# Sources of Multiple Reduplication in Salish and Beyond[1]

*Justin M. Fitzpatrick*
*MIT*

Data that could adjudicate between theories of multiple reduplication are relatively hard to come by. The languages of the Pacific Northwest provide a fertile ground for testing and development of such theories. I examine patterns from Lushootseed (Salish) and Nitinaht (Wakashan), as well as typologically and geographically distant languages like Thao and Manam (Austronesian) and develop a theory where reduplicated forms are the result of linearization of non-linear phonological forms (Raimy 2000). I argue that non-linearity can appear in lexical representations as well as in poly-morphemic forms and that linearization applies at several points in a derivation, rather than solely at the interface with phonetics. The latter point challenges Urbanczyk's (2001) monostratal Optimality Theoretic analysis of Lushootseed and suggests that several strata of linearity and non-linearity intervene between lexical and surface phonological forms.

## 1. Introduction

In this paper I propose, following Raimy (2000), that phonological precedence can be non-linear. That is, a single segment may immediately precede or follow more than one segment. I expand this proposal, however, to allow for non-linearity at multiple points in a morpho-phonological derivation. After outlining my assumptions in section 2, I begin in section 3 with a case of lexical non-linearity. I argue there, drawing on evidence from Manam (Austronesian), that stored lexical phonological representations can contain non-linear precedence. I then go on in section 4 to argue that phonological derivations can contain multiple levels of non-linearity and linearization. Evidence for this proposal comes from patterns of multiple reduplication in Lushootseed (Salish), which I argue fail to be captured

[1]I wish to thank the audiences at the 2nd North American Phonology Conference, the 27th Penn Linguistics Colloquium, and the 1st Escola de Verão de Lingüística Formal da América do Sul, where portions of this work have been presented. I have benefitted greatly from discussions (email and otherwise) with Dawn Bates, Eugene Buckley, John Frampton, Morris Halle, Aaron Iba, Alec Marantz, Galen Pickard, Eric Raimy, John Stonham, and especially Andrew Nevins, my co-conspirator during much of this work. All remaining shortcomings are of course my own. The section on Manam appeared in slightly different form in Fitzpatrick and Nevins (2004). This work was supported under a National Science Foundation Graduate Research Fellowship.

by the approach of Urbanczyk (2001). Finally, I show how both phenomena can coexist in the same system; I argue, drawing on Stonham (1994), that Nitinaht (Wakashan) contains evidence for lexical reduplication and for multiple levels of linearization.

## 2. An Explicit Representation of Precedence

Raimy (2000:12) 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.

(1) 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 (2000) proposes that precedence relations be explicitly represented in phonology: A→B is read 'A immediately precedes B.' In this system, (2a) and (2b) both represent what is traditionally represented as [kæt].[2]

(2) 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 in section 2.1 and section 4).

[2]For the moment, immediate precedence relations only seem motivated between X-slots (or a CV skeleton) on the timing tier. Thus the representations in (2) are shorthand for (i).

(i) START→X→X→X→END
k æ t

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 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, either immediately or transitively) 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.[3] 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 (2000) 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.

I will develop a theory in which representations like (3b) are possible, and where such non-linear representations are mapped to linear representations that include repetition of phonological material (Raimy 2000). Under this mapping, which I will call *linearization*, (3b) is mapped to *kal-kaldíŋ* ‘goats’. Some of the empirical motivation for this treatment of reduplication can be found in Raimy (2000), Fitzpatrick and Nevins (2004), and Raimy and Idsardi (1997). Further investigation of this model in the realm

[3] 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.

of affixation (especially infixation, cf. Yu 2003) can be found in Fitzpatrick (to appear). 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).

### 2.1. Linearization

By hypothesis, a non-linear representation such as (3b) must be *linearized* if it is to be legible to sensory-motor systems. *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.

Linearization will be treated in more detail in section 4. However, a brief introduction is given here to facilitate the understanding of the following sections. The linearization procedure 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. It does this in a local manner; at each segment, as-yet untraversed backward-pointing links are chosen over forward-pointing links.[4] 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.

Thus, (3b), repeated as (4a), is linearized as (4b). Linearization begins at the START symbol and proceeds forward along the 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).

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

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

One can think of Completeness as a hedge on the linearizer's bets. At any point with multiple outgoing immediate precedence relations, taking the backward-pointing loops at the earliest possible point ensures that these loops are spelled out. If they were not, it is quite possible there will be no other chance to spell them out. I argue below that Completeness is not a linearization-specific principle, but is a general principle that governs all computations over precedence structures. This hypothesis is stated in (5), where *choice point* is understood as a segment with more than one immediately preceding or following segment.

[4] A global form of Completeness, in which different possible linearizations of a form are compared, is also theoretically possible. I pursue the local version here since it is computationally more tractable.

(5) At a choice point, all precedence-based computations follow backward-pointing “loops” first.

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. 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.[5]

(7)

| | Input: (4a) | Completeness | Economy |
|---|---|---|---|
| ☞ | kal-kaldìŋ | | k-a-l |
| | kaldìŋ | (l-k)! | |
| | kal-kal-kaldìŋ | | k-a-l-k!-a-l |

Finally, in the case of multiple “loops,” as in (8) (from Lushootseed, discussed in more detail in section 4), *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 ⇒[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).

[5]These linearization constraints are not to be thought of on a par with normal OT constraints, since they cannot be reranked.

(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 |

This approach to linearization differs markedly from Raimy's (2000) version, which makes crucial use of morphological recency in determining which relations are "spelled out" first. Ultimately, it is an empirical question whether morphological ordering information plays an active role in linearization.

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 into the original stem. Therefore, this approach can truly claim to contain no reduplication-specific mechanisms (*pace* Downing 2001).

## 3. Non-linearity in the Lexicon: the case of Inherent Reduplication

Most of the works cited above that propose precedence-based explanations of morpho-phonological phenomena have focussed on ways in which non-linear representations that result from the spell-out of poly-morphemic structures provide insight into puzzling phonological facts. In this section, I suggest that mono-morphemic non-linearity is also possible. That is, lexical entries for phonological forms can themselves contain "loops." That is, there are *inherently reduplicated* roots. After introducing inherent reduplication here for Manam (Austronesian) in this section and multiple levels of linearization for Lushootseed (Salish) in the next, I will show how the two interact in Nitinaht (Wakashan) in section 5.

If it is true that languages prefer to encode segmental repetition as reduplication, as Zuraw (2000) argues, then inherent reduplication may be the norm. That is, when a learner observes a mono-morphemic form that contains the repetition of phonological material, she may posit inherent reduplication as a first hypothesis. For example, Thao (Blust 2001) contains various forms that never appear in non-repeated form.[6]

[6] Since the language independently uses reduplication for iterative/habitual formation (Chang 1998), the fact that many of these words denote repetitive actions (chewing, shaking, etc.) might lead a Thao speaker to posit non-lexical reduplication here. If this is the case, these would not be examples of mono-morphemic inherent reduplication. Rather, these words would be inherently marked with the iterative morpheme, perhaps akin in some way to bound roots like *mit* (e.g., permit, remit, submit) in English are inherently marked to require a prefix.

(10)

| | | | |
|---|---|---|---|
| | karkar | 'chew' | (*kar) |
| | kashkash | 'scratch, scrape, hoe, rake' | (*kash) |
| | qucquc | 'tie, bind' | (*quc) |
| | shakshak | 'grope, feel for' | (*shak) |
| | shishi | 'shake something (back and forth)' | (*shi) |

However, I have not discovered that these forms behave in any special way with respect to other morphological processes.[7] Therefore, at least some of these Thao forms may simply be examples of words that, quite accidentally (from a synchronic point of view), contain repeated phonological material. In order to make an argument that lexical forms contain inherent reduplication, rather than accidental repetition of phonological material, one must find a way in which forms that contain repetition behave differently from other forms. If this difference in behavior can be explained by invoking inherent reduplication, it seems likely that their lexical entries contain a representation of reduplication. The question will then be *what is the nature of this representation?*. I turn now to such a case.

Manam (Austronesian) contains a pattern of rightward bimoraic reduplication, which creates adjectival forms, agentives, and "continuative, progressive, persistive" aspects of verbs (Lichtenberk 1983:608).[8] Color terms are also derived through reduplication from a noun, and mean something like "the color of X."

[7]It is noteworthy, perhaps, that these forms constitute a large portion of the Thao forms that undergo *triplication*, as in (ia). (note especially that *kashkash* becomes simply *kash* in the triplicated form.) However, forms that do not appear to be inherently reduplicated also undergo triplication (ib). The latter can also undergo regular patterns of reduplication, however. These forms show Ca- reduplication/triplication. -m- is the actor focus marker.

| | | | |
|---|---|---|---|
| (i) | a. | shishi 'shake' | sh-m-a-sha-shishi 'shake repeatedly' |
| | | kashkash 'scratch' | k-m-a-ka-kash 'to scratch repeatedly' |
| | b. | shkash 'fear' | makit-shka-shkash 'gradually grow fearful' |
| | | | makit-shka-shka-shkash 'gradually be overcome with fear' |
| | | zay 'turn' | qata-za-zay 'turn the head from side to side' |
| | | | qata-za-za-zay 'ceaselessly turn the head from side to side' |

Without further research, it is hard to say whether this difference is significant. It may simply reflect the limited data available.

[8]Lichtenberk (1983) identifies eight patterns of reduplication, among them rightward disyllabic and rightward final closed syllable reduplication. I follow Buckley (1998) in conflating these into one pattern. Several of Lichtenberk's other patterns might also be combined.

(11)

| | | |
|---|---|---|
| (11) | salaga ‘be long’ | salaga-laga ‘long (sg.)’ |
| | malipi ‘work’ | malipi-lipi ‘work (cont.)’ |
| | mo.ita ‘knife’ | moita-ita ‘cone shell (knife-like)’ |
| | dara ‘blood’ | dara-dara ‘red (blood-like)’ |
| | ʔarai ‘ko ginger’ | ʔarai-rai ‘green (the color of ʔarai leaves)’ |
| | zama ‘tomorrow’ | zama-zama ‘pertaining to tomorrow’ |
| | pile ‘speaker’ | pile-pile ‘speaker’ |
| | malaboŋ ‘flying fox (generic)’ | malabom-boŋ ‘ko flying fox’ |
| | ʔulan ‘desire (vb.)’ | ʔulan-laŋ ‘desirable’ |

But this process does not apply as expected to a certain class of stems (12).

| | | | |
|---|---|---|---|
| (12) | ragogo ‘be warm’ | ragogo-go ‘warm’ | *ragogo-gogo |
| | ʔoʔo ‘be plentiful’ | ʔoʔo -ʔo ‘many, much’ | *ʔoʔo-ʔoʔo |
| | rere ‘like’ | rere-re ‘like *contin.*’ | *rere-rere |
| | lele ‘look for’ | lele-le ‘look for *contin.*’ | *lele-lele |
| | wawa ‘discolored skin’ | wawa-wa ‘white’ | *wawa-wawa |

Interestingly, the members of this class of stems contain repeated material at the end of the form. No examples of such forms have been found to undergo the pattern in (11), where the final foot of the surface form in the left-hand column is repeated in the right. Buckley (1998) argues convincingly that the different behavior of these forms is not due to haplology. Though one could posit a constraint on quadruple syllable repetition, such as (13), phonologists have as of yet found no case in which languages can count up to four. Therefore, one should look elsewhere for an explanation before positing such a constraint.

(13) $*\sigma_i\sigma_i\sigma_i\sigma_i$

Certainly it cannot be the case that Manam prohibits sequences of segmentally identical feet, since this is exactly what is found in ‘sa(làga)(lága)’. Contiguous identical syllables are similarly accepted in the language, as in ‘ragogogo’. If one insisted on treating these facts as a case of haplology, this leaves only a suspicious constraint like (14).

(14) $$* \quad \underset{\sigma_i \qquad \sigma_i}{\overset{\text{Foot}_j}{\diagup \quad \diagdown}} \qquad \underset{\sigma_i \qquad \sigma_i}{\overset{\text{Foot}_j}{\diagup \quad \diagdown}}$$

Rather than posit such a parochial constraint, Buckley argues that the roots in the *ragogo* class are inherently reduplicated. For Buckley, this means the lexical entry for *ragogo* is something like {*rago*, RED}, where RED is the empty morpheme that, in correspondence-theoretic OT, is given segmental content by base-reduplicant correspondence constraints (McCarthy and Prince 1995).

Correspondence (15) is a relation that can hold between strings. Specifically, we are concerned here with a correspondence relation that holds between substrings of the output, which is called base-reduplicant correspondence.

(15) *Correspondence* (McCarthy and Prince 1995:14)
Given two strings $S_1$ and $S_2$, *correspondence* is a relation $\Re$ from the elements of $S_1$ to those of $S_2$. Elements $\alpha \in S_1$ and $\beta \in S_2$ are referred to as correspondents of one another when $\alpha \Re \beta$.

Buckley invokes a conflict between INTEGRITY and REALIZEMORPHEME constraints to generate the observed result. INTEGRITY penalizes cases when a single segment in one string has multiple correspondents in another string. This constraint, in the input-output domain, legislates against the "breaking" of a single input segment into multiple output segments (e.g., a single input vowel becomes a diphthong). In the base-reduplicant domain, it legislates against having more than one copy of a base segment show up in output reduplicants. REALIZEMORPHEME stipulates that a morpheme in the input (even if null in the input) must have some phonological exponent in the output.

(16) a. INTEGRITY: No element of $S_1$ has multiple correspondents in $S_2$ (i.e., For $x \in S_1$ and $w,z \in S_2$, if $x \Re w$ and $x \Re z$, then $w=z$).
b. REALIZEMORPH: An input morphological category is expressed in the output.

In order for this result to hold, however, Buckley suggests that only one set of BR-correspondence indices is available, regardless of the number of RED morphemes in the input and output. Thus he in essence adds to (15) the restriction in (17).

(17) Identity of Indexation (an addendum to *(BR) Correspondence*, not formalized in Buckley 1997):
If $\alpha_i \Re \beta_j$ and $\alpha_i \Re \gamma_k$ then $j = k$ & If $\alpha_j \Re \gamma_i$ and $\beta_k \Re \gamma_i$ then $j = k$.

Given these assumptions, the correct result, ragogo-go (*ragogo-gogo) is optimal (RM = REALIZEMORPH; INTEG = INTEGRITY).[9]

(18) After Buckley's (1997) tableau 24

| | /rago+$\underline{\text{RED1}}$=$\sigma$+$\underline{\underline{\text{RED2}}}$=Ft/ | RM | INTEG | $\underline{\underline{\text{R2}}}$=Ft |
|---|---|---|---|---|
| | a. $r_1a_2g_3o_4$ + $\underline{g_3o_4}$ + $\underline{\underline{g_3o_4g_3o_4}}$ | | ***!* | |
| ☞ | b. $r_1a_2g_3o_4$ + $\underline{g_3o_4}$ + $\underline{\underline{g_3o_4}}$ | | ** | * |
| | c. $r_1a_2g_3o_4$ + $\underline{g_3o_4}$ | *! | | |

[9] (Double) underlining in these examples shows the output strings that correspond to the input RED morphemes. One would presumably need a sort of "morpheme correspondence" to enforce the REALIZEMORPH constraint, but I have not included this here, for the sake of simplicity.

Here the reduplicated form of {*rago*, RED} has two RED morphemes in the input {*rago*, RED1, RED2}, where each RED has particular requirements as to surface prosodic shape. INTEGRITY penalizes multiple reduplication of base segments, thus limiting the number of [go]s in reduplicant strings (the underlined and double underlined strings) and ruling out the expected surface form (18a). REALIZEMORPHEME ensures that a single copy of *go* is made (18b) as an output for RED2, rather than null output (18c). Were the restriction on correspondence relations in (17) not made, one could not prohibit *rago-go-gogo*. The optimality of such a form is shown in the following tableau (19). Here no segment in the output has more than one output correspondent with the same index, and so INTEGRITY is irrelevant.

(19)

| | /rago+$\underline{\text{RED}}$=$\sigma$+$\underline{\underline{\text{RED}}}$=Ft/ | RM | $\underline{\underline{\text{RED}}}$=Ft |
|---|---|---|---|
| ☞ | a. $r_1a_2g_{3,a}o_{4,b}$ + $\underline{g_{3,c}o_{4,d}}$ + $\underline{\underline{g_ao_bg_co_d}}$ | | |
| | b. $r_1a_2g_{3,a}o_{4,b}$ + $\underline{g_3o_4}$ + $\underline{\underline{g_ao_b}}$ | | *! |
| | c. $r_1a_2g_3o_4$ + $\underline{g_3o_4}$ | *! | * |

This account raises an important question, however: There is no evidence that stems in the *ragogo* class are polymorphemic. Therefore, it is somewhat dissatisfying, and probably empirically wrong, to treat their lexical entries are containing a separate RED morpheme. If the correspondence-theoretic RED is to be treated invariantly as a morpheme, we should not be too quick to place inherent reduplication under the RED umbrella (to borrow a phrase from A. Nevins).

If the forms that undergo exceptional reduplication (i.e., those in (12)) are inherently non-linear, however, this pattern is readily explicable. Under this approach, the underlying form of *ragogo* does not contain an empty RED morpheme. Rather, its lexical entry is as follows.

(20) START→r→a→g→o→END Linearized as: [ragogo]

But, although (20) accounts for the surface form *ragogo*, how does it account for unexpected reduplicated forms like *ragogo-go* (**ragogo-gogo*)? Though this may appear to be a case of reduplicative allomorphy, the representation itself, coupled with the assumptions regarding linearization outlined above, provide a natural non-allomorphic solution.

In cases without lexical loops (e.g., *salaga* ⇒ *salaga-laga*), the reduplicative morpheme results in the addition of an immediate precedence relation between the final segment and the onset of the penultimate syllable (21).

(21) START → s → a → l → a → g → a END

Linearized as: [salagalaga]

To compute the placement of the link, the morphology counts back two vowels (or one foot), following precedence relations. This is spelled out explicitly in the following algorithm (22).

(22) Algorithm 1: RedupFinalTwoMoras

```
1   let CurNode =  END;
2   let RedupedMoras = 0;

3   while (RedupedMoras < 2)
4         PrevNode = precededby(CurNode);
5         if Moraic(prevNode)
6            RedupedMoras ++;
7         CurNode = PrevNode;
8         CurNode = Onset(CurNode); % Onset() returns CurNode
9                                   % if CurNode has no onset.

10  AddPrecedence(Imprec(END),CurNode)
```

This formalization simply makes explicit how the grammar searches for the final bimoraic foot of a stem. It begins at the END symbol (line 1), having found zero moras at this point (line 2). While two moras have not yet been found (line 3), the grammar moves backward through the precedence structure (line 4), beginning at END. If a mora is found (line 5), we note this (line 6)[10] and proceed further back in the structure (line 7). Line 8 helps the algorithm deal with onsets (which are non-moraic), and line 10 adds the required precedence link once the final bimoraic foot has been found.

For *salaga*, this algorithm proceeds as follows, and results in the addition of $a \rightarrow l$ to the precedence structure:

(23)

| `RedupedMoras` | `CurNode` |
|---|---|
| 0 | END |
| 1 | a |
| 1 | g |
| 2 | a |
| 2 | l |

Result:
`AddPrecedence(a,l)`

Given this result for *salaga*, we might wonder why we do not observe *ragogo* ⇒ **rago-go-rago*, which is how (24) would appear after linearization. (24) would be formed by adding a precedence relation between the final segment (o) and the segment that is fourth from the end (r), analogous to the *salaga* case.

[10] The symbol ++ should be read “increase by one.”

(24) START→r→a→g→o→END Linearized as: [ragogorago]

But here the principles of linearization, which were spelled out briefly above, play a key role. These principles ensure that, when faced with a choice point in a precedence structure, backward-pointing relations, i.e., "loops", are followed first (5). Thus, from `END` in (25) the computation (26) goes to /o/ and increments the vowel counter by one, then to /g/, then to what precedes it, namely /o/ (recall that back-pointing loops are always taken first), at which point the vowel counter reaches two. The consonant preceding the second counted vowel (/g/ in this case) will be set as the head of the new link, and the final segment of the input, /o/, is set as tail, resulting in the addition of a new precedence relation, as in (27).

(25) START→r→a→g→o→END

(26)

| `RedupedMoras` | `CurNode` |
|---|---|
| 0 | END |
| 1 | o |
| 1 | g |
| 2 | o |
| 2 | g |

Result:
`AddPrecedence(o,g)`

(27) START→r→a→g→o→END Linearized as: ragogogo

Since only two loops are present, linearization will result in three copies of the syllable [go], rather than four. This result is achieved by making explicit the algorithm for finding the final two moras (or bimoraic foot) of the stem and by applying the general principle of Completeness, which ensures that "loops" are always taken first, to all algorithms that apply to precedence structures. As Gene Buckley (p.c.) notes, this approach is preferable to the correspondence-theoretic account since it does not coopt the empty RED *morpheme* for the representation of a *mono*-morphemic root. Furthermore, this approach explains the unexpected behavior of the *ragogo* class; this is not a case of haplology, allomorphy, or the result of limits on the base-reduplicant correspondence relation. Rather, it is the result of normal computation over precedence structures. The words in the *ragogo* class reduplicate as *ragogo-go* since, due to the stems' inherent non-linearity, the final monomoraic syllable is itself the final bimoraic foot.

## 4. Poly-Morphemic Non-Linearity

In the last section I argued that lexical representations of phonology may contain non-linear precedence, and that this fact allows us to explain the sometimes unexpected behavior of such lexical items. I now turn to cases of polymorphemic non-linearity. That is, cases of non-linear precedence that result from the spell-out of multiple morphemes. In section 4.1 I provide a detailed analysis of each of three reduplication patterns in Lushootseed (Bates et al. 1994; Urbanczyk 2001). I then go on in section 4.2 to analyze cases where more than one of these reduplicative morphemes appears in the same form. I argue that the monostratal OT analysis of Urbanczyk (2001) is insufficient for modeling all of the patterns when the full range of vowel alternations is considered. Urbanczyk (2001) divides Lushootseed morphemes into two classes (affix and root) in order to control for the shape of these patterns. However, this division incorrectly predicts patterns of stress-based [i]-insertion. My solution is a different classification of morphemes and the possibility of multiple linearization in some cases. That both [i]-insertion and multiple linearization are sensitive to these classes I take as support for this analysis.

### 4.1. Three Patterns of Reduplication in Lushootseed

I will discuss three patterns of reduplication in Lushootseed and their interaction with each other and with certain vowel insertion properties of the language.[11] These three patterns are exemplified in (28).[12]

(28)

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

Fitzpatrick (to appear) outlines a system for the lexical specification of affixes using precedence relations. In this system, lexical entries for affixes follow the schema in (29). This schema simply encodes the position of an affix, in the same way as a lexical diacritic or a morpheme-specific alignment constraint would.

(29) PREC→PHON→FOLL

a. PREC: Segment that immediately precedes the affix

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

[12] [ʔí-ʔi-bəʃ-tx$^w$] contains the causative transitive suffix *-tx$^w$*. Several different morphemes are spelled out using the OOC pattern, including *out of control*, *particularization* (ia), and *human count* (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'

b. PHON: Phonological content 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 ($\text{Seg}_1$ and $\text{Seg}_f$), first and last vowel ($\text{V}_1$ and $\text{V}_f$) and consonant ($\text{C}_1$ and $\text{C}_f$), certain prominence-based anchor points such as the stressed vowel ($\acute{\text{V}}$), and the segments that immediately follow and immediately precede these points (ImFoll(X) and ImPrec(X)). Other possible anchor points might also include include initial and final foot and syllable. This approach to anchor points owes much to Alan Yu's (2003) work on possible infixal insertion points.[13] In this schema, a prefix would be represented as in (30a) and a suffix as in (30b).

(30) a. START → /PREFIX/ → Initial Segment ($\text{Seg}_1$)
b. Final Segment ($\text{Seg}_f$) → /SUFFIX/ → END

That is, a prefix with segmental content /PREFIX/ immediately follows the START symbol and immediately precedes $\text{Seg}_1$. A suffix that is pronounced /SUFFIX/ immediately follows the final segment ($\text{Seg}_f$) and immediately precedes the END symbol.

An infix has PREC and FOLL values that are neither initial nor final. For example, Atayal animate actor focus, as described in Yu 2003, has the form in (31).

(31) $\text{C}_1$ → m → ImFoll($\text{C}_1$)
a. qul ⇒ q**m**ul
b. hŋuʔ ⇒ h**m**ŋuʔ

Reduplication fits nicely into this approach to the lexical specification of the phonological shape of vocabulary items. When PREC follows FOLL in the stem to which a morpheme's phonological exponent attaches, linearization leads to the repetition of phonological material. Examples are shown in (32).[14]

(32) a. Pangasinan *plural*: ImFoll($\text{C}_1$) → $\text{C}_1$
niog ⇒ ni-niog (From: START→n→i→o→g→END)
amigo ⇒ a-mi-migo (From: START→a→m→i→g→o→END)
b. Indonesian *plural*: $\text{Seg}_f$ → $\text{Seg}_1$
buku ⇒ buku-buku (From: START→b→u→k→u→END)

[13] See Fitzpatrick (to appear) for arguments that the same set of anchor points is used for specification of infixal position and reduplicant shape.
[14] When PHON is null I represent the vocabulary item as PREC→FOLL.

This unification of affix types: prefixes, suffixes, infixes, and reduplication, departs from current work following Steriade (1988), as well as work in correspondence-theoretic OT, which treat reduplication as total repetition of a form, with truncation of certain parts of one copy. This also departs from Raimy's (2000) proposal that reduplication is uniformly the result of readjustment.

A preliminary specification of the vocabulary items for Lushootseed reduplication (28) that lead to these patterns is given in (33).[15]

(33)
a. DIM: $V_1 \rightarrow Seg_1$
b. DIST: ImFoll($V_1$)$\rightarrow C_1$
c. OOC: ImFoll($\acute{V}$) $\rightarrow \acute{V}$

For any form to which a single reduplicating exponent has been added, such as (34), Completeness will force material to be repeated in the output form. As above, I compare this form with impossible linearizations in a tableau (35).

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

(35)

| Input: (34) | 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 (28)). The underlying forms for [ʔíb-ʔib-əʃ] 'walk all about' (DIST) and [ʔ-íb-ib-əʃ] 'pace back and forth' (OOC) are shown in (36).

(36)
a. [DIST √ʔibəʃ]: START → ʔ → i → b → ə → ʃ → END
Linearized as: [ʔ-íb-ib-əʃ]

b. [OOC √ʔibəʃ]: START → ʔ → í → b → ə → ʃ → END
Linearized as: [ʔ-íb-ib-əʃ]

Since the appearance of schwa is predictable in Lushootseed, I follow Bates (1986) and other scholars of Salish in positing that schwas are not present an underlying forms. I assume that some set of phonological constraints (PHONCON), which crucially dominates DEP(ə), ensures their

[15] 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 truly prefixed forms, the vocabulary item could be reanalyzed as ImFoll($V_1$)$\rightarrow V_1$.

insertion in the correct contexts.[16] Furthermore, Lushootseed stress appears on the leftmost non-schwa vowel. I formalize this with a gradient alignment constraint StressLeft. If all vowels are schwa, stress appears on the left-most schwa. The ranking shown in (37) for *bədáʔ* 'child' ensures this outcome. Dep(¬ə) is shorthand for all Dep-V constraints, where V is any vowel but schwa. The ranking Dep(¬ə), PhonCon, *ə́ ≻ StressLeft, Dep(ə) does not allow epenthesis of non-schwa vowels (like [i]). Therefore, stress is non-initial.

(37)

| /bdaʔ/ | Dep(¬ə) | PhonCon | *ə́ | StrLeft | Dep(ə) |
|---|---|---|---|---|---|
| bdáʔ | | *! | | | |
| ☞ bədáʔ | | | | * | * |
| bə́daʔ | | | *! | | * |
| bídaʔ | *! | | | | |

I will assume that schwas are inserted into the base root form before reduplicative morphemes are spelled 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'.

(38) [ooc √čx̌]: start → č → ə → x̌ → end

Linearized as: [čə́x̌əx̌] 'smash up, crack to pieces'

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 (33). These forms require the addition of an elsewhere case to our dist and dim vocabulary items.[17]

(39)
a. dim: If $\sigma_1$ is CV, then $V_1 \rightarrow C_1$; *elsewhere*: $C_1 \rightarrow C_1$
b. dist: If $\sigma_1$ is CV, then ImFoll($V_1$)$\rightarrow C_1$;
*elsewhere*: ImFoll($C_1$)$\rightarrow C_1$
c. ooc: ImFoll($\acute{V}$) $\rightarrow \acute{V}$

With this revision in hand, we can examine how the analysis extends to reduplication of CCV- stems. For example, √č'ƛ'aʔ surfaces as [č'ƛ'áʔ] 'rock'. The underlying form for [dist √č'ƛ'aʔ] = č'ƛ'č'ƛ'áʔ 'rocks scattered about' is given in (40). Linearization and (lack of) schwa insertion is shown in (41).

[16] Schwa insertion can also be stated in a rule-based format. My 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 I am exploring a new representational vocabulary that could, in principle, be adopted in a variety of frameworks.

[17] If reference to the second C of a form is possible, as suggested to me by Chuck Cairns (p.c.), then dist could be simply $C_2 \rightarrow C_1$.

(40) [dist √č'ƛ'aʔ]: START→č'→ƛ'→a→ʔ→END

(41)

| (40) | Dep(¬ə) | PhonCon | StrLeft | Dep(ə) |
|---|---|---|---|---|
| ☞ č'ƛ'č'ƛ'áʔ | | | | |
| č'əƛ'č'əƛ'áʔ | | | *!* | ** |
| č'íƛ'č'əƛ'aʔ | *! | | | * |

However, while no vowels besides schwa and lexically present vowels appear in OOC and DIST forms, diminutives formed from CCV- (and CVV-) initial stems 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- (42b) and CVV-initial DIM forms (42c).

(42) 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'

Based on these differences with respect to [i]-insertion, I 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 (43a)). Class B, on the other hand, is evaluated using the slightly different ranking in (43b), where Dep(i) is demoted below StressLeft.

(43) a. Dep(¬ə) , Max$_{IO}$, PhonCon, *ə́ ≻ StressLeft, Dep(ə)
b. Dep(¬ə/i) , Max$_{IO}$, PhonCon, *ə́ ≻ StressLeft ≻ Dep(i) ≻ Dep(ə)[18]

Since (43b) allows i-epenthesis, this ranking leads to the following derivation of [DIM √č'ƛ'aʔ] ⇒ [č'íč'ƛ'aʔ]. The derivation of (42c) is identical in all relevant respects.

(44) START → č' → ƛ' → a → ʔ → END

(45)

| (44) | PhonCon | *ə́ | StrLeft | Dep(i) | Dep(ə) |
|---|---|---|---|---|---|
| ☞ č'íč'ƛ'aʔ | | | | * | |
| č'ə́č'ƛ'aʔ | | *! | | | * |
| č'ič'ƛ'áʔ | | | *! | * | |
| č'č'ƛ'áʔ | *! | | | | |

[18]Ranking of Dep(i) and Dep(ə) is not required here. However, I maintain the Dep(i) ≻ Dep(ə) ranking proposed by Urbanczyk (2001).

Turning to Cə- stems, the constraint ranking for Class B predicts correctly that [i] can be inserted for stress. We could treat this as insertion of the Place features of [i] into the placeless schwa. If schwas had not yet been inserted in the stems to which DIM is applied, the derivation of Cə-initial forms would proceed in the same way. However, I am assuming schwas are inserted in root forms, based on the fact that the placement of OOC often depends on the placement of schwas.

(46) START → t→ə → l → a → w → END

(47)

| (47) | $\text{MAX}_{\text{IO}}$ | *ə́ | STRLEFT | DEP(i) | DEP(ə) |
|---|---|---|---|---|---|
| ☞ títəlaw | | | | * | |
| tə́təlaw | | *! | | | |
| tətílaw | | | *! | * | |

In (47) $\text{MAX}_{\text{IO}}$ is evaluated existentially (Struijke 2000; Raimy and Idsardi 1997). That is, $\text{MAX}_{\text{IO}}$ 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 $\text{MAX}_{\text{IO}}$ in any of the candidates in (47). However, existential evaluation in this case is not an extra stipulation imposed on faithfulness constraints, as it is in Struijke (2000). Rather, existential evaluation of MAX is a result of the fact that there is no base or reduplicant string in the output form under my assumptions. $\text{MAX}_{\text{IO}}$ is defined as normal (48), and existential evaluation results with no additional stipulations.

(48) Every segment of the input has a correspondent in the output.
i.e., $\forall \text{S}_x \exists \text{S}_y[\text{S}_x \in \text{INPUT} \rightarrow [\text{S}_y \in \text{OUTPUT} \wedge \Re(\text{S}_x,\text{S}_y)]]$, where $\Re$ is the IO-Correspondence relation. (McCarthy and Prince 1995:16)

Urbanczyk (2001) proposes that DIM and OOC are in a class together (for Urbanczyk, these are "affixes"), with DIST in a different class (Urbanczyk's "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) (*kʷ'əq* ⇒ *kʷ'əqíq*), whereas all other OOC examples show no i-insertion (see (49)). Similarly, DIST forms do not allow i-insertion:

(49)

| | | | |
|---|---|---|---|
| | OOC | | |
| | bək$^{w}$’ ‘all’ | bə́k$^{w}$’ək$^{w}$’ ‘they all’ | *bək$^{w}$’ík$^{w}$’/*bík$^{w}$’ək$^{w}$’ |
| | čəx̌ ‘split’ | s-čə́x̌əx̌ ‘smashed up’ | *s-číx̌əx̌/*s-čəx̌íx̌ |
| | d$^{z}$ək$^{w}$’ ‘travel’ | d$^{z}$ə́k$^{w}$’ək$^{w}$’ ‘...changing all the time’ | *d$^{z}$ík$^{w}$’ək$^{w}$’/*d$^{z}$ək$^{w}$’ík$^{w}$’ |
| | d$^{z}$əx̌ ‘move’ | d$^{z}$ə́x̌əx̌ ‘visit’ | *d$^{z}$íx̌əx̌/*d$^{z}$əx̌íx̌ |
| | DIST | | |
| | ʔəɬəd ‘eat’ | ʔə́ɬʔəɬád ‘many eat’ | *ʔíɬʔəɬad/*ʔəɬʔíɬad |
| | bədáʔ ‘child’ | bə́dbədáʔ ‘children’ | *bídbədaʔ/*bədbídaʔ |
| | bəd ‘thin’ | bə́dbəd ‘thin (board) | *bídbəd/*bədbíd |

Under Urbanczyk’s constraint ranking (50) (compiled from various tableaux in Urbanczyk (2001:ch. 4)), however, we would expect [i]-insertion and stress shift in all of the forms in (49a), contrary to fact. The candidate *bək$^{w}$ík$^{w}$* is not considered in Urbanczyk (2001). (For Urbanczyk, the second -VC sequence is the reduplicant, while the first is part of the base.) Here IO-Dep-Rt 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.)

(50)

| {bək$^{w}$, ooc} | IO-Dep-Rt | *ə́ | StrLeft | BR-Dep-Afx |
|---|---|---|---|---|
| ☞ bək$^{w}$ík$^{w}$ | | | * | * |
| bə́k$^{w}$ək$^{w}$ | | *! | | |
| bík$^{w}$ək$^{w}$ | *! | | | |

Urbanczyk’s incorrect prediction comes from failure to correctly identify the classes of affixes. Urbanczyk’s root and affix classes were required in order to predict reduplicant shape. However, these classes do not correctly predict [i]-insertion. In the next section we turn to further motivation for the Class A/Class B distinction.

### 4.2. Multiple Reduplication: The Case for Intra-Derivational Linearization

> “Cyclicity is a stipulation in derivational algorithms. It is a *nice* stipulation, but still a stipulation, whose need we might question.” (M. Brody, in MIT colloquium talk, 25 April 2003)

A key demonstration of linearization comes in the introduction of multiple reduplication. Lushootseed allows forms that contain multiple reduplicative affixes. I will discuss the phonological forms that correspond to the morphological structures in (51) and some permutations of these structures.

(51) a. [dist [ooc $\sqrt{}$root]] / [ooc [dist $\sqrt{}$root]]
b. [dim [ooc $\sqrt{}$root]]

c. [DIM [DIST $\sqrt{\text{root}}$]]
d. [DIST [DIM $\sqrt{\text{root}}$]]

Turning to (51a), 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 (52).

(52) a. [DIST [OOC $\sqrt{\text{root}}$]]
b. [OOC [DIST $\sqrt{\text{root}}$]]

Examples of (52a) have semantics that could be paraphrased as something like 'many X's involved in random, ineffectual action' (cf. Bates et al. (1994:xvii)). Examples include $saq'^{w}aq'^{w}saq'^{w}$ 'many flying around, wheeling in the sky' (from $saq'^{w}$ 'fly') and $g^{w}\acute{a}x^{w}ax^{w}g^{w}ax^{w}$ 'a lot of strolling about' (from $g^{w}\acute{a}x^{w}$ 'go for a walk').

It is also important to note that if [DIST [OOC $\sqrt{\text{root}}$]] forms were derived by cyclic, serial application of each reduplication pattern, one would expect outputs like $saq'^{w}saq'^{w}aq'^{w}$, contrary to fact (53). Therefore, these forms must be derived non-cyclically.

(53)

| Base | saq'$^{w}$ |
|---|---|
| OOC | saq'$^{w}$aq'$^{w}$ |
| DIST | saq'$^{w}$saq'$^{w}$aq'$^{w}$ |
| Output | saq'$^{w}$saq'$^{w}$aq'$^{w}$ |

Turning to (52b), it seems that OOC has little or very subtle effect on the semantics of DIST forms. Examples include *bálalbali* '(suddenly) I am forgetful' (cf. *báli* 'forget', *bəlbáli* 'I am forgetful') and $g^{w}\acute{a}dadg^{w}ad$ 'discuss, converse, talk over' (cf. $g^{w}ad$ 'talk', $g^{w}\acute{a}dg^{w}ad$ '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 linearization does not apply between the spell-out of the two morphemes. Lack of linearization is already motivated by the failure of the derivation in (53). Thus both hierarchical structures in (52) will lead to the phonological representation in (54) and linearization as in (55).

(54) START→b→a→l→i→END

(55)

| | Input: (54) | 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 (55c), which fails to include the $l \rightarrow b$ relation in the output. Economy rules out (55d) since this form results from unnecessarily crossing the $l \rightarrow a$ relation more than once. Shortest rules out (55b) since this form would result if the $l \rightarrow b$ relation were crossed before the shorter $l \rightarrow a$ 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 $l \rightarrow a$ link is shorter. Recall that, descriptively, Shortest is operative when there are two backward-pointing precedence relations, choosing the one whose endpoint is closer in terms of transitively preceding the source (an "inner loop").[19]

Due to the invariant nature of linearization, the same surface form will arise in a non-cyclic derivation, regardless of the morpheme hierarchy. In contrast, under a cyclic derivation, only the [OOC [DIST √root]] form is predicted to surface with the observed shape.

(56) Table 1

| | Cyclic | Non-cyclic |
|---|---|---|
| [DIST [OOC √]] | * | ✓ |
| [OOC [DIST √]] | ✓ | ✓ |

[DIM [OOC √root]] forms demonstrate 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 (57), due to a local preference for back-pointing links, $a \rightarrow d$ is spelled out before $a \rightarrow y$.

(57) START→d→a→y→ʔ→END

(58)

| | Input: (57) | SHORTEST | COMPLETE | ECONOMY |
|---|---|---|---|---|
| ☞ a. | dá-day-ayʔ | | | d-aa-y |
| b. | dáy-a-dayʔ | | (a-d)! | aday |

Here candidate (58b) 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 (59a), rather than (59b).

(59) START→s→a→xʷ→END

a. sásaxʷsaxʷ
b. *sáxʷsasaxʷ

[19]In this, and ideally all cases, "distance" can be computed over the output tape; that is, over already linearized material.

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').[20]

(60) START→b→ə→d→a→ʔ→END

(61)

| | Input: | $\text{MAX}_{\text{IO}}$ | *ə́ | STRESSLEFT | DEP(i) |
|---|---|---|---|---|---|
| ☞ a. | bíbədbədaʔ | | | | * |
| b. | bəbədbədáʔ | | | *!** | |
| c. | bəbídbədaʔ | | | *! | * |
| d. | bə́bədbədaʔ | | *! | | |

Unlike the situation we saw with DIST-OOC forms, in which both hierarchical orders 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/]] surfaces as [bíbibədaʔ] 'young children', rather than *[bíbədbədaʔ]. This [$C_1$V$C_1$V$C_1$V...] pattern is observed in numerous other forms:

(62)

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

The failure of the non-cyclic derivation is summarized in the following table:

(63) Table 2

| | Cyclic | Non-cyclic |
|---|---|---|
| [DIM [DIST √]] | ✓ | ✓ |
| [DIST [DIM √]] | ✓ | * |

In order to capture Table 1 (56) and Table 2 (63), I will draw on the distinction between Class A and Class B morphemes that was established in the last section based on [i] insertion and stress. Specifically, I propose that linearization applies between Class B morphemes (e.g., DIM) and Class A morphemes (e.g., DIST, OOC). This assumption does not affect the results for forms that contain DIM and OOC (57), since both [DIM [OOC root]] and [OOC [DIM root]] forms would result in the observed pattern, regardless of

[20]The linearized form is evaluated using the constraint ranking of the highest morpheme, in this case DIM.

cyclicity.

(64) Table 3

| | Cyclic | Non-cyclic |
|---|---|---|
| [Dim [Ooc $\surd$]] | ✓ | ✓ |
| [Ooc [Dim $\surd$]] | ✓ | ✓ |

However, under this assumption, the derivation of the [DIST [DIM ROOT]] form [bíbibədaʔ] proceeds as in (65). Compare this to the linearization of (60).

(65) 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:[21] [bíbibədaʔ]

Thus we find further evidence for a distinction between Class A and Class B morphemes: besides allowing [i]-insertion and stress shift, Class B morphemes also trigger linearization, whereas Class A morphemes do not appear to trigger linearization.

It is possible that linearization applies quite generally between 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." (emphasis not in original)

The facts regarding the differential behavior of Class A and B reduplication patterns with respect to stress allow us the theoretical option of positing "stratal linearization" (to use Mester's terminology). Linearization of this sort allows us to deal with the apparent lack of intervening linearization in OOC-DIST forms, while at the same time accounting for variable morpheme ordering in DIM-DIST forms and the corresponding differences in surface form. However, though the Class A/Class B distinction

[21] 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]).

in Lushootseed allows us this analytical option, one would like independent evidence for the possibility of intra-derivational linearization. I turn in the next section to clearer evidence of this sort from Nitinaht (Wakashan), where I also revive the theme of inherent non-linearity in lexical forms.

## 5. Inherent Reduplication meets Multiple Linearization

The Wakashan language Nitinaht makes extensive use of suffixation for both "lexical" and "inflectional" purposes. That is, Nitinaht contains not only familiar agreement and tense suffixes ("inflectional" suffixes), but also suffixes that are predicates and modifiers ("lexical" suffixes).

Stonham (1994:40) observes that in Nitinaht "[a] certain subset of the lexical suffixes require that certain effects on the shape of the root be manifested, either length on the root vowel, reduplication of some portion of the root, or some combination of the above."[22] That is, when certain suffixes are added to a form, we observe concomitant reduplication of the initial CV of the root or vowel lengthening, or both. There are roughly forty suffixes of this sort, exemplified in (66).[23]

| (66) | | Unreduplicated | Reduplicated |
|---|---|---|---|
| | a. | ƛ'ic-ak | (CV) ƛ'iƛ'ic-ak'uk |
| | | 'whiteness' | 'flour' |
| | | (white+DUR) | (white+resembles) |
| | b. | yaq-'aq | (√V:) yaːq-'ałuk-'aq |
| | | 'the who one...' | 'the one who ... likes' |
| | | (REL-DEF) | (REL-like...-DEF) |
| | c. | hita-quł | (CV + √V:) hihi:ta-ʔac'uł |
| | | 'face' | 'foot' |
| | | (LOC-at face) | (LOC-at the foot) |
| | d. | baƛ-a: | (CV:, i.e. CV+√V:) baːbaƛ-'ad |
| | | 'tied' (tie-DUR) | 'tied at the top' (tie-at the end) |
| | e. | yaq-quy | (CV:+√V:) ya:ya:q-tiš-a:w-quy |
| | | 'the one which...' | 'the who to use as guide' |
| | | (REL-COND) | (REL-use...as guide-should-COND) |

When suffixed with durative *-ak*, *ƛ'ic* 'white' surfaces without repetition of the initial CV, while when predicative *-ak'uk* 'resembles ...' is added to *ƛ'ic*, the initial CV must be repeated (*ƛ'ic-ak'uk). But suffix-induced stem-alteration is not limited to CV reduplication. Various types

[22] My large debt to Stonham's work should be obvious in this section.

[23] No minimal pairs could be found for the CV:+√V: pattern due to the relative rarity of the suffixes involved. However, comparison with forms like (i) show that the reduplication-triggering affix is *-tiš* 'use ... as guide'.

(i) a. ʔu-kʷaqł-ʔa łicłic 'It's called a mat' (REFER+call...+ 3INDIC mat')
b. ʔuːʔuː-uk-tiš-aʔƛ-'a 'X was using it as a guide' (REFER+DUR+use...as a guide+now+3INDIC)

of stem-alteration are required by suffixes of this sort. (66a) shows CV reduplication, while (66b) shows only lengthening of the stem vowel. (66c) shows both CV reduplication and lengthening of the second surface vowel. (66d) shows CV reduplication with lengthening of the first surface vowel. Finally, (66e) shows a combination of all possibilities: CV reduplication with lengthening of both vowels. If the root vowel is already long (i.e., lexically specified as long), then (66a) and (66d) both look like (66e).

Stem changes are triggered by these morphemes even when other suffixes intervene linearly between the reduplication-triggering suffix and the root.

| (67) | Unreduplicated | Reduplicated |
|---|---|---|
| | ƛ'uqʷ-či:-a:ʔdɬ | (CV:) ƛ'u:ƛ'uqʷ-a:ʔdɬ-a:p |
| | wide-pull-along a long object | wide-along a long object-very |
| | | 'X's legs are too big' |

The patterns in (66) reduce to all logical combinations of CV reduplication and vowel lengthening.

| (68) | | No reduplication | CV reduplication |
|---|---|---|---|
| | No lengthening | All other cases | (a) |
| | Root V: | (b) | (c) |
| | Reduped V: | N.A. | (d) |
| | Both V: | N.A. | (e) |

Affixes that require reduplication may co-occur in Nitinaht. However, a key difference between Lushootseed and Nitinaht is that when affixes of the type shown above co-occur in Nitinaht, only one repetition of the first CV surfaces (69) (reduplication-triggering affixes are underlined).

| (69) | | | |
|---|---|---|---|
| | a. | ƛ'u:ƛ'uqʷ-a:ʔdɬ-a:p | 'X's legs are really big' |
| | | *ƛ'u:ƛ'u:ƛ'uqʷ-a:ʔdɬ-a:p | |
| | b. | sa:sa:tq-'aqsiɬ-a:p | 'X's eyes were really itchy' |
| | | *sa:sa:sa:tq-'aqsiɬ-a:p | |
| | c. | ba:baɬ-aski-yabɬ-a:p | X is really cold on the shoulders' |
| | | *ba:ba:baɬ-aski-yabɬ-a:p | |
| | d. | ƛ'i:ƛ'i:daqʷ-aqs-ib-k'uk | 'X resembles whale's baleen' |
| | | *ƛ'i:ƛ'i:ƛ'i:daqʷ-aqs-ib-k'uk | |

Furthermore, a combination of the requirements of each suffix (reduplication, vowel length) surface in the multiply-suffixed form, suggesting that these are separate requirements. Due to this fact, and the fact that the full range of reduplication-plus-vowel length combination is possible, I will focus only on reduplication.

The facts in (69) suggest that there is a requirement in Nitinaht against multiple reduplication – a sort of haplology. Morphological haplology, in which an affix or clitic is absent when the stem to which it attaches already contains the phonological material that would be added by the affix, is attested in many languages and language families (Stemberger 1981).

Interestingly, certain words in Nitinaht, such as *kakaw'ad* 'killer whale', contain repeated initial CV strings. One can ask, as we did for Thao and Manam in section 2, whether this is a case of accidental repetition of segments, or of inherent reduplication. As with Manam, there is reason to believe that these words are inherently reduplicated: When the reduplication-triggering affix -*ataχ* 'hunt...' appears with this form, no additional repetition of the initial CV is possible.

(70) a. kakaw'ad 'killer whale' (*kaw'ad)
b. kakaw'ad+ataχ 'hunting killer whales'
c. *kakakaw'ad+ataχ

This fact shows two things. First, Nitinaht has inherently reduplicated lexical items, just as we saw for Manam. Second, the haplology constraint in Nitinaht cannot be based solely on affixation. That is, we could not say that in Nitinaht certain affixes contain the following instructions.

(71) Reduplicate the initial CV if the stem does not already contain a reduplication-triggering affix.

Rather, Nitinaht must have a ban on the addition of instructions to reduplicate if such instructions are already present. Crucially, this restriction must pay no attention to the source of the instructions. One approach that fits these constraints would draw on the representational possibilities afforded us by the nonlinear treatment of reduplication. A reduplication-triggering affix such as -*ataχ* would be listed as follows, where the structural description for reduplication holds throughout the language.

(72) a. Suffix: -ataχ
b. Reduplication:
(i) SC: $V_1 \rightarrow C_1$
(ii) SD: $V_1 \nrightarrow C_1$
(i.e., $V_1 \rightarrow C_1$ unless this relation already exists)

This theory of haplology in Nitinaht does not ban triple repetition of segmental material in a surface string. (Such a ban would not be surface true, as we will see.) Instead, this condition must have access to abstract precedence structures. My claim is that the language allows the introduction of a new precedence relation only if this precedence relation does not already exist.

This predicts that if an affix is not able to see that there was reduplication at an earlier level of the derivation, multiple reduplication will be possible. Precisely such a situation is possible if linearization intervenes between two reduplicating affixes.

In fact, this prediction is born out. Nitinaht allows other types of reduplication besides the type triggered by the suffixes shown above. For example, the distributive surfaces as initial CV reduplication without an accompanying suffix.

(73)

| Unreduplicated | Reduplicated |
|---|---|
| ʔiːč’-ib<br>old-thing | ʔi-ʔiːč’-ib<br>DIST-old-thing ‘a bunch of old ones’ |
| ƛ’iχ-ʔaː<br>red-on the rocks | ƛ’i-ƛ’iχ-ʔaː<br>DIST-red-on the rocks<br>‘There’s red ones all over the rocks.’ |

The distributive morpheme can co-occur with reduplication-inducing suffixes in Nitinaht,[24] and also with inherently reduplicated stems (74).

(74) ka-kakaw’ad-ataχ ‘hunting killer whales here and there’
DIST+killer whale+hunt...

Here *-ataχ* is an affix of type (a) from (66), which should trigger CV reduplication, but does not in the case of inherently reduplicated *kakawad*. But this form also contains the distributive, which is observed as initial CV reduplication. Both appear in the same form, resulting in a $C_1V_2C_1V_2C_1V_2$... form. This confirms that Nitinaht does not have a ban on three consecutive identical CV sequences. We do not observe a sequence of four identical syllables, as we would expect from inherent reduplication + suffix-triggered reduplication + distributive reduplication, but nor do we observe just a single repetition, as would be expected if there were a general ban on repetition in surface strings in Nitinaht.

So why is exactly this much multiple reduplication possible here, but not in (69), where multiple reduplication-triggering suffixes did not result in multiple reduplication? The solution is that *-ataχ* and DIST are different types of morphemes. The stem /kaw’ad/ is inherently reduplicated. Therefore, when *-ataχ* is added, it does not contribute an extra loop to the representation. However, by the time DIST applies, the representational evidence that [kakaw’ad] was reduplicated has been erased. How? As with Lushootseed multiple reduplication, linearization provides the answer. The derivation of [ka-kakawad-ataχ] is given in (75).

(75)
a. Root: START→k→a→w’→a→d→END
b. Add affix: a → t → a→χ
START→k→a→w’→a→ d →END
c. Reduplication not triggered by affix: SD not met.
d. Linearize: START→k→a→k→a→w’adataχEND
e. Add DIST: START→k→a→k→a→w’adataχEND

[24]Clearly, one would want an example of such a form. Unfortunately, I have been unable to find one in my sources (Haas and Swadesh 1932; Stonham 1994).

f. Linearize: [kakakaw'adataχ]

This derivation exemplifies a case of lexically-stored reduplication, precedence-based (rather than surface string-based) haplology, and multiple levels of affixation and linearization. Nitinaht permits double reduplication only in those instances where each repetition is the result of affixal or inherent non-linearity at different levels of the grammar, with linearization applying between these levels.

But do we have any independent reason to think different types of affixes exist in Nitinaht, as we did for Lushootseed? Haas and Swadesh (1932) and other work establish a difference between "stem suffixes", such as those in (66) and "word suffixes", such as the distributive. A clustering of phenomena justify the separation of Nitinaht affixes into these two groups.

First, stem suffixes can occur in various orders, and can appear multiple times in a single form (though this of course affects the meaning), while word-suffixes appear only once in a form (at the maximum), and appear rigidly outside of stem suffixes.

On more phonological grounds, epenthetic vowels appear before glottalized consonants in stem suffixes but not in word suffixes. Furthermore, an epenthetic vowel appears between the last stem suffix in a form and the following word suffix if the former is consonant-final and the latter is consonant-initial. Stonham (1994) also cites lenition and "hardening" phenomena, described in Haas and Swadesh (1932) and Boas (1947) as further evidence for a distinction between morpheme types in Nitinaht.

With these differences in hand, we have evidence for the distinction between DIST (a word suffix) and reduplication-triggering suffixes like -*ataχ* (stem suffixes) that provide for the possibility of intervening linearization.

## 6. Summary

Once non-linear precedence structures are allowed in the grammar, we can ask at what levels such representations can appear. I proposed that non-linear representations can appear not only due to the spell-out of polymorphemic structures, but even in lexical representations. One argument for this came from Manam final foot reduplication. Patterns of multiple reduplication in Lushootseed provided further exemplification of non-linear precedence and linearization. I presented a complete analysis of the at times quite complex patterns of Lushootseed multiple reduplication and argued that stress and [i]-insertion in Lushootseed provide evidence for a classification of two types of affixes (A and B) and the application of linearization between affixes of different types. I also argued that this two-level analysis of Lushootseed is preferable to the single-level model of Urbanczyk (2001) since the latter incorrectly predicts [i]-insertion in OOC forms.

These two proposals – the availability of non-linear representations in the lexicon and the possibility of multiple levels of non-linearity and linearization – come together in Nitinaht. I showed that Nitinaht has a type of reduplicative haplology that is active not only in affix-related reduplication, but in lexical reduplication as well. If reduplication-triggering affixes come from different levels, however, we see that no haplology affect arises. This can be explained if, as in Lushootseed, linearization applies between one type of affix (the *ata$\chi$* type) and another (exemplified by DIST). Since this type of haplology cannot be treated as a surface ban on repeated CV sequences, this is evidence for a more abstract representation of reduplication. Such a treatment of reduplicative haplology would be difficult or impossible in a monostratal model or a model where haplology was based on linear surface strings.

## References

Bates, Dawn. 1986. 1986. In *The Proceedings of the BLS 12*, 1–13.

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

Blust, Robert. 2001. Thao triplication. *Oceanic Linguistics* 40:324–335.

Boas, Franz. 1947. Kwakiutl grammar with a glossary of suffixes. In *Transactions of the American Philossophical Society*, ed. Zellig Harris and Helene Boas Yampolsky, volume 37.

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

Buckley, Eugene. 1998. Integrity and correspondence in Manam double reduplication. In *The Proceedings of NELS 28*, 1, 317–346.

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

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, Mass.: MIT Press.

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

Fitzpatrick, Justin. to appear. A typology of possible affix types. In *Papers from EVELIN 1*, ed. Andrés Salanova. MIT Working Papers in Linguistics.

Fitzpatrick, Justin, and Andrew Nevins. 2004. Linearizing nested and overlapping precedence in multiple reduplication. In *The Proceedings of the 27th Penn Linguistics Colloquium*, 75–88.

Haas, Mary, and Morris Swadesh. 1932. A journey to the other world. *International Journal of American Linguistics* 7:195–208.

Kenstowicz, Michael. 1994. *Phonology in generative grammar*. Malden, Mass.: Blackwell Publishers.

Lichtenberk, Frantisek. 1983. *A grammar of Manam*. University of Hawaii Press.

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.

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 78th LSA.

Raimy, and Idsardi. 1997. A minimalist approach to reduplication in OT. In *The Proceedings of NELS 27*.

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

Stemberger, Joseph. 1981. Morphological haplology. *Language* 57:791–817.

Steriade, Donca. 1988. Reduplication and syllable transfer in Sanskrit and elsewhere. *Phonology* 5:271–314.
Stonham, John. 1994. *Combinatorial morphology*. John Benjamins.
Struijke, Caro. 2000. Why constraint conflict can disappear in reduplication. In *The Proceedings of NELS 30*.
Urbanczyk, Suzanne. 2001. *Patterns of reduplication in Lushootseed*. Garland Outstanding Dissertations.
Yu, Alan. 2003. The morphology and phonology of infixation. Doctoral Dissertation, University of California, Berkeley.
Zuraw, Kie. 2000. Aggressive reduplication. Presented at 74th LSA Meeting.

Department of Linguistics and Philosophy
MIT, Bldg. 32-D866
77 Massachusetts Ave.
Cambridge, MA 02139
USA

*jfitzpat@mit.edu*