(tala)noa(tala)(noa)		$\sigma/\sigma\sigma\sigma\sigma!$	

And yet though this constraint ranking produces the correct output for /RED + talanoa/, it fails for /RED + butaʔo/. As shown below, this constraint ranking picks out *bubutaʔo*, while the attested form is *butabutaʔo*. Spaelti's analysis predicts that Boumaa Fijian reduplication should vary from a bimoraic foot to a single light syllable, depending on the size of the base.

(34)

/RED + butaʔo/	Align (32)	ALLFTR	MAX-BR
☞ (**bu**bu)(taʔo)		$\sigma\sigma$	taʔo
☜ **(buta)**bu(taʔo)	*!	$\sigma/\sigma\sigma\sigma$	a

Spaelti (1997) fixes this hole in the analysis by positing a requirement that the base itself must constitute a prosodic word, which is minimally bimoraic. In the end, of course, this is a question of fact, and Spaelti provides no evidence for this assertion. But even were it true, the conspiracy of these constraints does nothing other than restate the fact: this reduplication pattern repeats the initial bimoraic foot of the stem. Any hole in the analysis will be due to a failure to state this fact. Any approach that claims that the shape of a reduplicant is determined not by an idiosyncratic specification of the shape for each morpheme, but by language-general constraints operative elsewhere in the grammar, makes the claim that *reduplicants should*

[14] Alignment constraints are to be read as follows: Align(X,P,Y,Q) = for all X, the P-edge of X is aligned with a Q-edge of Y. E.g., Align(Stem,L,σ,L) = for all stems, the left edge of the stem coincides with the left edge of a syllable.

[15] Left-edge placement is taken for granted, but could be insured with another reduplication-specific alignment constraint that would rule out, e.g., (tala)(noa)(noa).

show variation in their shape as seen above. And yet this does not seem to be the way reduplication works. I argue here that explicit specification of reduplication shape is necessary. Reduplication is the result of a morpheme. Morphemes, by their nature, contain certain idiosyncratic information (e.g., segmental content, shape, etc.). Thus we should expect reduplication patterns to have some degree of idiosyncracy.

Generality Level 2: Generalized Templates

As we have seen, the lofty goal of deriving position and shape of all reduplicants from general constraints runs into problems quickly. In part due to this, and for other reasons, Generalized Template Theory (GT) was developed (Urbanczyk 2001, among others). GT accepts a certain level of idiosyncracy and claims that "the phonological exponence of a reduplicative morpheme is derivable from its morphological classification" (Urbanczyk 2001:11). That is, reduplicative morphemes are segmentally-empty and are specified for morpheme type. Reduplicants can be roots or affixes and prefixes or suffixes. There are no constraints like RED=Foot or RED=σ.[16] Instead, A morpheme is specified as, for example, {RED, root, prefix}. This means that this particular morpheme is sensitive to root-specific constraints. The prefix part of the specification is a left-alignment constraint.

This approach is useful in analysis of languages with multiple patterns of reduplication. However, it is plain to see that this particular version of GT allows only four reduplication patterns.

(35)

	Prefix	Suffix
Root	Rt, Prefix	Rt, Suffix
Affix	Aff, Prefix	Aff, Suffix

This predicts that a language can have at most four types of reduplication, though others might be added by allowing some fixed segmentism to the mix. More importantly, however, it predicts that a language can have at most one of each type. This prediction should be easy to falsify. For example, a language with two or more types of root reduplication of the prefixal sort could not be modelled in GT. More basically, we should ask whether we think human language is limited, by definition, to four types of reduplication.

While a full cross-linguistic study of languages with multiple reduplication patterns could turn up a counterexample to the four pattern maximum, the data in the following section brings other questions to the fore. I argue next that full idiosyncracy is in fact required in some cases, and so GT and other systems that seek to derive reduplicant shape and content from general principles are too weak.

[16]These constraints lead to well-known typological problems.

Generality Level 3: Idiosyncracy

In this section I argue that languages do at times contain very strange patterns of reduplication that cannot be described with surface-oriented statements. I focus on a pattern of reduplication in Thao[17] and show that it requires idiosyncratic statements in the lexicon.

Chang (1998) provides an analysis of Thao reduplication patterns within Correspondence-Theoretic OT and points out several difficult problems for such an approach. Chang first discusses so-called "Full Reduplication," which appears as the repetition of the full stem, save any final consonant.

(36)

STEM	RED FORM	GLOSS
kari	k-m-ari-kari	'to dig up/ repeatedly, habitually
ribuq	m-ribu-ribuq	'stir/ continuously'
psaq	ma-psa-psaq	'to kick/ repeatedly'
łun	mak-łułun	'mucus/ keep blowing the nose'
zai	m-in-zai-zai	'advice/ say repeatedly'
tup	kan-tu-tup	'follow/ persistently'
du	maš-du-du	'to pass s.t. along'

This appears to be a case of full reduplication (maximal satisfaction of MAX-BR) that obeys NOCODA.[18]

(37)

		/RED+ribuq/	NOCODA	MAX-BR
☞	a.	ribu-ribuq	*	w
	b.	ribuq-ribuq	**!	

However, as with the Fijian example above, this is not quite right. This analysis predicts variation of the shape of the reduplicant depending on the shape of the stem. Consonant-initial stems like *ribuq* will not copy the final coda, while vowel initial stems like *acan* will.

(38)

		/RED+acan/	NOCODA	MAX-BR
☞	a.	acan-acan	*	
☜	b.	acaacan	*	n!

And yet *acan* reduplicates as *aca:can*, not as the predicted **acanacan*. Thus once again appeal to general constraints fails to correctly predict the shape of the reduplicant. Since, according to Chang, the two a's in *ac**a**-**a**can* coalesce into a single heavy syllable nucleus, we do not have the option of stating a constraint like the following, which is parallel to Spaelti's patch on the Fijian case above.

[17] Thao is an Austronesian language spoken around Sun-Moon Lake in central Taiwan.

[18] There are no stem-internal codas in the stems that undergo full reduplication. We will see below that this is crucial.

(39) ALIGN-LEFT(Stem,σ)

Chang decides to solve this problem by encoding the fact that the reduplicant does not contain a consonant corresponding to the coda of the base with a brute-force constraint:

(40) ALIGN-RIGHT(RED, V)

This of course does nothing other than state that this particular reduplicant ends in a vowel.

The picture becomes much more complicated, however, when we turn to "rightward reduplication," which is treated separately by Chang, but seems in fact to be the same pattern caused by the same morpheme. As with "full reduplication," all the forms below show intensity, repetition, and/or habituality, depending on the stem.

(41)	šnara	pa-šnara-nara	'ignite/burn s.t. repeatedly
	kikałi	kikałikałi	'ask/ around'
	qriuʔ	q-un-riu-riuʔ	'steal/ habitually'
	patihaul	matihau-haul	'spell/ cast a spell'
	ag.qtu	agqtu-qtu	'contemplate/ think about'
	ar.faz	m-arfa-rfaz	'fly/ continuously'
	m-ig.kmir	igkmi-kmir-in	'grasp/ roll into a ball'
	bu.qnur	mia-buqnu-qnur	'anger/ be irritable'

Upon seeing this pattern, it is clear that "Full Reduplication" is just the subcase of "Rightward Reduplication" where there is no material left of the string copied by the reduplicant. And yet this additional left-edge material makes an OT Correspondence-Theoretic analysis nearly impossible. Chang presents a candidate analysis including standard constraints like *COMPLEX and NOCODA as well as templatic constraints like RED=$\sigma_\mu\sigma_\mu$[19] and RED=σ_μ, which apply allomorphically depending on the input stem. Furthermore, she finds that more parochial constraints like ALIGN-LEFT(RED,C_1^2) (i.e., align the left edge of RED with at least one and at most two consonants) and BR-INDEX($\leq$4) (i.e., the base-reduplicant correspondence contains at most four segments) are necessary to capture the full range of data.

Clearly the resulting list of constraints and rankings provides little or no insight into the pattern itself. All of this is simply trying to encode an idiosyncratic reduplicant shape that simply won't conform to the expectations developed for phonotactic and alignment constraints elsewhere. The intensive-repetitive morpheme in Thao results in reduplication that repeats a string starting at the final vowel of the stem and going left maximally two consonants and one (more) vowel. In fact, this pattern is very difficult to describe even given the theory developed here, which has been accused of being overly permissive (Downing 2001).

[19]The final version of this is R=$\mu.\mu$ (i.e., the reduplicant is heterosyllabically bimoraic).

Interestingly, this pattern can be described easily in the framework of Marantz (1982): the Thao pattern is a CVCV skeletal affix that is infixed after the final vowel. The string to the left is then copied and segments are associated right-to-left.[20] This results in the patterns below.

(42) CVCV

ɬqizi + ɬ q i z i ⇒ ɬqiziqizi

(43) CVCV

qriu + q r i u (ʔ) ⇒ qriuriuʔ

(44) C VCV

patihau + pati h a u (l) ⇒ patihauhaul

(45) C VCV

agqtu + ag q t u ⇒ agqtuqtu

(46) CVCV

arfa + a r f a (z) ⇒ arfarfaz

(47) C VCV

buqnu + bu q n u (r) ⇒ buqnuqnur

This pattern could be stated in precedence notation only by allowing certain extra anchor points: BEGIN: final vowel, END: left two segments from final vowel. If this segment is a vowel, include the preceding C. This is clearly not pretty, but it does limit the possibility of counting to a small number (2), which is already known to be used elsewhere in the grammar (e.g., footing). If we find that a precedence statement of this sort should not be allowed, however, the case of Thao might lead us to allow both precedence-based "looping" reduplication, and affixation of empty CV structure, as in Marantz (1982).

The point is, however, that this pattern of reduplication is strange enough to defy analysis using "natural" constraints. It forces us to adopt full idiosyncracy for the description of the pattern.

4.3. The Syntax-Morphology Interface

In the last section we saw a pattern of reduplication in Thao that stretches the precedence-based system of morphological spell-out perhaps to the

[20]Recall that association is "phoneme driven." That is, one begins with the segmental tier, seeking a C or V to match a segment, then proceeds to the next segment.

point of breaking. This would suggest that either the system is too weak or it is not the only way that languages can show reduplication. However, the system may also be too strong. I have relied throughout on the assumption that the placement of a morpheme with respect to a stem is somewhat independent of that morpheme's place in the syntax. I assume throughout a type of mirror theory, in which hierarchical relations among morphemes do have an effect on a phonological form, but in principle a morpheme can be spelled out as prefix, suffix, infix, reduplication, etc.

However, this may not be true. One could certainly imagine a theory that ties syntactic position more closely to morphological spell-out. For example, prefixes might have fundamentally different syntax than suffixes.[21] Furthermore, though we find, for example, plurality marked with both prefixes (e.g., Bantu) and suffixes (e.g., Indo-European), it might turn out that even here semantic differences have escaped us and language does not even contain this much freedom of variation. These questions are simply meant to show that the theory developed here could be challenged on many fronts.

5. Conclusions

We started with a fundamental question: why are there prefixes, suffixes, infixes, and reduplication? Why are there these morphological processes and not some other set? I have developed a theory of morphological spell-out based on Raimy's (2000a) notion of immediate precedence relations. This theory, it turns out, when explored to its full extent, results in a typology of formally possible affix types that closely mirrors some of human language's most common morphological processes. I argue that a combination of the formal possibilities allowed in the precedence-based system with the anchor-point theory of Yu (2003) provides the correct inventory of possible affix types. This marriage of theories answers the always-important question of what constrains the system. Finally, I showed that a certain amount of idiosyncracy, such as that allowed in the present theory, is necessary for the description of all attested morphological patterns (a subset of the possible patterns). I ended, however, by questioning whether the theory developed here might not be both too weak and too strong in various ways.

References

Bates, D., Hess, T. and Hilbert, V. (1994). *Lushootseed Dictionary*, University of Washington Press.

Broselow, E. (1983). Subjacency in morphology, *Natural Language and Linguistic Theory* **1**(3): 317–346.

Chang, M. L. (1998). Thao reduplication, *Oceanic Linguistics* **37**(2): 277–297.

[21] It is already well-known that in some languages prefixes and suffixes have markedly different phonology, which could perhaps be linked to syntax.

Chomsky, N. (1995). *The minimalist program*, MIT Press, Cambridge, Mass.
Dixon, R. M. W. (1988). *A Grammar of Boumaa Fijian*, University of Chicago Press.
Downing, L. (2001). Review of Raimy 2000, *Phonology* **18**(3): 445–451.
Embick, D. and Noyer, R. (2001). Movement operations after syntax, *Linguistic Inquiry* **32**: 555–595.
Fitzpatrick, J. (2004). The representation and linearization of non-linear precedence. Ms., MIT.
Fitzpatrick, J. (to appear). A theory of possible affix types, *in* A. Salanova (ed.), *Papers from EVELIN 1*, MIT Working Papers in Linguistics.
Fitzpatrick, J. and Nevins, A. (2004). Linearizing nested and overlapping precedence in multiple reduplication, *The Proceedings of the 27th Penn Linguistics Colloquium*, pp. 75–88.
Fitzpatrick, J. and Nevins, A. (to appear). Phonological occurrences: Relations and copying, *The Proceedings of the 1st and 2nd North American Phonology Conferences*.
Halle, M. (2004). Reduplication. Ms., MIT.
Halle, M. and Marantz, A. (1993). Distributed Morphology and the Pieces of Inflection, *in* K. Hale and S. J. Keyser (eds), *The view from Building 20*, MIT Press, Cambridge, Mass.
Kenstowicz, M. (1994). *Phonology in Generative Grammar*, Blackwell Publishers, Malden, Mass.
Marantz, A. (1982). Re Reduplication, *Linguistic Inquiry* **13**: 483–545.
McCarthy, J. and Prince, A. (1995). Faithfulness and reduplicative identity, *in* J. Beckman, L. W. Dickey and S. Urbanczyk (eds), *Papers in Optimality Theory*, Vol. 18 of *University of Massachusetts Occasional Papers*, pp. 249–384.
Mester, A. (1988). *Studies in Tier Structure*, Garland Outstanding Dissertations.
Nevins, A. and Iba, A. (2004). A selectionist learner for parametrically ambiguous reduplicants, Presented at the 78th LSA.
Raimy, E. (2000a). *The Phonology and Morphology of Reduplication*, Mouton de Gruyter.
Raimy, E. (2000b). Remarks on backcopying, *Linguistic Inquiry* **31**(3): 541–552.
Raimy, E. (2004). Primitives of affixation, *The 3rd North American Phonology Conference*.
Raimy and Idsardi (1997). A minimalist approach to reduplication in OT, *The Proceedings of NELS 27*.
Smith, N. and Tsimpli, I.-M. (1994). *The Mind of a Savant*, Blackwell.
Spaelti, P. (1997). *Dimensions of Variation in Multi-pattern Reduplication*, PhD thesis, University of California, Santa Cruz.
Steriade, D. (1988). Reduplication and syllable transfer in Sanskrit and elsewhere, *Phonology* **5**: 271–314.
Urbanczyk, S. (2001). *Patterns of Reduplication in Lushootseed*, Garland Outstanding Dissertations.
Yu, A. (2003). *The Morphology and Phonology of Infixation*, PhD thesis, University of California, Berkeley.
Yu, A. (to appear). Understanding infixes as infixes, *The Proceedings of the 1st and 2nd North American Phonology Conferences*.

Dept. of Linguistics and Philosophy
MIT, Bldg. 32-D866
77 Massachusetts Ave.

Cambridge, MA 02139
jfitzpat@mit.edu