Type-C Emphatic Reduplication in Sakha

May Pik Yu Chan*

1. Introduction

Sakha (Yakut) is a Turkic language spoken in the Republic of Sakha (Yakutia) in Russia. While it shares many phonological processes with other Turkic languages such as vowel harmony, Sakha is unique in that it preserves many aspects of older Turkic phonology. This includes contrastive vowel length distinctions (Krueger 1962) and a productive emphatic reduplication process. In the present study, we adopt an Optimality Theory (OT) analysis to capture the latter. Emphatic reduplication refers to partial reduplication involving (quasi-)fixed segmentism in Sakha adjectives. This means that part of the root is reduplicated, and the reduplicant also contains sound(s) that are not part of the root. In Sakha, the main form of emphatic reduplication applied to adjectives for intensification consists of a linker /p/ sound, which is the focus of this work. We aim to capture formally the choice of the linker /p/ and optional OCP effects in reduplication, as well as vowel shortening effects.

2. Background and Data

There are multiple forms of reduplication in Sakha. In nouns and verbs, full reduplication of the whole stem is found with the particle da inserted. Examples include kinige da kinige ‘what a book’ (book PARTICLE book) and kyler da kyler ‘laughs a lot’ (laugh-AOR PARTICLE laugh-AOR). This also occurs with adjectives, as in uraax da uraax ‘very far’ (far PARTICLE far) (Vinokurova 2005).

Reduplication without the particle is possible with adjectives. Vinokurova (2005) reports that among adjectives in Sakha, reduplication can apply to the initial syllable or the entire word. Both cases of reduplication intensify the adjective. Regarding full reduplication, the internal structure of the adjective affects whether it is possible to modify the vowel of the adjective to mark further intensification. When the adjective contains a bound root and a suffix, the vowel of the suffix can be modified. On the other hand, when adjectival root words that do not contain any suffixes get intensification, the reduplicant also contains sound(s) that are not part of the root. In Sakha, the main form of emphatic reduplication applied to adjectives for intensification consists of a linker /p/ sound, which is the focus of this work. We aim to capture formally the choice of the linker /p/ and optional OCP effects in reduplication, as well as vowel shortening effects.

There are multiple subtypes of partial reduplication in Sakha. Examples above show different linker sounds between the partially reduplicated forms.

Reduplicated form

<table>
<thead>
<tr>
<th>Adjective</th>
<th>Reduplicated form</th>
<th>Modified reduplicant</th>
</tr>
</thead>
<tbody>
<tr>
<td>*uʃugei ‘good’</td>
<td>uʃugei- uʃugei ‘very good’</td>
<td>uʃugiiu- uʃugei ‘very very good’</td>
</tr>
<tr>
<td>ʃugas ‘close’</td>
<td>ʃugas-ʃugas ‘very close’</td>
<td>ifuugu-ʃugas ‘very very close’</td>
</tr>
<tr>
<td>*ʃkii ‘dense’</td>
<td>*ʃkii-ʃkii ‘very dense’</td>
<td>*ʃkii-ʃkii</td>
</tr>
<tr>
<td>kieŋ ‘wide’</td>
<td>*kieŋ-kieŋ ‘very wide’</td>
<td>*kieŋ-kieŋ</td>
</tr>
</tbody>
</table>

There are adjectives that allow for both full and partial reduplication. For example, the full reduplicated form of uʃugei ‘good’ can be either uʃugei- uʃugei or ubus- uʃugei, with the identical meaning of ‘very good’. Similarly, kieŋ ‘wide’ can be reduplicated as kip- kieŋ or kieŋ- kieŋ to mean ‘very wide’. However, not every adjective allows both forms of reduplication.

There are multiple subtypes of partial reduplication in Sakha. Examples above show different linker sounds between the partially reduplicated forms ubus- uʃugei ‘very good’ and kip- kieŋ ‘very wide’, namely /s/ and /ʃ/. Stachowski (2014) distinguishes six types of emphatic reduplication among Turkic languages. This includes type-C, type-CV, type-CVC, type-V, type-ma and type-rV. Among these forms of partial reduplication, four exist in Sakha. In the present work, we focus on ‘Type-C reduplication’, which refers to reduplication that has a single consonant as a linker, such as /p/. Since the linker consonant /p/ is generally invariable, this is an example of fixed segmentism (Alderete et al. 1999). There are three main observations that this study aims to address: (i) the general pattern of fixed segmentism in type-C reduplication, (ii) optional OCP effects, and (iii) monophthongization or vowel shortening effects in the reduplicant. A traditional OT (Prince and
Smolensky 1993) approach will be used to capture these generalizations.

Data from the present study are gathered from two main sources. A consultant, who is a native speaker of Sakha in her twenties, provided judgements. Other data are obtained from Stachowski (2014), who compiled them from several other sources, mainly from a time range of 1907-1930. Thus, judgements by our consultant do not always correspond to those in the literature. Data regarding non-adjectival reduplication are obtained from Vinokurova (2005). The compiled adjectival data may be found on Open Science Framework (OSF): https://osf.io/gyqfj/.

3. The Phonological Distribution and Status of /p/

Sakha contains eight monophthongs, including the phones /i/, /y/, /ɯ/, /u/, /e/, /œ/, /ɔ/, /a/, which contrast in [+/- high] and [+/- round] features. There are four diphthongs involving a high vowel and a non-high vowel that agrees in frontness and backness (e.g. /ie/, /yœ/, /ua/ and /uɔ/). The monophthongs all have contrastive long counterparts. Some but not all consonants in Sakha have contrastive geminate counterparts as well.

In native Sakha vocabulary, bilabial stops in word initial position may only be voiced (Krueger 1962). Exceptions include onomatopoeic words and loans (e.g. paxaj ‘Ugh!’ or park ‘park’). Intervocically, only the geminate /p:/ and voiced /b/ are contrastive (see 2a). Word-finally, singleton voiceless /p/ may occur, but not voiced /b/. Examples (2b) and (2c) show that word final /p/ or /p:/ in the root is realized as voiced /b/ intervocically, regardless of what vowel-initial morpheme is attached to the root.

(2) a. sippax ‘blunt, dull’ sibax ‘plaster’
    xappax ‘cover, lid’ xabax ‘bladder’
    b. sap ‘thread’ saba ‘his thread’
    tap ‘to hit’ tabar ‘he hits’
    c. kiliep ‘bread’ kiliebi ‘bread-acc’
    mɔrkup ‘carrot’ morkułu ‘carrot-acc’

Also noteworthy, is that both the literature and our consultant’s orthography consider the linker as a <p> and not a <b> when the reduplicant is followed by an onset-less root, suggesting that the linker /p/ in emphatic reduplication is treated as a coda.

(3) Root Reduplicated form Meaning of root
    uhun up-uhun long
    ɯraas up-ɯraas clean
    afinity ap-afﬁnity hungry

To capture this generalization that the linker /p/ is considered as a coda despite its intervocalic status, the following constraints are proposed.

(4) *C[+VOICE] ]σ : Voiced consonants do not align with the right edge of a syllable.
(5) *VPV: Intervocalic /p/ is disallowed.
(6) ALIGNL(STEM, σ): The left edge of the stem must align to the left edge of a syllable boundary.

<table>
<thead>
<tr>
<th>RED-uhun</th>
<th>*C[+VOICE] ]σ</th>
<th>ALIGNL(STEM, σ)</th>
<th>*VPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) up-uhun</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b) ub-uhun</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c) u-buhun</td>
<td></td>
<td>*!</td>
<td></td>
</tr>
</tbody>
</table>

Tableau 1. Lack of intervocalic voicing in reduplication.

By ranking *C[+VOICE] ]σ above *VPV, we capture the generalization that linker /p/ is preferred over intervocalic voicing, which occurs in other intervocalic environments created by suf-
fixation. In Tableau 1, (b) is eliminated as /b/ is voiced and is at the syllable edge, it violates the higher ranked \[C[+Voice]\] \(\sigma\). (c) does not have a voiced segment at the right syllable edge, however the change in syllabification is disfavored by ALIGNL(STEM, \(\sigma\)), as the stem and the left syllable edge is misaligned. Consequently, the winning candidate is up-uhun.

4. The Morphological Status of the Linker /p/

In the following section, we examine the morphological status of /p/. We start by reviewing existing work that captures fixed segmentism cross-linguistically. Fixed segmentism in OT has been addressed in two main ways, including a phonological basis or a morphological basis of reduplication (Alderete et al. 1999). We then discuss both types of fixed segmentism and argue that type-C emphatic reduplication in Sakha is best captured by a morphological account of fixed segmentism.

4.1. A Phonological Basis of Fixed Segmentism

The phonological basis of fixed segmentism is intertwined with the notion of the emergence of the unmarked. In OT, there are two main types of constraints, faithfulness and markedness constraints (Prince and Smolensky 1993). The former governs corresponding input and output structures, and the latter evaluates the output structures. These constraints are inherently in tension with one another, and it is the ranking of faithfulness and markedness constraints that constitutes a grammar. Among faithfulness constraints, correspondences can be between input and output forms (underlying-surface mapping) or between base and reduplicants (McCarthy and Prince 1994). Considering IO (input-output) correspondences and BR (base-reduplicant) correspondences, a general schema of the emergence of the unmarked may be generalized as follows:

(7) Reduplicative TETU constraint ranking schema (McCarthy and Prince 1994):

\[ \text{FAITH-IO} >> \text{MARKEDNESS-CONSTRAINT} >> \text{FAITH-BR} \]

Under this schema, the structure that is prohibited by the markedness constraint could generally surface in the language. This is because \text{FAITH-IO} ranks over the markedness constraint. However, in the case of reduplication, the unmarked structure is preferred over the structure of the base surfacing, as \text{FAITH-BR} is ranked below the markedness constraint. Consequently, an unmarked structure that usually does not override underlying forms is preferred in the reduplicant. This is one type of the emergence of the unmarked (TETU).

In Alderete et al. (1999)’s general discussion of reduplication, they explain that fixed segments derived from TETU are defaults under a phonological analysis of emphatic reduplication. This is because the TETU ranking causes the reduplicant to alter the base which results in less marked structures. This is not to say that fixed segmentism must be invariant; instead, variation in realization may occur depending on the language’s constraint hierarchy (Alderete et al. 1999).

In order to determine default segments, a universal markedness constraint hierarchy is required. One important markedness ranking hierarchy to consider for the present study would be a theory to determine the default place of articulation. For consonants in general, place of articulation is an important factor in markedness. A common default segment would be the glottal stop, though coronals may also realize as default (Alderete et al. 1999). The universal place-markedness hierarchy was proposed (Prince and Smolensky 1993) to capture this generalization of coronals being less marked cross-linguistically, and subsequent work has extended this generalization to include pharyngeal/laryngeal sounds (Lombardi 1997). The ranking in (8) asserts that a pharyngeal or laryngeal sound is generally least marked, followed by coronal sounds. Labials and dorsal sounds are more marked and are not expected to pattern as default segments.

(8) Place markedness hierarchy (Lombardi 1997, McCarthy and Prince 1994):

\[ *\text{PL/LAB}, *\text{PL/DORS} >> *\text{PL/COR} >> *\text{PL/PHAR} \]

Against this background, Alderete et al. (1999: 357) describe a number of properties for phonological fixed segmentism. First, there should be a subset relation between the reduplicant’s inventory and that of the entire language. Second, there should be a cross-linguistic correlation between restrictions on the reduplicant’s inventories and restrictions on the entire language; this is a natural consequence of universal markedness constraints. Third, there should be consistency
between fixed segmentism and independent evidence of default status in the language. Finally, there should be potential for conditioned variability of ‘fixed’ segmentism, which occurs when there are multiple forms of BR faithfulness constraints, with some being more specific than others.

Although a TETU analysis is compatible with many cases of fixed segmentation, it does not generate the correct predictions for Sakha, mainly because /p/ is not a default segment. Bilabials are not unmarked cross-linguistically, and there are no independent reasons for /p/ to be treated as an unmarked default segment in Sakha, which is a condition for the TETU analysis. On the contrary, singleton voiceless bilabial stops are very restricted in the language, as they do not appear in intervocalic nor onset positions in native Sakha vocabulary. Under the place-markedness hierarchy (Prince and Smolensky 1993), labials are generally marked, and work by de Lacy (2006) has also shown that labials and dorsals cannot be promoted in the place-markedness hierarchy. Therefore, /p/ cannot be a default segment and the phonologically based analysis of fixed segmentism is problematic for capturing Sakha data.

Another observation regarding markedness in Sakha type-C reduplication concerns the coda status of /p/. NOCODA is a markedness constraint that favors open syllables, as codas are marked. However, the fixed segment in the reduplicated form involves a marked coda in Sakha. Even if NOCODA was ranked lower alongside base-reduplicant faithfulness constraints, the grammar would still generate incorrect predictions. Equal ranking of NOCODA and MAX-BR would not be able to explain why a coda is preferred over an onset. While ranking NOCODA below MAX-BR would predict the correct syllable structure, ranking a markedness constraint below MAX-BR would be theoretically inconsistent with TETU. Furthermore, it still would not explain why changing the original consonant to /p/ is preferred.

4.2. A Morphological Basis of Fixed Segmentism

The morphological melodic overwriting approach considers the fixed segment to be an alternative affix on its own. In the following section, we first introduce the morphological analysis of fixed segmentism with an example from English from Alderete et al. (1999). We then show why this successfully accounts for the pattern of fixed segmentism in Sakha.

An overwriting analysis was first proposed by McCarthy and Prince (1996), and its OT implementation was then proposed by Alderete et al. (1999). In overwriting, the identity of the fixed segment in reduplication is listed by morphology and not generated by phonology (McCarthy and Prince 1996, Yip 1992). An example of overwriting in English would be “schm” words, such as table-schmable. Here, the /ʃm/ onset is considered as a prefixal morpheme, and the identity of this fixed /ʃm/ is not phonologically determined. Three of the main properties of affixal overwriting as organized by Alderete et al. (1999) are summarized below.

The first property of affixal overwriting is that fixed segments display faithfulness properties of an affix. This means that strings can contain marked structures, such as the English /ʃm/, which is not phonotactically natural in the language. This also means that contrasts among overwriting strings are possible, especially with sounds that do not correspond to default segments. For example, in Hindi, overwriting with /w/- and /s/- occurs without phonological conditioning in the choice of segment (Singh 1969). Second, overwriting strings have alignment properties like those of affixes and are aligned to either the left or right periphery. Third, the overwriting string may alternate by suppletion or allomorphy. This means that they are context-sensitive enough to participate in phonological processes that may affect other affixes, but the choice of alternate form is not phonologically predictable. For example, dissimilatory constraints may cause alternation of overwriting affixes (Yip 1998): in Telugu /gi-/ alternates with /pi-/ when the root starts with a /gi/ (e.g. gilaka-pilaka ‘rattle’).

Alderete et al. (1999: 356) account for the overwriting analysis of fixed segmentism that affects the reduplicant in OT by ranking input-output faithfulness constraints above base-reduplicant faithfulness constraints, which captures the data in English. In the following tableau, recreated from Alderete et al. (1999), (b) violates MAX-IO because the prefix /fm/- has been deleted. (c) and (d) also have MAX-IO violations because the onset /t/ from the base in both forms have been deleted. (a) still has one violation of MAX-BR because the onset /t/ that was copied from the base is deleted. However, it is the best candidate among the four as FAITH-BR is low-ranked.

<table>
<thead>
<tr>
<th>/table-RED-fm/</th>
<th>MAX-IO</th>
<th>MAX-BR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
We now turn back to Sakha to see how the melodic overwriting account of fixed segmentism applies. To begin with, the fixed segment /p/ does match the aforementioned properties of affixes. The first property is that the fixed /p/ is marked both in the language and generally in the place markedness hierarchy, and it therefore requires faithfulness constraints to surface unchanged. For example, /p/ does not become voiced in intervocalic positions. This is unlike other processes of affixation subject to bilabial stop voicing, like adding an accusative marker to roots with bilabial codas (e.g. *kiliebi* ‘bread’-ACC).

A second property of affixes is that contrasts among overwriting strings are possible. The following three examples from Stachowski (2014) show some reduplicated forms that use non-/p/ linkers. These are not phonologically conditioned changes. For example, it is not that a root onset /s/ triggers the reduplicant linker to become a /j/ instead of /p/, as the examples *sojuu* ‘a cry, weep’ and *siikej* ‘crude, raw’ have the same environment but different linker output forms. There is also no obvious relationship between the segments /s/, /j/, /n/ and /p/, suggesting that these linkers are better considered as contrastive forms of the overwriting string.

A third property is edge alignment of the affix. Although the linker itself looks like an infix, the coda status of /p/ suggest that it is aligning to the right edge of the reduplicant. A constraint that captures this would be $\text{ALIGNR(FixedSEG, PRWD)}$, which states that the fixed segment and the prosodic word should align at the right edge. By extension, this means that the underlying order of the morphemes should be maintained. Together with $\text{FAITH-IO}$ and $\text{FAITH-BR}$ constraints, the reduplication pattern could be formalized. We assume /p/ to be part of the underlying form here, taking affixal status and therefore subject to input-output constraints. In tableau 3, candidate (b), which shares two segments with the root, violates $\text{MAX-BR}$ only once, as only /a/ lacks a base-reduplicant correspondence. However, it also violates $\text{MAX-IO}$ once, because the linker /p/, which is subject to input-output constraints, is deleted. Since $\text{MAX-IO}$ is higher ranked, candidate (b) fails. Candidate (c) has all the necessary segments in the output; however, the linker morpheme is not right aligned to the reduplicant (or left aligned to the base), which causes it to violate the high ranked $\text{ALIGNR(FixedSEG, PRWD)}$ constraint. Consequently, the correct output form (a) wins.

<table>
<thead>
<tr>
<th>Root</th>
<th>Reduplicated form</th>
<th>Meaning of root</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>tastun</em></td>
<td><em>tast-tastun</em></td>
<td>outer, outsider</td>
</tr>
<tr>
<td><em>sojuu</em></td>
<td><em>soj-sojuu</em></td>
<td>a cry, weep</td>
</tr>
<tr>
<td><em>siikej</em></td>
<td><em>sin-siikej ~ sip-siikej</em></td>
<td>crude, raw</td>
</tr>
</tbody>
</table>

(9) Root | Reduplicated form | Meaning of root |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><em>tastun</em></td>
<td><em>tast-tastun</em></td>
<td>outer, outsider</td>
</tr>
<tr>
<td><em>sojuu</em></td>
<td><em>soj-sojuu</em></td>
<td>a cry, weep</td>
</tr>
<tr>
<td><em>siikej</em></td>
<td><em>sin-siikej ~ sip-siikej</em></td>
<td>crude, raw</td>
</tr>
</tbody>
</table>

Nonetheless, it is possible to consider the reduplicant as a prosodic word and therefore not subject to bilabial stop voicing.
The final property of overwriting affixal strings is that they may alternate by suppletion or allomorphy. This also appears to be the case for Sakha when optional OCP effects emerge. When the root begins with a bilabial, the linker /p/ alternates into other forms for dissimilatory effect. This is discussed in section 5.

A natural question that arises from the overwriting analysis is what status the fixed segment /p/ has. There are three logical interpretations. (i) /p/ is part of the base, (ii) /p/ is part of the reduplicant, and (iii) /p/ is part of neither the base nor the reduplicant. The original definition of the base would be the strings immediately following the prefixed reduplicant or the strings immediately preceding a suffixed reduplicant (McCarthy and Prince 1993b). Under this definition, /p/ could be considered part of the base, as it on the immediate right edge to the reduplicative prefix. This possibility was not discussed in Alderete et al. (1999). Instead, the tableau in Alderete et al. (1999) assumes that /p/ is part of the reduplicant (i.e. table-schmable and not table-schmable). As correctly pointed out in a footnote in Alderete et al. (1999), both /p/ being part of the reduplicant or not are possibilities, though no evidence in favor of one definition over the other was known.

Overall, this section has shown that Sakha type-C reduplication is best captured by a melodic overwriting account of reduplication. Under this analysis, input-output faithfulness constraints ranking over base-reduplicant faithfulness constraints would suffice. This is because /p/ is considered an affix and is subject to input-output based constraints as well. On the other hand, a phonological analysis that relies on markedness constraints to determine the fixed segment is not consistent with Sakha phonology. In the following sections, we will continue to capture more details of Sakha type-C reduplication, including OCP effects and vowel shortening or monophthongization.

5. OCP Effects in Sakha Type-C Reduplication

As mentioned above, Sakha has optional OCP effects in the form of the linker. Note, however, that these data are not consistent across the literature and our consultant, such as in the translation of the word bosxo ‘straight, upright/free’. This is unsurprising, as most of the data from the literature were collected around or before 1930, and our consultant is a young speaker of modern Sakha. For example, our consultant does not recognize the words budaan ‘misty, foggy’ and buťax ‘unclean’.

Regarding the alternation in the linker sound, data from the literature suggests that there are OCP effects. When root begins with a bilabial plosive, the linker sound alternates to /s/ or /r/. In the case that the root begins with a bilabial nasal, the linker sound alternates to stops with other places of articulation. For similar sounding sounds like buťaan ‘slow’, our consultant reduplicates the word as bu̯p-buťaan ‘very slow’. Two consecutive bilabial stops with different voicing features are considered legal by our consultant, which provides further evidence that the reduplicant with the linker sound is a separate phonological word. Generally, our consultant consistently uses /p/ as a linker, as in (11).

<table>
<thead>
<tr>
<th>(10)</th>
<th>Stachowski (2014)</th>
<th>From our consultant</th>
</tr>
</thead>
<tbody>
<tr>
<td>bɔsxo</td>
<td>bɔs/-r-bɔsxo</td>
<td>straight, upright</td>
</tr>
<tr>
<td>budaan</td>
<td>bus-budaan</td>
<td>misty, foggy</td>
</tr>
<tr>
<td>manjaŋ</td>
<td>mak/-manjaŋ</td>
<td>white</td>
</tr>
</tbody>
</table>

Focusing only on the OCP effects that are observed, OCP effects in Sakha may also be analyzed in terms of suppletive allomorph selection; specifically, ‘phonologically conditioned suppletive allomorphy’ (PCSA). A main reason for this is that no synchronic rule exists to cause p → s, r, k, t in the language. Therefore, while the motivation for alternation is phonological, the alternation

\[2\] Note that in the case of \(bɔsxɔ\), which is translated as ‘free’ by our consultant, the ‘reduplicated’ form to mean the intensified ‘very free’ is bu̯ɔr-bɔsxɔ. Noticeably, the added mora (/u/ instead of /ɔ/) is not a phonological process; instead, bu̯ɔr is another lexical word meaning ‘sand’. Therefore, bu̯ɔr-bɔsxɔ means that something is as costless as sand, and is a lexicalized compound and not a productive reduplication form.
itself is better explained by listed allomorphy. In a traditional OT approach to the phonology-morphology interface, this relationship could be captured by the ranking of phonological and morphological constraints. This is described in McCarthy and Prince (1993a, 1993b), where ‘P constraints’ (phonological constraints) outrank ‘M constraints’ (morphological constraints). This schema captures the general observation that some allomorphs are chosen by phonological conditions. This ‘P >> M’ schema has been used to capture reduplication data.


Paster (2006) argues that in the case of reduplication, the general phonological constraint (or constraint family) would be BR-CORRESPONDENCE. This enforces base and reduplicant similarities. As for the morphological constraint, MORPHEXPR could prevent the realization of the reduplicant, if it does not form phonologically legal outputs.

(13) MORPHEXPR: An input morphological category is expressed in the output.

However, if it is /p/ alone as an affix that is alternating, this general BR-CORRESPONDENCE >> MORPHEXPR does not capture the OCP effect. Instead, the phonological principle that applies to the linker /p/ affix itself would be OCP. While MORPHEXPR would still be an important constraint to prevent the deletion of the linker, another M-constraint is required. We name this constraint LINKER = /P-/:

(14) LINKER = /P-/: The reduplicant is linked to the base by the affix /p-/.

Note that in the analysis so far, we have not resolved whether the linker should be treated as a suffix to the reduplicant, a prefix to the base, or an infix to both. In the present formulation, we place the dash to the right of the linker, suggesting a prefix status, though an infix account would also be consistent with the analysis below. Note also that the formulation of this constraint is not novel, and echoes Kager’s (1996) ‘GENITIVE = /-n/’ constraint for Djabugay genitive suffix allomorphy. The relative ranking of OCP >> LINKER = /P-/ (>> MORPHEXPR) could at least partially capture our data.

<table>
<thead>
<tr>
<th>/RED-{p-, t-, k-, s-}-ma^\text{\textbackslash n}an/</th>
<th>OCP-LABIAL</th>
<th>LINKER = /P-/</th>
<th>MORPHEXPR</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) mat-ma^\text{\textbackslash n}an</td>
<td>✓</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>(b) mak-ma^\text{\textbackslash n}an</td>
<td>✓</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>(c) map-ma^\text{\textbackslash n}an</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(d) ma-ma^\text{\textbackslash n}an</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Tableau 4. P >> M model for OCP effects in Sakha reduplication

While the P >> M model captures OCP effects that disfavor /p/ outputs when preceding a labial, this analysis makes no predictions with respect to whether it should be /t/ or /k/, or /s/ or /s/. Adding constraints like LINKER = /K/ or LINKER = /T/ would not resolve this issue because there are no phonological patterns among the adjectival roots that may condition the surface form of the linker, except that it should not be a bilabial if the root begins with a bilabial. In other words, while PCSA is optimizing and preexisting P constraints may partially explain the data, it cannot fully capture OCP effects in Sakha type-C reduplication.

Paster (2006) predicts related issues in the P >> M model, based on her survey of over 600 grammars, and concludes that a subcategorization model may be a better model to capture PCSA cross-linguistically. Specifically, she shows that the P >> M model makes four predictions: (i) that PCSA is ‘optimizing’ and preexisting P constraints may capture it, (ii) that PCSA is sensitive to surface and not underlying forms, (iii) that there is bidirectionality of phonological conditioning between stem and affix, and (iv) conditions on allomorphs could be located anywhere on the word.
Prediction (i) poses issues for data from Sakha, as OCP itself is optimizing, but not the selection of non-/p/ linkers. A consequence of this is that output conditioning may be insufficient to capture our data, as *mar-mayän and mat-mayän equally satisfy OCP in the output. Furthermore, Sakha does require adjacency of the linker affix to the root, and it is the root that conditions the affix and not the other way round. In contrast to the P >> M model, a subcategorization model would predict that (i) the PCSA does not have to be optimizing, (ii) that PCSA can respond to phonological constraints, and (iii) a stem could be irregular, such that linkers other than /p/ could be chosen depending on subcategorization requirements. In the case of Sakha, this means that OCP could force irregularities when the base has an initial-bilabial sounds, though irregularities could exist for any base.

(15) budaan bus-budaan misty, foggy
tastuŋ tas-tastuŋ outer, outsider

We can thus hypothesize that subcategorization may be more suited for OCP effects found in Sakha reduplication, as it does not require phonological optimization, and it is input conditioned. This means that it is possible for linker allomorphy to result from different subcategorization requirements for each linker-form. As the phonological requirements for the [r/t/k/s-] form are more specific than the requirements of the [p-] form, the [p-] form is treated as the elsewhere case. Based on the few examples of OCP in Sakha available, we can hypothesize that [s-] is used before bilabial plosives, [t/k-] are used before bilabial nasals, and [r-] may be treated as exceptions. This distribution could be formulated as follows:

(16) Construction A
[t/k- [♯C[LAB, +NAS]]adj base ]linker

Construction B
[s- [♯C[LAB]]adj base ]linker

Construction C
[p- [ adj base ]linker

Construction A indicates that a non-labial stop version of the linker morpheme should be used when the phonological constraining element is a labial nasal. Construction B shows a less restrictive subcategorization requirement, in which [s-] could be attached to labials that are not nasals. Construction C is the least restrictive, showing that [p-] is attached elsewhere. This non-OT account roughly captures the limited data on OCP that we have, despite not being able to perfectly capture the distribution of [r-]. Further work and data would be required to test the phonological generalizations made.

Overall, this section shows that Sakha type-C reduplication has optional OCP effects. We attempted to capture OCP effects under the P >> M model and subcategorization model by assuming phonologically conditioned suppletive allomorphy, which is consistent with the melodic overwriting analysis developed in section 4. In the following section, we show that while an allomorphic analysis better captures the distribution of the linker /p/, TETU effects still exist in the reduplicant.

6. Vowel Shortening and Monophthongization

In this final section, we capture the effects of vowel shortening in the reduplicant. Apart from the compound example of buor-boxxa ‘very free, as free as sand’, the reduplicant always has a monomoraic short vowel, regardless of whether the root has a diphthong or long vowel. When the root has a long vowel, such as siyës ‘thin’, the vowel in the reduplicant is shortened. Therefore, only sip-siyës is grammatical, but not *siip-siyës. For diphthongs, only the first mora can be part of the reduplicant (i.e. only the high vowel of the diphthong). Consequently, kip-kiëŋ ‘very wide’ is a grammatical form, but not *kep-kiëŋ or *këp-kiëŋ.

(17) Root Reduplicated form Meaning of root
kieŋ kip-kiëŋ wide
kyœx kyp-kyœx blue
This pattern is consistent with a TETU analysis of reduplication. Specifically, having a core syllable as a CV-minimal prefix would be unmarked. We capture the pattern by following the TETU constraint ranking schema (FAITH-IO >> MARKEDNESS-CONSTRAINT >> FAITH-BR). To posit a constraint that favors a minimal core syllable, *VV may disfavor long vowels and diphthongs when ranking MAX-IO above *VV and MAX-BR below *VV. In the following tableau, all candidates violate *VV at least once due to the root. On top of that, (b) and (c) both violate *VV once more, as the reduplicant contains another diphthong. Candidate (c) also violates MAX-IO because the linker /p/, which is considered as its own suffix is deleted, however it does not violate MAX-BR, as the reduplicant is a replica of the base. Candidate (d) adds another /p/ segment making it geminate, thus violating DEP-IO. Candidate (a) does not have high ranked Input-Output violations but instead violates the low ranked MAX-BR twice and is the winning candidate.

(18) *VV: long vowels and diphthongs are disallowed

<table>
<thead>
<tr>
<th>Candidate</th>
<th>RED-p-kieŋ</th>
<th>MAX-IO</th>
<th>DEP-IO</th>
<th>*VV</th>
<th>MAX-BR</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td>kip-kieŋ</td>
<td></td>
<td>*</td>
<td>e, η</td>
<td></td>
</tr>
<tr>
<td>(b)</td>
<td>kiep-kieŋ</td>
<td></td>
<td>**!</td>
<td>η</td>
<td></td>
</tr>
<tr>
<td>(c)</td>
<td>kieŋ-kieŋ</td>
<td>p!</td>
<td>**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d)</td>
<td>kipp-kieŋ</td>
<td>p!</td>
<td>*</td>
<td>e, η</td>
<td></td>
</tr>
</tbody>
</table>

Tableau 5. Vowel shortening

Overall, this section shows that although TETU could not account for the fixed segment itself, a TETU-based analysis is sufficient to capture the vowel shortening effect in Sakha reduplication. Taken together, fixed segmentism in Sakha is better explained with phonological and morphological processes combined.

7. Conclusion

The present work uses a classical OT approach to capture type-C reduplication in Sakha. The coda status of the linker /p/ is a marked structure in Sakha and cross-linguistically. However, we have shown that the linker /p/ is still considered as coda, and that the reduplicant and /p/ may be a separate phonological word. Based on the analysis that /p/ is a coda, we have shown that a phonological approach to fixed segmentism does not capture the data. On the contrary, the linker /p/ is better considered as an affix under a melodic overwriting approach. We also capture optional OCP effects via constraint ranking and discuss variants to the bilabial linker in terms of phonologically conditioned allomorphy. Lastly, we show that although a TETU account of reduplication is insufficient overall to capture the data, the vowel shortening effect in the reduplicant could be accounted for via TETU.

Future work may also consider other forms of Sakha reduplication. For example, it is conceivable that type-CVC emphatic reduplication may also be analyzed under a morphological account, as string sequences like /-bVs/ are not inherently unmarked either. Similarly, the effects of the adjective’s structure affecting the grammaticality of degree modification in full adjectival reduplication is also sensitive to morphology. Formal accounts of these processes remain a topic for future investigation.

References


