Phonological Occurrences: Relations and Copying

Justin Fitzpatrick and Andrew Nevins[*]
Massachusetts Institute of Technology

Abstract

We formally and empirically examine the phonological relation *immediate precedence*, its properties, and further relations that can be derived from it. When each *occurrence* of a phonological segment in a non-linear multiprecedence structure is reconcatenated into a linear sequence of copies, *reduplication* results (Raimy 2000a). A number of consequences arise in the representation of affixes, in modelling allomorphy in Hindi and Turkish echo reduplication, and in explaining patterns of multiple reduplication in Lushootseed (Salish). We offer directions for further analysis of reduplication that distinguish conditions on multiprecedence from other wellformedness modules without "the emergence of the unmarked" as a central tenet.

1 The Fundamental Role of Immediate 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 OT Constraints like Linearity and Contiguity also make explicit use of immediate precedence information.

Raimy proposed that these crucial precedence relations be explicitly represented in phonology: A→B is read 'A immediately precedes B.' In this system, (1a) and (1b) both represent what is traditionally represented as [kæt].[1]

(1) 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

*Many thanks to audiences at the 2nd North American Phonology Conference, the 27th Penn Linguistics Colloquium, and participants of the 10th EGG Summer School course "Multiple Introductions: Precedence for Syntacticians, Dominance for Phonologists," where portions of this work have been presented. We have benefited immensely from discussions with Klaus Abels, David Adger, Eugene Buckley, John Frampton, Morris Halle, Aaron Iba, Victor Manfredi, Alec Marantz, Eric Raimy, Charles Reiss, and Bert Vaux. Only the authors are responsible for the facts, opinions and mistakes contained here. Partial support for this work was provided by a National Science Foundation Graduate Research Fellowship.

[1] For the moment, immediate precedence relations only seem motivated between X-slots (or a CV skeleton) on the timing tier. Thus the representations in (1) are shorthand for (i).

(i) START→X→X→X→END
          | | |
          k   æ   t

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 in section 3).

The distinction between precedence and immediate precedence is crucial. *Precedence* is the set of all pairs <A,B> such that A precedes B, either immediately or transitively; *immediate precedence* is the set of all pairs <A,B> such that A→B. Thus for (1), the precedence relation is (2a), while the immediate precedence relation is shown in (2b).

(2) a. *Precedence*: {<START,k>, <START,æ>, <START,t>, <START,END>, <k,æ>, <k,t>,<k,END>, <æ,t>, <æ,END>, <t,END>}
b. *Immediate Precedence*: {<START,k>, <k,æ>, <æ,t>, <t,END>}

As for the question of which of these relations is the *primitive* relation from which the other can be derived, answers are largely conceptual in nature. Immediate precedence seems more basic as it is more "immediate," and phonological processes seem to be overwhelmingly between immediately adjacent segments. The "more general" relation, (transitive) precedence, is readily defined inductively: *a* precedes *b* if *a* immediately precedes *b* or *a* immediately precedes some segment *c* that precedes *b*. The coprecedence and cosuccession relations (discussed in section 2) are also derived from immediate precedence. For concreteness, then, we adopt immediate precedence as the primitive relation. Phonological occurrences are defined in terms of immediate precedence. For a segment X, each of X's immediate precedence relations is an occurrence of X.

Since in most representations precedence relations have always been 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 questions 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). It has often been assumed that 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.[2] 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 (3a) (from Ilokano, Kenstowicz 1994:624) are legible to extra-linguistic systems, might it be the case that, within phonology, nonlinear representations like (3b) are allowed?

(3) 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 (3a) is linear, the set of precedence relations in (3b) 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. By hypothesis, non-linear representations such as (3b) are mapped to linear representations that include repetition of phonological material. Under this mapping (discussed in detail below) (3b) is mapped to *kal-kaldíŋ* 'goats'.

Certain parallels arise between immediate precedence in phonology and *immediate dominance*, the primitive relation in phrase structure. *Syntactic occurrences* are defined in terms of immediate dominance relations

[2]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.

("motherhood"). *Transitive dominance* is a derived relation, as are *co-dominant* ("XPs in the chain")[3] and *co-dominated* ("sisters").[4] Just as a single phonological segment can have multiple occurrences in phonology, syntactic elements can be in multiple positions; that is, they can be dominated by more than one element. For example, at one level of representation, distinct syntactic objects TP and vP may each immediately dominate a DP subject, resulting in *multidominance* (the existence of multiple occurrences of a single syntactic object). At another level of representation (e.g., "pronunciation"), there is no multidominance, and only single occurrences of a given syntactic object exist. The mapping between multidominance structures and "pronunciation" structures thus involves linearization of a sort: multiple occurrences are converted into one or more copies. Another type of mapping applies in the computation to logical form (LF): at LF multiple syntactic occurrences are interpreted separately, one occurrence as an operator, another as a variable.

There is no *a priori* reason why the basic relations are the ones they are, and why they exhibit variance in their formal properties at different "stages" in a linguistic computation (e.g., why sisterhood is symmetric at one level ("Merge") and asymmetric at another (in Fregean LF)). The mapping between these levels of representation is of course, the interesting part. In one syntactic mapping, c-command, a relation at one level, corresponds to precedence among syntactic objects at another, according to the property of asymmetry (Kayne 1994). In the phonological mapping alluded to in (3b), immediate precedence among phonological objects is mapped onto immediate precedence, such that transitive precedence has the property of linearity.

If multiple occurrences are possible in both the syntactic and phonological domains, why, then, do we observe the repetition of phonological material, but not of syntactic constituents? In the phonological mapping, linearization of a representation in which an element (transitively) precedes itself results in reduplication, as in (3b). But syntactic constituents that have been moved/remerged are generally pronounced and interpreted only once. The answer lies in the nature of syntactic merger and projection. Syntactic merger creates hierarchical objects; A and B are put together to create (project) C. Phonological precedence structures, however, are not hierarchical. While there can be "looping" precedence representations, there are no "looping" trees. Since merger of two syntactic objects creates a third object which immediately dominates the first two, the creation of new immediate dominance relations cannot occur between any two syntactic objects already present in a phrase marker. There is no new sisterhood relation without the introduction of a new syntactic object, an intermediary, of which they are co-dominees. Hence there cannot be any way for an element x to transitively dominate a y that transitively dominates x. Therefore, it is possible to map a multidominant syntactic phrase marker to a tree that does not include multidominance, without repetition of syntactic material. The same is not true for multiprecedent phonological structures.

Empirical benefits of the relational model of precedence are found in explanations of over-/under- application in terms of quantification over occurrences (Raimy 2000b), dialect variation in reduplication in terms of ambiguous precedence structures (Nevins and Vaux 2003), formal comparison of the precedence structures in reduplication vs. infixation (Raimy 2000a), and formal specification of inherent reduplication and retriplication (Fitzpatrick and Nevins 2003). In the present paper, we demonstrate that the derived interprecedence relations of coprecedence and cosuccession provide explanations for allomorphy in "avoidance" cases of fixed-segment reduplication (section 2). We illustrate the conditions on linearization and the interaction of multiprecedence with morphological structure and phonological well-formedness in the analysis of Lushootseed multiple reduplication in section 3. The fourth section concludes as we discuss research in the multiprecedence-wellformedness interface that must be done for future analyses of reduplicant "markedness effects". These generalizations are still necessary in models without an output substring labelled "reduplicant" on which particular conditions must hold.

1.1 Vocabulary Insertion and the Specification of New Relations

We assume here a realizational model of phonology along the lines of Halle and Marantz (1993), *et. seq.* In such a theory syntax precedes morphology and builds hierarchical structure containing morpho-syntactic features. These features are morphologically spelled out through vocabulary insertion. That is, a morpheme

[3] If Y immediately dominates X and Z immediately dominates X, than Y and Z are co-dominant.

[4] Further relations may be derived. For example, c-command can be derived in terms of lack of a transitive dominance relation between two objects (Abels 2003).

(morpho-syntactic terminal node) is given a phonological exponent in a cyclic manner, starting with the most deeply embedded morpheme and moving up.

The explicit immediate precedence relations introduced above provide a natural way to express the idiosyncratic form of vocabulary items. A given node might be spelled out as a free morpheme, in which case it is not specified to attach to a stem, or as a bound morpheme, where some attachment site is indicated. In the former case a morpheme will have its own START and END symbols; in the latter, the lexical (vocabulary) entry will specify what immediately precedes and immediately follows the vocabulary item's phonological material. Such vocabulary items have the form in (4).

(4) PREC→MEL→FOLL

- a. PREC: Segment that immediately precedes the affix
- b. MEL: Phonological content/melodic material of the vocabulary item (possibly null)
- c. FOLL: Segment that immediately follows the affix

The contents of PREC and FOLL are taken from a small set of natural phonological anchor points, including the first and last segment, first and last vowel and consonant, certain prominence-based anchor points such as the stressed vowel, and the immediately following segment (ImSucc(Seg), i.e., the immediate successor of Seg) and immediately preceding segment (ImPredec(Seg), i.e., the segment that immediately precedes Seg).[5] In this schema, a prefix would be represented as in (5a) and a suffix as in (5b).

(5)
- a. START → MEL → $\text{Seg}_{initial}$
- b. Seg_{final} → MEL → END

An infix has PREC and FOLL values that are not initial or final. For example, Atayal animate actor focus, as described in Yu (2003:13), has the form in (6).

(6) C_1 → m → ImSucc(C_1)

- a. qul ⇒ q**m**ul
- b. hŋuʔ ⇒ h**m**ŋuʔ

Reduplication fits naturally into this approach. When PREC is transitively preceded by FOLL in the stem to which a morpheme's exponent attaches, linearization (discussed below) leads to the repetition of phonological material. Examples are shown in (7).[6]

(7)
- a. Pangasinan plural (Yu 2003:18ff.): ImSucc(C_1) → C_1
  - (i) niog ⇒ ni-niog
  - (ii) amigo ⇒ a-mi-migo
- b. Indonesian *plural*: Seg_{final} → $\text{Seg}_{initial}$
  - (i) buku ⇒ buku-buku

When (7b) is spelled out with the stem /buku/, (8) is formed. Here there continues to be a single *b*, *u*, *k*, and *u* in the representation. However, *b* is immediately preceded by more than one element. In such a case we say *b* has *multiple occurrences*; here *b* has an occurrence in which it immediately follows START and one in which it immediately follows *u*.[7]

(8) START→b→u→k→u→END

[5]This approach to anchor points owes much to Alan Yu's (2002; 2003) work on possible infixal insertion points, including the work presented at NAPhC 2. See Fitzpatrick (2003) for arguments that the same set of anchor points is used for specification of infixal position and reduplicant shape.

[6]When MEL is null we represent the vocabulary item as PREC→FOLL.

[7]Raimy (2000a) treats the introduction of reduplicative precedence relations are the effect of a *readjustment rule* that applies in the environment or a (possibly null) affix. His approach has the advantage of accounting for cases in which a particular reduplication pattern invariably accompanies an overt affix. In the system presented here, it is possible to allow the introduction of immediate precedence relations through readjustment. However, we will not focus on this possibility here.

Since any legitimate precedence relation can appear in the lexical specification of a vocabulary item, we can provide a very straight-forward representation of certain affixes. For example, as proposed in Chomsky and Halle (1968), the English prefixes *a-* and *su-* are accompanied by the repetition of the consonant that they precede.[8] Thus the addition of /a-/ to /kɛpt/ results in [akkɛpt], which, through velar softening,[9] produces [aksɛpt]. Similarly, /sʌ-/ applied to /gɛst/ gives rise to [sʌggɛst] ⇒ [sʌgdʒɛst].

(9) Vocabulary item for /su-/:

START ⟶ s ⟶ ʌ ⟶ Seg$_{initial}$

For this vocabulary item, PREC the start symbol (START), MEL is /sʌ/, and FOLL is the initial segment. Additionally, an immediate precedence relation is specified from FOLL to FOLL. Vocabulary insertion of /sʌ-/ onto the stem /gɛst/ generates the representation in (10a). Linearization leads to (10b), and velar softening applies to produce the surface form (10c).

(10) a. START ⟶ s ⟶ ʌ ⟶ g ⟶ ɛ ⟶ s ⟶ t ⟶ END

b. START → s → ʌ → g → g → ɛ → s → t → END
c. START → s → ʌ → g → dʒ → ɛ → s → t → END

We have seen how attention to the fundamental phonological relation of immediate precedence allows the form of vocabulary items to be stated in such a way that prefixes, suffixes, infixes, and reduplication all have the same fundamental form. This unification is a welcome result, as these morpheme-types, particularly reduplication, have often required disparate theoretical machinery.

2 Across Occurrences: Interprecedence Relations

We turn our attention now to allomorphy in *fixed segment reduplication.*[10] Examples of *echo reduplication* in English, Hindi, and Kampuri are given in (11).

(11)

	Language	Fixed Segment		
a.	English	ʃm	table-shmable	eel-shmeel
b.	Hindi	v	paanii-vaanii	aam-vaam
c.	Kampuri	s	ghar-sar	

Alderete et al. (1999) argue that cases of fixed segment reduplication should be treated as the result of reduplication of the base plus competition between a segment of the base and the fixed segment for a position in the output form. Though space does not allow a thorough review of this claim, Nevins and Wagner (2001) provide extensive arguments against this competition-based overwriting approach. We propose that, just as with other vocabulary items, fixed segment reduplication involves the idiosyncratic specification of phonological content and precedence-based position. Thus, for example, Hindi *v*-reduplication would be listed as in (12).

(12) Seg$_{final}$ → v → V_1

However, the fixed segment in these patterns of echo reduplication is not invariant. When the reduplicated word begins with segments identical to the fix segment(s), allomorphy arises. Thus, rather than *Schmidt-*

[8]For Chomsky and Halle, [a-] and [sʌ-] have underlying forms /ab-/ and /sʌb-/, with place assimilation of the final consonant.

[9]The process of velar softening changes *k* to *s* and *g* to *dʒ* when they immediately precede front vowels. Evidence for this includes pairs such as *placate-placid*, *critic-criticize*, *rigid-rigor*, and *analogous-analogize*.

[10]Alderete et al. (1999) make the distinction between morphological fixed segmentism and phonological fixed segmentism based on the markedness/epenthetic-predictability of the putative fixed segment; we limit this section to the former.

Schmidt, English speakers with productive shm- reduplication often change the fixed segment to shp- or something similar.[11] In Hindi, v-initial words are reduplicated with a [ʃ] fixed segment. Kampuri uses a [t] when reduplicating [s]-initial words (Kenstowicz 1994:633).

(13)

	Language	FixSeg		
a.	English	ʃp	shmaltz-shpaltz	*shmaltz-shmaltz
b.	Hindi	ʃ	vakil-ʃakil	*vakil-vakil
c.	Kampuri	t	saati-taati	*saati-saati

This phenomenon presents a peculiar type of long distance-conditioned dissimilatory allomorphy. Most dissimilation effects are local. In the more familiar cases of locally-determined, phonologically-conditioned allomorphy, the exponence of a morpheme is conditioned by something in a local relation to the morpheme or segment in question. For example, the Spanish definite article surfaces as the masculine form *el*, rather than the feminine *la* when it immediately precedes a feminine noun that begins with a stressed *a* (14).

(14) a. el burro 'the donkey' (*masc.*), la mesa 'the table' (*fem.*),
b. el águila 'the eagle' (*fem.*), *la águila
c. el agua [el ágwa] 'the water' (*fem.*), *la agua

However, only a local relation will trigger this allomorphy. If the stressed *a* is further away from the article, the feminine form arises.

(15) a. la abarca [la abárka] 'the sandle' (*fem.*), *el abarca
b. la pequeña hacha [la pekéɲa átʃa] 'the little axe' (*fem.*), *el pequeña hacha

Other types of allomorphy do at times appear to be conditioned at a distance. For example, the Latin adjectival suffix *-alis* surfaces as *-aris* when preceded by another lateral.

(16) a. nav-alis 'naval', *navaris
b. sol-aris 'solar', *solalis
c. milit-aris 'military', *militalis
d. sepulchr-alis 'funereal', *sepulchraris

This dissimilation process is seemingly non-local; its triggering element may be an unbounded 'distance' away within the stem. In these examples, several segments, including consonants, intervene between the suffix *l/r* and the stem *l/r*. However, with the advent of autosegmental phonology, this effect was explained locally: the feature(s) in question are next to each other on their tier. The non-locality of this dissimilation effect is only apparent. Lateral dissimilation, like other dissimilation processes, obeys locality along the relevant dimension, as confirmed by the existence of examples such as *flor-alis* 'floral' (**floraris*) and *litor-alis* 'of the shore' (**litoraris*), where only the closest liquid matters for the determination of the suffixal content.

Allomorphy in echo reduplication does not show this type of unboundedness. Let us take the Hindi case (*paanii-vaanii, vakil-ʃakil*) as an example. The fixed segment in Hindi echo reduplication, which generally appears as [v], surfaces as [ʃ] precisely when the stem begins with [v]. On the surface, this stem-initial consonant is not in any local relation with the fixed segment. In fact, other [v]s may appear closer on the surface to the fixed segment, and yet these do not trigger allomorphy.

(17) yav 'barley' ⇒ yav-vav

The question is, then, why is it precisely the initial segment that triggers the observed alternation? On the surface, the exponent of the morpheme and this initial segment are not in a local relation that would allow simple formulation of the dissimilation process. A formal mechanism for capturing this alternation is hard to come by. AntiFaithfulness will not account for why it is precisely the FixSeg that is compromised and no other. That is, why would **vakil-vak**t*** not be an option for echo reduplication of *vakil*? Moreover,

[11] See Nevins and Vaux 2003 for detailed discussion of the different "dialects" of shm- reduplication.

no ranking of non-identity with respect to a Rhyme constraint will explain dissimilation in all cases. For example, Turkish has a pattern of partial reduplication (Kelepir 1999) in which the initial syllable is repeated with a fixed [p] inserted after the first vowel.

(18) a. ciliz 'alive', ci**p**-ciliz, 'alive (emphatic)'
b. yeni 'new', ye**p**-yeni, 'new (emphatic)'

A dissimilation effect also arises here. When the stem consonant that follows the first stem vowel is an obstruent, the fixed segment surfaces not as the obstruent [p], but as the sonorant [m].

(19) a. boʃ 'empty', bo**m**-bo**ʃ**, 'empty (emphatic)'
b. dik 'steep', di**m**-di**k**, 'steep (emphatic)'

This dissimilation pattern raises the same questions we asked regarding echo reduplication. Why is it precisely this segment that triggers the observed alternation? Is there any natural relation between the exponent of the morpheme and this segment such that the dissimilation processes can be stated simply?

For both echo reduplication and partial reduplication, non-linear representations offer a solution. The unlinearized representations for Hindi *paanii-vaanii* and *vakil-ʃakil* are given in (20a) and (20b), respectively. These representation make one similarity between the fixed segment and the stem-initial segment apparent: both immediately precede the initial vowel.

(20) a. START→p→a:→n→i:→END
v

b. START→v→a→k→i→l→END
ʃ

Turning to Turkish partial reduplication, the representations for *cip-ciliz* and *dim-dik* are given in (21a) and (21b), respectively. Once again, non-linear representations show that the relevant segments (in this case the fixed segment [p/m] and the second stem consonant) do have something in common—they both immediately follow the first vowel in the stem.

(21) a. START→c→i→l→i→i→z→END
p

b. START→d→i→k→END
m

In order to capture these precedence structure similarities, let us define two natural *interprecedence* relationships: *coprecedence* and *cosuccession*

(22) a. $\trianglerighteq$(x, y) (**x and y are coprecedent**) iff x→a and y→a.
b. $\trianglelefteq$(x, y) (**x and y are cosuccessors**) iff a→x and a→y.

That is, if two segments immediately follow the same segment, they are coprecedent. If they both immediately follow the same segment, they are cosuccessors. Using these relations, dissimilation in the Hindi *vakil-ʃakil* (**vakil-vakil*) case and the Turkish *dim-dik* (**dip-dik*) case are readily captured through rules or constraints with *locally* defined structural descriptions. A possible rule for the Hindi case is given in (23a). The Turkish p→m alternation is captured by (23b).

(23) a. v ⇒ ʃ / $\trianglerighteq$(v, _)
b. p ⇒ m / $\trianglelefteq$(_, [+obstruent])

Just as with local dissimilation on a particular autosegmental tier, non-linear representations expose the underlying locality of seemingly non-locally conditioned dissimilation. The derived interprecedence relations of coprecedence and cosuccession help explain why the dissimilating segment is always not, for example, the final consonant. *Vakil-vakit* would never be a candidate for dissimilation since the segment *l* is not in an interprecedence relation with the fixed segment.[12]

2.1 Economy and Ineffability

Besides partial reduplication, Turkish also has a process of full echo reduplication. But unlike Hindi echo reduplication, Turkish echo reduplication shows no avoidance allomorphy.

(24)	a.	kitap	kitap-mitap, 'books and so forth'
	b.	masa	*masa-masa (no allomorphy), 'tables and so forth'

Here FixSeg = /m/ and we see a pattern much like Hindi v-reduplication. However, reduplication is impossible with [m]-initial forms like (24b). They are simply ineffable and their meaning must be expressed periphrastically.

Under the assumption that negative evidence is not available in language learning, this is an interesting problem. Primary linguistic data (PLD) such as [kitap-mitap] or [mez-vez] would presumably lead the Turkish or Hindi learner, respectively, to posit a morpheme whose exponent is the insertion of a fixed segment that immediately follows the final segment and immediately precedes the initial vowel. But this simple statement generates *[vakil-vakil] for reduplication of /vakil/ and, *mutatis mutandis*, *[masa-masa] for /masa/. How does the learner *know* this form is not available?

We believe the answer can be found in the nature of the representations for these cases. Note that for the linearized form [vakil-vakil] there is an underlying representation (25b) containing fewer segments and relations than (25a), which is the representation created by the spelling out of the relevant vocabulary item.

(25) a. START→v→a→k→i→l→END   Linearized as: vakil-vakil

v

b. START→v→a→k→i→l→END   Linearized as: vakil-vakil

An adaptation of the proposals of Fox (2000) regarding economy and interpretation may be warranted here. In such a system, operations apply only if they have some designated interpretive effect. For example, a semantically vacuous quantifier raising operation is impossible.

Though we do not provide a specific implementation here of Fox's interpretive economy, the intuition is that in *[vakil-vakil] and *[masa-masa], there is no evidence for the fixed segment, that is, the segmental content of the morpheme is lost completely, and so the forms are ungrammatical as echo reduplication.[13]

Since this fact cannot be learned from positive evidence, we assume it arises from an innate property of human language, the nature of which is the focus of ongoing research.

3 Linearization: Multiple Precedence Mapped to Copies

By hypothesis, non-linear multiprecedence structures must be mapped to linear representations, ultimately legible to the sensory-motor interface. *Linearization* is the reconcatenation of a multiprecedence representation, containing "loops", into a single-precedence, linear representation.

[12]Note that in our system it is also possible to write a rule that adjusts the initial C, rather than the FixSeg. An additional context in the structural description, namely $\trianglelefteq$(_,END) for Hindi and $\trianglerighteq$(_,START) for Turkish, ensures that it is the FixSeg that dissimilates in these cases. However, this may not be necessary since phonologically-conditioned allomorphy does not generally affect already spelled out material in the stem (Bobaljik 2000). Assuming cyclic insertion of vocabulary items, the stem will have been given phonological form before the correct allomorph for echo reduplication is chosen.

[13]PLD such as [vakil-ʃakil] will give the Hindi learner a way to avoid the effects of economy through allomorphy.

We propose that linearization is guided by three natural principles: completeness, economy, and shortest. *Completeness* ensures that the segments and relations in the input are maximally spelled out in the output. (This could be thought of a MAX(Immediate Precedence).[14]) Informally, *economy* essentially says "do no more than is necessary." This condition, which is analogous to *STRUC, helps avoid gratuitous output that is not necessary for satisfaction of the other conditions. Finally, in the case of certain choice points, that is, segments that immediately precede multiple elements, *shortest* dictates that inside loops are spelled out before outside loops. We will see below how these conditions are computed.[15]

It should be noted that linearization is not simply a patch necessitated by the treatment of reduplication as the result of looping precedence relations. In fact, some linearization procedure is implicit in most treatments of affixation (especially infixation). Therefore, this approach can truly claim to contain no reduplication-specific mechanisms (*pace* Downing 2001).

3.1 Lushootseed Reduplication

In this section we provide a concrete illustration of our linearization algorithm. We also argue that, at least in some languages, certain forms result from derivations involving multiple linearization operations.

We will discuss three patterns of reduplication in Lushootseed and their interaction with each other and with certain vowel insertion properties of the language.[16] These three patterns are exemplified in (26).[17]

(26)	a.	Root:	ʔíbəʃ	'walk, travel'
	b.	Out-of-control (OOC):	ʔ-íb-ib-əʃ	'pace back and forth'
	c.	Distributive (DIST):	ʔíb-ʔib-əʃ	'walk all about'
	d.	Diminutive (DIM):	ʔí-ʔi-bəʃ-tx^w	'walk (something) a bit'

A preliminary specification of the vocabulary items that lead to these patterns is given in (27).[18]

(27) a. OOC: ImSucc($\acute{V}$) $\rightarrow$ $\acute{V}$
b. DIST: ImSucc(V_1)$\rightarrow$$C_1$
c. DIM: $V_1 \rightarrow Seg_{initial}$

For any form to which a single reduplicating exponent has been added, such as (28), *completeness* will force material to be repeated in the output form. In order to show that linearization produces the correct output form, we can compare this form with impossible linearizations in a tableau.[19]

(28) [DIM √ʔibəʃ]: START → ʔ → i → b → ə → ʃ → END

[14] The linearization algorithm uses only the relations and segments given to it as input. Since no new precedence relations are added during linearization, this is analogous to undominated DEP.

[15] We have implemented a version of this linearization algorithm that both maps non-linear representations to linear ones and abstracts non-linear patterns from (sets of) surface reduplicated forms. See Nevins and Iba (2004) and `http://swank.mit.edu/~reduplicator/` for more information.

[16] All data is from Bates et al. (1994). We owe much to the work of Bates (1986), Broselow (1983), and Urbanczyk (2001).

[17] Several different morphemes are spelled out using the OOC pattern, including *out of control* (24b), *particularization* (ia), and *personification* (ib) (used for counting people).

(i) a. dibəɬ 'we, us' ⇒ dibibəɬ 'just us'
b. c'uk^{w}s 'seven' ⇒ c'uk^{w}uk^{w}s 'seven people'

[ʔí-ʔi-bəʃ-tx^w] contains the causative transitive suffix *-tx^w*.

[18] There is one exception in Bates et al. (1994) to this characterization of OOC: wəlíʔ ⇒ [wələlíʔ-il]. However, our vocabulary item does account for certain cases where stress is non-initial: dx^w-ʔəhád ⇒ [dx^w-ʔəhádad] and ʔək^{w}yíq ⇒ [ʔək^{w}yíqiq]. If, as Yu (2003) suggests, these examples are truely prefixed forms, the vocabulary item could be reanalyzed as ImSucc(V_1)$\rightarrow$$V_1$.

[19] These conditions on linearization cannot be reranked to produce a factorial typology, a key feature of OT constraints. Thus their presentation in tableau form is purely for the sake of exposition. 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.

(29)

Input: (28)	SHORTEST	COMPLETENESS	ECONOMY
☞ ʔi-ʔibəʃ			ʔ-i
ʔibəʃ		(b-ʔ)!	
ʔi-ʔi-ʔibəʃ			ʔ-i-ʔ-i!

This analysis extends to other cases of single reduplication involving OOC, DIST, and DIM vocabulary items on CVC-initial stems (e.g., all the forms in (26)). However, other stems require further discussion.

Since the appearance of schwa is predictable in Lushootseed, we follow Bates (1986) and other scholars of Salish in positing that schwas are not present an underlying forms. We assume that some set of phonological constraints (PHONCON) ensures their insertion in the correct contexts. Furthermore, Lushootseed stress appears on the leftmost non-schwa vowel (STRESSLEFT).[20] If all vowels are schwa, stress appears on the left-most schwa. The ranking shown in (30) (*bədáʔ* ‘child’) ensures this outcome. DEP(¬ə) is shorthand for all Dep-V constraints, where V is any vowel but schwa.[21] The ranking DEP(¬ə), *ə́ ≺ STRESSLEFT ≺ DEP(ə) does not allow epenthesis of non-schwa vowels (like [i]). Therefore, stress is non-initial.

(30)

/bdaʔ/	DEP(¬ə)	PhonCon	*ə́	STRESSLEFT	DEP(ə)
bdáʔ		*!			
☞ bədáʔ				*	*
bə́daʔ			*!		*
bídaʔ	*!				

We will assume cyclic schwa insertion applies here. That is, schwas will be inserted after each morpheme spell-out. Were this not the case, it would be very difficult to express the placement of morphemes such as OOC in roots like √čx̌ ⇒ [čəx̌] ‘split’.

(31) [OOC √čx̌]: START → č → ə → x̌ → END   Linearized as: [čə́x̌əx̌]

This analysis correctly predicts the form of CVC-initial stems. However, some Lushootseed roots are not CVC-initial, even after schwa insertion. Non-CVC-initial stems do not follow the generalizations in (27). These forms require a slight addition to our vocabulary items.

(32)
a. OOC: ImSucc(V́) → V́
b. DIST: If σ_1 is CV, then ImSucc(V_1)→C_1; *elsewhere*: ImSucc(C_1)→C_1
c. DIM: If σ_1 is CV, then V_1 → C_1; *elsewhere*: C_1 → C_1

With this revision in hand, we can examine how our analysis extends to reduplication of CCV- stems. For example, √č’ƛ’aʔ surfaces as [č’ƛ’áʔ] ‘rock’. Its DIST form (‘rocks scattered about’) is given in (33).

(33) [dist √č’ƛ’aʔ]: START→č’→ƛ’→a→ʔ→END

(34)

(33)	DEP(¬ə)	PHONCON	STRESSLEFT	DEP(ə)
☞ č’ƛ’č’ƛ’áʔ				
č’əƛ’č’əƛ’áʔ			*!*	**
č’íƛ’č’əƛ’aʔ	*!			*

However, DIM forms of CCV- (and CVV-) initial forms do not behave in the same way. Instead, we observe epenthesis of an [i] to split up the initial C_1C_1 cluster. This [i] then bears main word stress. [i]-epenthesis also applies in Cə- (35b) and CVV-initial DIM forms (35c).

[20] For the purpose of this exposition, STRESSLEFT can be read as a gradient alignment constraint.

[21] Our presentation of this analysis in an OT framework in no way denies the possibility or plausibility of analyses in other frameworks. The form of presentation does emphasize that we are exploring a new representational vocabulary that could, in principle, be adopted in a variety of frameworks.

(35) a. č'ƛ'áʔ 'rock', č'íč'ƛ'aʔ 'little rock'
b. təláw-il 'run', títəlaw'-il 'jog'
c. s-duuk^w 'knife', s-díduuk^w 'small knife'

We suggest that the three morphemes under discussion here are separated into two classes: Class A (DIST and OOC), and class B (DIM). These two classes are subject to different phonological rules and constraints. One difference is that Class A morphemes are spelled out using the constraint ranking shown above (repeated in (36a)). Class B, on the other hand, is evaluated using the slightly different ranking in (36b).

(36) a. DEP(¬ə) , MAX$_{IO}$, PHONCON, *ə́ ≺ STRESSLEFT ≺ DEP(ə)
b. DEP(¬ə/i) , MAX$_{IO}$, PHONCON, *ə́ ≺ STRESSLEFT ≺ DEP(i), DEP(ə)

Since unlike (36a), (36b) allows i-epenthesis, this ranking leads to the following derivation of [DIM $\sqrt{}$č'ƛ'aʔ] ⇒ [č'íč'ƛ'aʔ]. The derivation of (35c) is identical in all relevant respects.

(37) START → č' → ƛ' → a → ʔ → END

(38)

(37)	PHONCON	*ə́	STRESSL	DEP(i)	DEP(ə)
☞ č'íč'ƛ'aʔ				*	
č'ə́č'ƛ'aʔ		*!			*
č'ič'ƛ'áʔ			*!	*	
č'č'ƛ'áʔ	*!				

If schwa insertion does not apply to the stem to which DIM is applied, the derivation of Cə-initial forms proceeds in the same way. However, we are assuming cyclic schwa insertion, based on the fact that the placement of OOC often depends on the placement of schwas (see below). Still, the constraint ranking for Class B predicts correctly that [i] can be inserted for stress (we treat this as insertion of the Place features of [i]).

(39) START → t→ə → l → a → w → END

(40)

(40)	MAX$_{IO}$	*ə́	STRESSLEFT	DEP(i)	DEP(ə)
☞ títəlaw				*	
tə́təlaw		*!			
tətílaw			*!	*	

Here we are assuming the MAX$_{IO}$ is evaluated existentially (Struijke 2000; Raimy and Idsardi 1997). That is, MAX$_{IO}$ is satisfied if for each segment/feature in the input there is a corresponding segment/feature in the output. Thus there are no violations of MAX$_{IO}$ in any of the candidates in (40).[22]

Urbanczyk (2001) proposes that DIM and OOC are in a class together (for Urbanczyk, these are "affixes"), with DIST in a different class (the "root" class). In Urbanczyk (2001), roots and affixes are subject to different constraints. However, if OOC were in the same class as DIM, we would expect [i]-insertion to appear in OOC forms. Only one unambiguous form that could be analyzed as i-insertion is listed in Bates et al. (1994) (*k^w'əq* ⇒ *k^w'əqíq*), whereas all other OOC examples show no i-insertion (41a). Similarly, DIST forms do not allow i-insertion (41b).

(41) a. bək^{w}' 'all' ⇒ bə́k^{w}'ək^{w}' 'they all'
čəx̌ 'split' ⇒ s-čə́čəx̌ 'smashed up'
d^{z}ək^{w}' 'travel' ⇒ d^{z}ə́k^{w}'ək^{w}' '...changing all the time'
d^{z}əx̌ 'move' ⇒ d^{z}əx̌ 'visit'

[22]This existential evaluation of MAX is a result of the fact that there is no base or reduplicant in the output form under our assumptions. MAX-IO is defined as in McCarthy and Prince (1995:16): "Every segment of the input has a correspondent in the output" (i.e., $\forall S_x \exists S_y [S_x \in \text{INPUT} \rightarrow [S_y \in \text{OUTPUT} \wedge \Re(S_x,S_y)]]$, where $\Re$ is the IO-Correspondence relation).

b. ʔəɬəd ‘eat’ ⇒ ʔəɬʔəɬád ‘many eat’
bədáʔ ‘child’ ⇒ bədbədáʔ ‘children’
bəd ‘thin’ ⇒ bə́dbəd ‘thin (board)

Under Urbanczyk’s constraint ranking (42) (compiled from various tableaux in Urbanczyk (2001:ch. 4)), however, we would expect [i]-insertion and stress shift in all of the forms in (41a). (For Urbanczyk, the second -VC sequence is the reduplicant, while the first is part of the base.) Here IO-Dep-$\sqrt{}$ is a set of Dep constraints that apply to root material, while IO-Dep-Afx applies to affixal material (see Urbanczyk (2001) for full discussion.)

(42)

{bəkʷ, ooc}	IO-Dep-$\sqrt{}$	*ə́	StrL	BR-Dep-Afx
☞ bəkʷíkʷ			*	*
bə́kʷəkʷ		*!		
bíkʷəkʷ	*!			

Urbanczyk’s incorrect prediction comes from failure to distinguish correctly between these classes of affixes. In the next section we turn to further motivation for the Class A/Class B distinction.

3.2 Multiple Reduplication Patterns

A key demonstration of linearization comes in the introduction of multiple reduplication, since these precedence relations are all present in the same representation. Lushootseed allows forms that contain multiple reduplicative affixes. We will discuss the patterns in (43).

(43) a. [dist [ooc $\sqrt{}$root]]
b. [dim [ooc $\sqrt{}$root]]
c. [dim [dist $\sqrt{}$root]]
d. [dist [dim $\sqrt{}$root]]

Turning to (43a), it is important to note that the data available in Bates et al. (1994) seems to suggest that forms with both dist and ooc morpheme may arise from either of the hierarchical structures in (44).

(44) a. [dist [ooc $\sqrt{}$root]]
b. [ooc [dist $\sqrt{}$root]]

Examples of (44a) should have semantics along the lines of ‘many X’s involved in random, ineffectual action’ (cf. Bates et al. (1994:xvii)). Examples include *saqʷ’aqʷ’saqʷ’* ‘many flying around, wheeling in the sky’ (from *saqʷ’* ‘fly’) and *gʷáxʷaʷgʷaxʷ* ‘a lot of strolling about’ (from *gʷáxʷ* ‘go for a walk’). If (44a) forms were derived cyclically, we would expect *saqʷ* ⇒ *saqʷaqʷ* ⇒ *saqʷsaqʷaqʷ*, contrary to fact.

ooc seems to have little effect on the semantics of dist forms. Examples of [dist $\sqrt{}$root] forms (44b) include *bálalbali* ‘(suddenly) I am forgetful’ (cf. *báli* ‘forget’, *bəlbáli* ‘I am forgetful’) and *gʷádadgʷad* ‘discuss, converse, talk over’ (cf. *gʷad* ‘talk’, *gʷádgʷad* ‘talk (a lot), speak up’).

Despite the variable hierarchy of the ooc and dist morphemes, all dist-ooc forms appear with the same pattern. This shape-invariance can be accommodated if we assume that linearization does not apply between the spell-out of the two morphemes. Thus both hierarchical structures in (44) will lead to the phonological representation in (45).

(45) START→b→a→l→i→END

(46)

	Input: (45)	Complete	Shortest	Economy
☞ a.	bal-al-bali			a-lb-a-l
b.	bal-bal-ali		*!	b-a-la-l
c.	bal-ali	(l-b)!		a-l
d.	bal-al-al-bali			a-l-!a-lb-a-l

Completeness rules out (46c), which fails to include the l⟶b relation in the output. Economy rules out (46d) since this form results from unnecessarily crossing the l⟶a relation more than once. Shortest rules out (46b) since this form would result if the l⟶b relation were crossed before the shorter l⟶a relation. Length is calculated at [l] with forward relations to [b] and [a] by comparison of material intervening between [b] and [l] and between [a] and [l]. While [a] must be crossed to get from [b] to [l], only the link between [a] and [l] must be crossed to get from [a] to [l]. Thus the l⟶a link is shorter.[23] Descriptively, Shortest is operative when there are two backwards-pointing precedence relations, choosing the one whose endpoint is closer in terms of transitively preceding the source (an "inner loop"). Due to the invariant nature of linearization, the same surface form will arise in a non-cyclic derivation, regardless of the morpheme hierarchy. Under a cyclic derivation, only the [ooc [dist √root]] form is predicted to surface with the observed shape.

Turning to [dim [ooc √root]], we find a demonstration of another aspect of completeness. Completeness provides a preference for relations that have not yet been spelled out. When more than one such relation exits a given segment, completeness favors back-pointing relations ("loops"). In the linearization of (47), due to a local preference for back-pointing links, *a⟶d* is spelled out before *a⟶y*.

(47) start→d→a→y→ʔ→end

(48)

	Input: (47)	Shortest	Completeness	Economy
☞ a.	dá-day-ayʔ			d-aa-y
b.	dáy-a-dayʔ		(a-d)!	aday

Here candidate (48b) would arise if start→d→a→y were spelled out, followed by the y→a loop, then the a→d loop, proceeded by d→a→y→ʔ→end.

For much the same reason (satisfaction of locally-computed completeness), [dim [dist √root]] surfaces as shown in (49a), rather than (49b).

(49) start→s→a→x^w→end

a. sásax^{w}sax^w
b. *sáx^{w}sasax^w

For forms whose first vowel is schwa, dim's status as a Class B morpheme comes into play. [dim [dist √bdaʔ]] surfaces as [bíbədbədaʔ], rather than [bəbədbədáʔ]. This is because dim forms allow the insertion of [i] (again, treated here as insertion of Place features), with accompanying leftward stress shift (bíbədbədaʔ 'litter (of animals); dolls).

(50) start→b→ə→d→a→ʔ→end

(51)

	Input:	Max_{IO}	*ə́	StressL	Dep(i)
☞ a.	bíbədbədaʔ				*
b.	bəbədbədáʔ			*!**	
c.	bəbídbədaʔ			*!	*
d.	bə́bədbədaʔ		*!		

[23]It may be the case that calculation of Shortest can always be done on previously linearized material, thus obviating the need for look-ahead.

Unlike the situation we saw with DIST-OOC forms, in which either hierarchical order produced the same form, the relative position of DIM and DIST has an effect on the phonological form. Since [DIST [DIM ROOT]] contains the same morphemes as [DIM [DIST ROOT]] shown above, application of both morphemes should result in the same form, unless linearization intervenes between the two. In fact, [DIST [DIM /bdaʔ/]] = [bíbibədaʔ] ‘young children’ rather than *[bíbədbədaʔ]. This [C_1VC_1VC_1V...] pattern is observed in numerous other forms:

(52)

	Root	[DIST [DIM √root]]
a.	χay	χa-χa-χay
b.	pastəd	pa-pa-pstəd
c.	duuk^{w}	di-di-duuk^{w}

We propose that linearization applies between the Class B DIM morpheme and the Class A DIST. The derivation of the [DIST [DIM ROOT]] form [bíbibədaʔ] proceeds as in (53).

(53) a. [DIM bdaʔ]: START→b→d→a→ʔ→END Linearized as: [bíbədaʔ]

b. DIST applies: START→b→i→b→ə→d→a→ʔ→END ⇒ [bibbibədaʔ]

c. Degemination applies:[24] [bíbibədaʔ]

Thus we have further evidence for a distinction between Class A and Class B morphemes: besides allowing [i]-insertion and stress shift, Class B morphemes trigger linearization, whereas Class A morphemes do not appear to trigger linearization. It is possible that linearization applies between the application of morphemes of different classes, perhaps along the lines of the cyclic/noncyclic distinction of Lexical Phonology. Mester (1988:178-179) arrives at a similar conclusion, reached through different representational assumptions about reduplication: “...the theory makes predictions about the point(s) in the derivation where linearization occurs. Linearization will take place whenever Tier Conflation/Bracket Erasure is invoked... If Tier Conflation/Bracket Erasure is stratal and not cyclic, there is thus a certain delay between morphological formation and morphological destructuring... Until Tier Conflation applies to [reduplicated forms]... that is, *until they exit their stratum of formation*, they remain nonlinearized.”

4 TETU without B-R Correspondence

The Multiprecedence-and-Linearization model of reduplication does not include any Correspondence-theoretic relationship between output base and reduplicant. The immediate question that arises is how “the emergence of the unmarked” (TETU) effects are captured. At first blush TETU is a generalization that perhaps *cannot* be captured by any theory that does not recognize a surface base and reduplicant. TETU is commonly found in analyses of reduplication with the ranking:

(54) IO-Faith ≫ PHONCON ≫ BR-Faith

In order to address this problem, we must ask what “unmarked” means. A review of the literature shows that TETU is not a homogeneous category. On the contrary, quite disparate phenomena are often grouped under this rubric. Separation of “TETU effects” into a clear typology shows that when taken on a type-by-type and case-by-case basis, these phenonema may find some natural analyses in a precedence-based approach.

If unmarked simply means “less”, then the precedence-based representation need not be challenged. Reduplicating readjustment is a new precedence relation that is *maximally* as long as the entire word, and minimally one segment. Unmarked-*qua*-less happens when the loop is short. Of course, the prosodic morphologist doth protest: why is “less” usually CV, or CVC. Usualness may be correlated with naturalness

[24] Urbanczyk (2001:283) keenly notes that, contra Broselow (1983), a degemination process is well-motivated empirically since single DIST forms that would create geminates (C_1VC_2 forms where C_1=C_2) generally surface as CV rather than CVC: [t’it’-əb] ‘bathe’, [t’**it’**it’-əb] ‘bathe for a while’ (*[t’**it’t’**it’-əb]).

of rules, but unnaturalness need not be ruled out: consider Mangarrayi, in which the repeated substring is a VC(C) sequence. Is that "unmarked"?

(55)

Vocabulary Item	PLURAL Form	Gloss
g**abu**ji	gababuji	'old persons'
y**ir**ag	yirirag	'fathers'
j**img**an	jimgimgan	'knowledgeable ones'

The Thao (Chang 1998) repetitive/habitual (56) allows for a maximally CVCV reduplication pattern, but on the surface this results in quite marked structure.

(56)
a. šnara 'to ignite' ⇒ pa-šnara-nara 'to burn s.t. repeatedly'
b. ag.qtu 'to contemplate' ⇒ agqtu-qtu 'think about'
c. ar.faz 'to fly' ⇒ m-arfa-rfaz 'to keep flying around'
d. m-ig.kmir 'to grasp in one hand' ⇒ igkmi-kmir-in 'to be rolled into a ball in one hand'

For example, *ar.faz* reduplicates as *ar.far.faz* (57), copying the coda of the first syllable and creating a new coda; *r.fa* is not a prosodic constituent in anyone's book.

(57) START→a→r→f→a→z→END

Thus, reduplication models without B-R correspondence or the prosodic-constituent-requirement are not missing out on any generalizations about the size of the reduplicant.

Other apparent instances of TETU include apparent segmental change. Consider *Non-Reduplicative Copying*. In cases of CV prefixation in which an "epenthetic" segment is preceded by a copy of the first vowel, the analysis need not be one of reduplication with replacement of the base vowel by a more unmarked nucleus. The case of Yoruba gerundive affixation has been analyzed by Kawu (1999) as prefixation of a high-toned vowel (or just a high tone), followed by a non-reduplicative repair copying. (See also Bissell 2002 for Miya):

(58) PREFIXATION→ íjo ONSET REPAIR→jíjo

In a similar manner, some cases of purported TETU in fixed segment reduplication (Alderete et al. 1999) can be fruitfully analyzed in such a way that the fixed segment is not part of the reduplicant at all, but arises through post-reduplication processes. One straightforward analysis of the Lushootseed *i*-epenthesis shown above is the following:

(59)
a. DIM inserts an immediate precedence relation: $C_1 \rightarrow C_1$
b. In addition, DIM adds a `Left (` to the metrical grid, requiring stress on the initial syllable.
c. Schwa-insertion to provide nuclei for unsyllabified segments occurs; however, *ə triggers a repair operation, inserting stressable *i* in place of the schwa.
d. The epenthesis is thus stress-driven and has nothing to do with the "reduplicant".

This analysis is an alternative to the OT analysis shown above, which does produce i-epenthesis without a base and reduplicant in the output.

Finally, consider cases of cluster-simplification, as in the Sanskrit perfect *pa-prath-a* (from *prath*). The precedence structure is as follows:

(60) START→p→r→a→th→END 'spread'

As far as we know, there are no cases of reduplication in any language such as *krand-kan.* The only putative counterexample is Chukchi (Krause 1980), *nute-nut.* However, McCarthy and Prince (1996) (as well as Bobaljik (2003) for Itelmen) demonstrate that an independent process of word-final apocope of final-stem vowels, is operative in the language. Hence it seems secure to propose the *UCLAP* generalization:

(61) ***U**nfaithful **C**opy **L**inearizes **A**s **P**refixal*: Linearization of a structure with segmental ”skipping” always places the unmarked copy as a surface prefix

“Unfaithful” in (61) refers to loss of segmental material. An illustration of the difference between segmental loss and “unmarkedness” comes from Marathi subtractive reduplication (Apte 1968), *amne-samne*, in which the initial onset is deleted in the first copy. Clearly this contravenes the edge-in constraint on association, as well as TETU, as ONSET is violated in the reduplicant.

Consider the implementation of UCLAP. The linearization procedure always favors *forward-pointing* links, those whose source transitively, but not immediately, precede their endpoint. In fact, this falls out as part of the definition of shortest:

(62) Shortest:

a. If x and y immediately precede z, and x transitively precedes y, linearize x to z first.
b. If x and y immediately follow z, and x transitively precedes y, linearize z to y first.

The unification of linearization in multiple reduplication (where markedness played no apparent role in arbitrating between [bíbədbədaʔ] and *[bədbíbədaʔ] and unfaithful-prefixing reduplication in Marathi *amne-samne* through the definition of Shortest suggests that transitive precedence plays a crucial role in the computation of reduplication.

Certainly, markedness may play a role in the explanation of phonological processes. As a blanket concept, however, it does not seem to govern the form of reduplicants. Generalizations about post-reduplicative repair related to independent phenomena in the language, and the UCLAP result of linearization are two cases in which no base-reduplicant correspondence plays a role in the analysis. Further work on the typology of TETU effects and their counterexamples may reveal each apparent instance to be part of a different aspect of phonological representations.

5 Contributions

Patrick Winston, an eminent researcher in artificial intelligence, is credited with remarking “If the representations are right, the rules will just follow.” With the adoption of multiprecedence representations, the description of phonological phenomena that would require complicated conjunctions of structural descriptions in the form of markedness constraints and structural changes in the form of anti-repair constraints can instead be literally read off the structure. The structural description for echo word dissimilation is made immediately apparent and natural by inspection of coprecedence relations. The structural change for linearization of the combination of two different reduplication patterns is immediately predictable based on the topography of precedence; shorter links are visible upon inspection. Focusing on Precedence as the relation of interest is not meant to replace other aspects of phonological representation; constituent relations such as Onset and Metrical Feet are indisputable elements in the competence of language users. These relations themselves may be subject to further inspection of their formal properties; indeed, ambisyllabicity and appendix phenomena might demonstrate that a relaxation of the one-to-one relation between segment and constituent membership holds at removed levels of representation. Onset itself may turn out to be expressible as a specialized subtype of immediate precedence, as immediate precedence is deducible from an Onset relation. Note, however, that precedence among constituents of different syllables (e.g. the nucleus of the first syllable and the onset of the second syllable in ***pa.tu*** cannot be deduced from any other relation; from this we conclude that the precedence relations here cannot be wholly deduced from constituent relations. Since phonology is stuck with them, we might as well make use of them in our explanations.

References

Abels, Klaus. 2003. The syntax of adposition stranding. Doctoral Dissertation, University of Connecticut, Storrs.

Alderete, John, Jill Beckman, Laura Benua, Amalia Gnanadesikan, John McCarthy, and Suzanne Urbanczyk. 1999. Reduplication with Fixed Segmentism. *Linguistic Inquiry* 30:327–364.

Apte, Mahadeo. 1968. Reduplication, echo formation, and onomatopoeia in Marathi. Poona: Deccan College Post-graduate and Research Institute.

Bates, Dawn. 1986. An analysis of Lushootseed diminutive reduplication. In *Proceedings of the Berkeley Linguistics Society 12*, 1–13.

Bates, Dawn, Thom Hess, and Vi Hilbert. 1994. *Lushootseed dictionary*. University of Washington Press.

Bissell, Teal. 2002. Emergence of the marked in Miya pluractional allomorphy. In *Phonological answers (and their corresponding questions)*, ed. A. Csirmaz, Z. Li, A. Nevins, O. Vaysman, and M. Wagner, volume 42. MITWPL.

Bobaljik, Jonathan. 2000. The ins and outs of contextual allomorphy. In *Proceedings of the Maryland Mayfest on Morphology*, ed. K. K. Grohmann and C. Struijke, volume 10, 35–71. University of Maryland Working Papers in Linguistics.

Bobaljik, Jonathan. 2003. Itelmen reduplication: Edge-in association and lexical stratification. Ms., McGill University. Material presented at the Montreal-Ottawa-Toronto Phonology Workshop.

Broselow, Ellen. 1983. Subjacency in morphology. *Natural Language and Linguistic Theory 1* 317–346.

Chang, M. Laura. 1998. Thao reduplication. *Oceanic Linguistics* 37:277–297.

Chomsky, Noam. 1995. *The minimalist program*. MIT Press.

Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of english*. MIT Press.

Downing, Laura. 2001. Review of Raimy 2000. *Phonology* 18:445–451.

Fitzpatrick, Justin. 2003. The representation and linearization of non-linear phonological structures. Ms., MIT.

Fitzpatrick, Justin, and Andrew Nevins. 2003. Linearizing nested and overlapping precedence in multiple reduplication. In *The Proceedings of the 27th Penn Linguistics Colloquium*.

Fox, Danny. 2000. *Economy and semantic interpretation*. MIT Press.

Halle, Morris, and Alec Marantz. 1993. Distributed Morphology and the Pieces of Inflection. In *The view from building 20*, ed. Kenneth Hale and Samuel J. Keyser. Cambridge, Mass.: MIT Press.

Kawu, Amhadu. 1999. Structural markedness and nonreduplicative copying. In *The Proceedings of NELS 30*. GLSA: Amherst.

Kayne, Richard. 1994. *The antisymmetry of syntax*. MIT Press.

Kelepir, Meltem. 1999. Turkish emphatic reduplication and antifaithfulness. In *The Proceedings of ConSOLE 7*. University of Bergen, Norway.

Kenstowicz, Michael. 1994. *Phonology in generative grammar*. Blackwell.

Krause, Scott. 1980. Topics in Chukchee phonology and morphology. Doctoral Dissertation, University of Illinois.

McCarthy, John, and Alan Prince. 1995. Faithfulness and reduplicative identity. In *Papers in optimality theory*, ed. Jill Beckman, Laura Walsh Dickey, and Suzanne Urbanczyk, volume 18 of *University of Massachusetts Occasional Papers*, 249–384. Graduate Linguistic Student Association.

McCarthy, John, and Alan Prince. 1996. Prosodic morphology 1986. In *Technical report*. Rutgers University.

Mester, Armin. 1988. *Studies in tier structure*. Garland Outstanding Dissertations.

Nevins, Andrew, and Aaron Iba. 2004. A selectionist learner for parametrically ambiguous reduplicants. Presented at the Annual Meeting of the Linguistic Society of America.

Nevins, Andrew, and Bert Vaux. 2003. Metalinguistic, schmetalinguistic: The phonology of shm- reduplication. In *The Proceedings of the Chicago Linguistic Society*, volume 39.

Nevins, Andrew, and Michael Wagner. 2001. Can correspondence fix fixed segmentism? Talk presented at LOT-NWCL, Utrecht, 7/01.

Raimy, Eric. 2000a. *The phonology and morphology of reduplication*. Mouton de Gruyter.

Raimy, Eric. 2000b. Remarks on backcopying. *Linguistic Inquiry* 31:541–552.

Raimy, Eric, and William Idsardi. 1997. A minimalist approach to reduplication in OT. In *The Proceedings of NELS 27*. GLSA. University of Massachusetts, Amherst.

Struijke, Caro. 2000. Why constraint conflict can disappear in reduplication. In *Proceedings of NELS 30*.

Urbanczyk, Suzanne. 2001. *Patterns of reduplication in Lushootseed*. Garland Outstanding Dissertations.

Yu, Alan. 2002. Understanding infixes as infixes. Presented at NAPhC 2.

Yu, Alan. 2003. The morphology and phonology of infixation. Doctoral Dissertation, University of California, Berkeley.

Department of Linguistics and Philosophy
E39-245, MIT
77 Massachusetts Avenue
Cambridge, MA 02139
jfitzpat@mit.edu
anevins@mit.edu