Enhancement and Diachrony in Reduplication

Jason D. Haugen
Oberlin College, USA

Abstract. This paper argues that, in addition to the well-known process of phonological reduction (grammaticalization), phonological enhancement is also a possible path of development for new reduplicative morphemes. By tracing the historical development of a new reduplicative morpheme in Tohono O’odham (Uto-Aztecans), it is shown that a novel reduplicant was innovated by means of vowel-lengthening (mora augmentation) for a special culturally-important class of “marked” plurals. The enhancement strategy of vowel lengthening for plurals allowed the marked class to be maintained in contrast to unmarked short vowel plurals on the one hand, but also maintained an inherited contrast with a heavy syllable reduplicant that triggers gemination in order to indicate ‘distributive’ or ‘repetitive’ on the other. Thus, this paper supports Urbanczyk (2005)’s contention that the notion of contrast enhancement can be extended to the morphological domain of reduplication, and shows how such enhancement serves as a means of avoiding homophony.

Keywords. Reduplication, diachrony, contrast enhancement, homophony, Tohono O’odham, Uto-Aztecans

1. Introduction

While the phenomenon of reduplication has received much attention from various (primarily synchronic) theoretical viewpoints in recent years, relatively little work has explicitly addressed the issue of the diachronic development of reduplicative morphemes. In such work that does address diachrony the prevailing view seems to be that reduplicative morphemes (hereafter, reduplicants) typically reduce (i.e. phonologically erode) over time. The prototypical proponents of this view, Bybee, Perkins and Pagliuca (1994), regard the development of reduplicative morphemes to be just another instantiation of grammaticalization more generally, so I will refer to this view as the grammaticalization approach to diachrony in reduplication.1 In Bybee et al.’s original formulation, the phonological reduction of reduplicants coincides with the broadening (or bleaching) of their semantic functions over time.

For example, Bybee et al. hypothesize the “fullest, most explicit form of reduplication, total reduplication, to be the originating point for all reduplications,

1 In subsequent work others have applied Bybee et al.’s grammaticalization approach in discussions of the diachronic development of reduplicative morphemes, including Niepokuj (1997) for Indo-European and other language families.
with the various types of partial reduplication as reductions and thus later developments from this fullest form” (p. 166). In support of this claim they present evidence from multi-pattern reduplication in Trukese (Austronesian) which purports to show a fossilized continuum of relative reduction of form: total reduplication, syllabic reduplication, and initial consonant doubling. They claim that the oldest form, total reduplication, retains the original semantic function of marking iterative or continuative aspect. The second historical stage, a syllable reduplicant reduced from total reduplication, marks habitual aspect, and the third stage, which involves the most reduced form, is an initial consonant doubling that functions as an intransitivizer. Bybee et al. thus claim that “the modern situation in Trukese suggests that reduplication can grammaticize more than once in the history of a language, and that forms produced by successive waves of grammaticization can co-exist, although the form and meaning of each one identifies its age” (p. 173).

We will discuss problematic cases for this particular view below. However, there are clear cases of reduplicants reducing in form without concomitant semantic bleaching. Garrett (2001), for example, presents a historical analysis of the development of variation in the reduplication patterns of Yurok (Algic), and reconstructs a CVCV- reduplicant for the morpheme called the repetitive. The other attested reduplication patterns used for the repetitive (CVC or CV:) are shown to derive from the historical loss of the second vowel. In word forms that retain the CVCV- pattern this syncope process has been blocked in a predictable environment: namely, contexts where the syncope would result in an otherwise illicit word-medial cluster of two sonorant consonants.

Similarly, semantic bleaching can occur without reduction of form. Shibasaki (2005) accounts for decategorization of reduplicated forms in the history of Japanese, where full reduplication transitioned from a verbal to an adverbal function, with an overlap in the two functions in Early Old Japanese but a predominance of the adverbal function in Late Middle Japanese. (Both functions are rare in Modern Japanese).

Finally, some comparative studies have proposed a reduction in reduplicative form without any reference to semantics at all. For example, in the most comprehensive survey of the comparative morphology of the Uto-Aztecan languages, Langacker (1977: 129-30) notes a variety of reduplication patterns in these languages, including CVCV-, CVC-, CV?-., CV~, and CV, each of which are attested for different functions in various Uto-Aztecan languages. Langacker conjectures that the CVCV-reduplicant may have been the original form, with the reduction of this leading to smaller reduplicants in different Uto-Aztecan languages over a period of time.

All of the studies above fall under the rubric of what I am calling here the grammaticization approach because they account for the development of new reduplicative morphemes by invoking either phonological reduction or semantic

---

2 See also Hurch and Mattes (2005), who question Bybee et al.’s proposal that all partial reduplication must derive from full reduplication. The account of the enhancement of reduplicative forms that I will sketch in section 3 below could be taken to support Hurch and Mattes over Bybee et al. on this point.
bleaching or both. The purpose of the present paper is to make two new contributions
to our understanding of the diachronic development of reduplicative morphemes.
First, in contrast to cases like Trukese, where different reduplicants have developed in
order to mark new semantic contrasts, I will show that different reduplicative
morphemes can converge in a process that I will call functional reduction. In such
cases, a semantic contrast between multiple reduplicative morphemes is lost, and the
reduplicants thus may appear in free variation for a single function. This is the
subject of Section 2. Then, in Section 3 I will show that in addition to
grammaticalization, as discussed above, another strategy is available for languages to
exploit in creating new reduplicants to mark novel semantic contrasts. I will call this
strategy enhancement of form, as it involves expanding the phonological shape of an
existing reduplicant, rather than reducing it, in order to mark a new semantic
category. In this section I will illustrate the multi-pattern reduplicative system of
Tohono O'odham (Uto-Aztecan), which includes two distinct patterns of heavy
syllable reduplication. I will argue that the heavy syllable reduplicant that triggers
gemination is inherited from an earlier stage of Uto-Aztecan reduplication, while the
long vowel heavy syllable reduplicant for a special class of "marked plurals" was an
O'odham-specific innovation involving the addition of a second mora to a short
vowel plural reduplicant. The development of the O'odham long vowel plural thus
serves as an example of an enhancement of reduplicative form. Section 4 concludes.

2. Reduction of Reduplicative Function

One consideration in the diachronic development of reduplication, which is possibly
concomitant with but also potentially separate from phonological reduction, is the
reduction of reduplicative function(s). Functional reduction involves the loss of
previous distinctions that were once made in the complex reduplicative system of an
earlier incarnation of a given language. To put it succinctly, functional reduction
involves the simplification of reduplicative systems.

One example of such reduction of function can be seen in the free variation that
exists for speakers of Mayo, a Taracahitic language of Southern Uto-Aztecan.
Reduplication is ubiquitous in the Uto-Aztecan (UA) family, and UA languages
typically contain multiple patterns of reduplication for a variety of different functions
(Langacker 1977, Haugen 2005, 2008). Further, there is often a lack of form-function
correspondence in the different patterns of reduplication that a UA language
contains. For example, Yaqui, another Taracahitic language and Mayo’s closest
known linguistic relative, has a variety of reduplicative patterns serving a variety of
functions but with no clear correlation between the form and function of the various
patterns (Harley and Amarillas 2003). Which reduplicant goes with which pattern
seems to be lexically listed for each stem (Haugen 2003). For example, three of the
reduplicants are: a copy of the first syllable of the stem; a CVCV reduplicant (for
some but not all stems of at least three syllables); and a heavy syllable reduplicant
triggering gemination. These reduplicative forms are illustrated in examples (1)-(3),
respectively. Each of these reduplicative patterns can indicate the notions of habitual
or progressive action, or be used in the imperative, as illustrated in examples (a)-(c), respectively, in each of (1)-(3).

(1) 
Yaqui Syllabic Reduplication (Harley and Amarillas 2003)

a. Aapo pahkowa-u hiva kit-kitte
   S/he fiesta-at always RED-knead
   ‘She always makes dough at fiestas’

b. Aapo papa-m bwia-po yeu bwak-bwakta
   S/he potato-PL earth-in out RED-pluck
   ‘S/he is plucking potatoes out of the soil’

c. Kat ae-t chep-chepte
   NEG.IMP it-on RED-step
   ‘Don’t step on it.’

(2) 
Yaqui CVCV Reduplication (Harley and Amarillas 2003)

a. Empo si manteka-ta tata-ta chiha-chihakta
   You very lard-ACC hot-ACC RED-splash
   ‘You (habitually) splash the hot lard too much’

b. Uu uusi muni-m chive-chivehta
   DET child bean-PL RED-spread.out
   ‘The child is spreading out the beans’

c. Kat aa kamu-kamukta
   NEG.IMP 3.SG.ACC RED-hold.in.mouth
   ‘Don’t keep it in your mouth’

(3) 
Yaqui Geminating Heavy Syllable Reduplication (Harley and Amarillas 2003)

a. Si=nee kowi-m hah-haita
   Very—I pig-PL RED-hate
   ‘I get very disgusted with pigs.’

b. Vempo huunum ho-hoteka-su saha-k
   They there RED-sit.PL-COMPL go-PERF
   ‘They started to sit there and finally left’

c. Kat a’avo am kik-kima
   Don’t there them RED-bring.in
   ‘Don’t bring them in (there)’

Similar examples exhibiting a lack of form-function correspondence are discussed in detail for another Southern UA language, Nahuatl (Corachol-Aztecan), by Tuggy (2003). Such examples call into question Bybee et al. (1994)’s strong claim that the synchronic forms of multi-pattern reduplicative systems evince the historical semantic changes that must have occurred in those systems (Haugen 2003, 2005).

Returning to the multi-pattern system of Mayo, the point that I would like to make here is that such semantic distinctions can be lost over time, in a process that I
am calling functional reduction. Mayo has at least two reduplicants, a light syllable and a heavy syllable, but there exists a functional semantic distinction between these reduplicants only for the eldest generations of speakers. There are two word classes in Mayo that are differentiated in how the second syllable is formed, either through copying into the second syllable of the stem (in the “accented” class), as in (4), or through gemination of the onset of the stem into the coda of the reduplicant (in the “accented” class), as in (5). The distinction between these classes is not relevant for our purposes here; for discussion see Hagberg (1993) and Haugen (2004). I will refer to the Mayo light syllable reduplicant as “RED1” and the heavy syllable as “RED2”.

(4) Mayo “unaccented” words (Hagberg 1993)

<table>
<thead>
<tr>
<th>Stem</th>
<th>RED1=σ_u</th>
<th>RED2=σ_uu</th>
<th>Unattested</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. bwa.ná</td>
<td>bwa.bwá.na</td>
<td>bwan.bwá.na</td>
<td>*bwab.bwa.na</td>
<td>‘cry’</td>
</tr>
<tr>
<td>b. bwi.ká</td>
<td>bwi.bwi.ká</td>
<td>bwik.bwi.ká</td>
<td>*bwib.bwi.ká</td>
<td>‘sing’</td>
</tr>
<tr>
<td>c. om.té</td>
<td>o.óm.te</td>
<td>om.óm.te</td>
<td>*o’.óm.te</td>
<td>‘hate’</td>
</tr>
<tr>
<td>d. no.ká</td>
<td>no.nó.ká</td>
<td>nok.nó.ká</td>
<td>*non.nó.ká</td>
<td>‘speak’</td>
</tr>
</tbody>
</table>

(5) Mayo “accented” words (Hagberg 1993)

<table>
<thead>
<tr>
<th>Stem</th>
<th>RED1=σ_u</th>
<th>RED2=σ_uu</th>
<th>Unattested</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. yú.ke</td>
<td>yú.yu.ke</td>
<td>yúy.yu.ke</td>
<td>*yûk.yu.ke</td>
<td>‘rain’</td>
</tr>
<tr>
<td>b. wom.te</td>
<td>wó.wom.te</td>
<td>wóm.wom.te</td>
<td>*wów.wom.te</td>
<td>‘be frightened’</td>
</tr>
<tr>
<td>c. nók.wa</td>
<td>nó.nok.wa</td>
<td>nók.nok.wa</td>
<td>*nón.nok.wa</td>
<td>‘known language’</td>
</tr>
<tr>
<td>d. nó.ka</td>
<td>nó.nó.ká</td>
<td>nón.nó.ká</td>
<td>*nók.nó.ká</td>
<td>‘know a language’</td>
</tr>
</tbody>
</table>

The point here is that the semantic distinction between RED1 and RED2 (in both word classes) has been lost for all but the eldest generation of speakers. Thus, for the younger generations there has been a recent development of free variation between light and heavy syllable reduplication forms.3

3. Enhancement of reduplicative form—Evidence from Uto-Aztecan

While reduction in form seems to be quite common as a general process of diachronic change for reduplicative morphemes cross-linguistically, in this section I will argue that the enhancement of form is also a strategy available for languages to develop new reduplication patterns. I will support this contention based on comparative evidence from the Uto-Aztecan language family that suggests that such a process occurred in Tohono O’odham.

3 Campbell and Muntzell (1989) discuss morphological reduction, category loss, and the leveling of allomorphy in languages undergoing severe shift and/or death, and the semantic reduction seen in Mayo reduplication might well be a consequence of its endangered status (see Moctezuma Zamarron 2001).
3.1. Reduplication in Tohono O'odham

Tohono O'odham, formerly known as "Papago", is a Southern Uto-Aztecan language in the Tepiman sub-group which is indigenous to southern Arizona, USA and northern Sonora, Mexico.

Tohono O'odham has three distinct and productive patterns of reduplication. The first is a light syllable (i.e. CV) prefixal reduplicant used to mark the majority of plural nouns of the language. Some examples are provided in (6):

(6) CV reduplication for Tohono O'odham unmarked plurals (Hill and Zepeda 1998)

<table>
<thead>
<tr>
<th>singular</th>
<th>plural</th>
<th>meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. bitokoi</td>
<td>bi-bitokoi</td>
<td>'Pinacate beetle'</td>
</tr>
<tr>
<td>b. ce-kol</td>
<td>ce-ckol</td>
<td>'squirrel'</td>
</tr>
<tr>
<td>c. cu:wí</td>
<td>cu-cu:wí</td>
<td>'jackrabbit'</td>
</tr>
<tr>
<td>d. gogs</td>
<td>go-gogs</td>
<td>'dog'</td>
</tr>
<tr>
<td>e. mi:stol</td>
<td>mi-mstol</td>
<td>'cat'</td>
</tr>
</tbody>
</table>

I follow Fitzgerald (2000)'s analysis of these data, wherein the reduplicant is viewed as a CV prefix which in some contexts triggers syncope in the base (cf. 6a, b, and e; see Fitzgerald 2000 for a full account of the conditions on the application or non-application of this syncope, which is beyond the scope of our concern here).

A second pattern of reduplication is a long vowel reduplicant that is used as the plural morpheme for a culturally significant class of less than 100 noun roots (Hill and Zepeda 1998). A sample is provided in (7):

(7) Tohono O'odham CVV reduplication for marked plurals (Hill and Zepeda 1998)

<table>
<thead>
<tr>
<th>singular</th>
<th>plural</th>
<th>meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ban</td>
<td>ba:-ban</td>
<td>'coyote'</td>
</tr>
<tr>
<td>b. mad</td>
<td>ma:-mad</td>
<td>'woman's child, younger sister's child'</td>
</tr>
<tr>
<td>c. ñem</td>
<td>ñe:-ñem</td>
<td>'liver'</td>
</tr>
<tr>
<td>d. bahi</td>
<td>ba:-bhai</td>
<td>'tail'</td>
</tr>
<tr>
<td>e. şon</td>
<td>şo:-şon</td>
<td>'trunk of a plant'</td>
</tr>
</tbody>
</table>

Henceforth we will refer to this class as "the long-vowel plurals". It is important to note that the appearance of the long vowel is not predictable based on any intrinsic properties of the nominal stems which take this plural. Rather, the members of this nominal class are semantically-motivated; see Hill and Zepeda (1998) for discussion.

---

4 This alternation between long and short vowel may not be limited to O'odham within Tepiman. Kager (1997) discusses the selection of long vowel reduplicants in Southeastern Tepehuan as being neither phonologically nor morphologically conditioned. Although Kager does not propose a semantically-based account of reduplicative allomorphy in Southeastern Tepehuan, its close familial relationship to O'odham makes the retention of a shared nominal classification system not unlikely. However, we will leave this important point for future research.
of the abstract conceptual schema and various metaphorical extensions that form the basis for this nominal classification. If this class was not marked with the long vowel reduplicant it would be indistinguishable from (i.e. homophonous with) the unmarked class in the plural. We return to this point below.

The third pattern of reduplication in O’odham is a heavy syllable that triggers gemination of the onset consonant of the base into the coda of the reduplicant. This pattern of reduplication is used either for marking distributivity of nouns and some verbs (8), or repetition in some verbs (9):

(8) Tohono O’odham heavy syllable distributives (Fitzgerald 2003)

<table>
<thead>
<tr>
<th>Singular</th>
<th>Plural</th>
<th>Distributive</th>
<th>Meanings</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. nowiu</td>
<td>no-nowiu</td>
<td>non-nowiu</td>
<td>‘ox’</td>
</tr>
<tr>
<td>b. nahagio</td>
<td>na-nhagio</td>
<td>nan-nhagio</td>
<td>‘earring’</td>
</tr>
<tr>
<td>c. hódai</td>
<td>hó-hodài</td>
<td>hoh-hodai</td>
<td>‘rock, stone’</td>
</tr>
<tr>
<td>d. ?a:g</td>
<td>?a-?ag</td>
<td>?a?-?ag</td>
<td>a pair of animal horns’</td>
</tr>
</tbody>
</table>

(9) O’odham heavy syllable répétitives (Fitzgerald 2003)

<table>
<thead>
<tr>
<th>Unitative</th>
<th>Repetitive</th>
<th>Meanings</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. giw</td>
<td>gig-giw</td>
<td>‘hit something’</td>
</tr>
<tr>
<td>b. himh</td>
<td>hih-him</td>
<td>‘laugh’</td>
</tr>
<tr>
<td>c. huhulgat</td>
<td>huh-hulgat</td>
<td>‘menstruate’</td>
</tr>
<tr>
<td>d. kow</td>
<td>kok-kow</td>
<td>‘dig something out of the ground’</td>
</tr>
<tr>
<td>e. jila</td>
<td>jiin-na</td>
<td>‘look in a certain direction’</td>
</tr>
</tbody>
</table>

The existence of two realizations of heavy syllable reduplication (i.e. as a heavy syllable either inducing gemination or appearing with a long vowel) raises an interesting issue for the Correspondence Theory of reduplication (McCarthy and Prince 1995), which posits that the emergence of a reduplicant should be based on a single ranking of markedness and faithfulness constraints. For example, in the early version of the theory both of the heavy syllable reduplicants of Tohono O’odham would have been specified with a template constraint like that in (10c), which would yield a heavy syllable reduplicant contrasting with a light syllable reduplicant, itself created with a different template (10b). Both (10b) and (10c) would be language-specific template constraints. (10a) gives the generalized definition of template as proposed by McCarthy and Prince (1993): 5

(10) Template Constraints

a. $M_{\text{cat}} = P_{\text{cat}}$ - A morphological category is realized as a prosodic category.

b. $\text{RED1} = \sigma_1$ - RED1 is a light syllable (i.e. composed of one mora).

c. $\text{RED2} = \sigma_2$ - RED2 is a heavy syllable (i.e. composed of two moras).

5 Although subsequent approaches to Correspondence Theory have abandoned templates in favor of other size restrictions on reduplicative forms (e.g. alignment to one edge of a prosodic category like “heavy syllable”, as in Hendricks 1999), the general point about the categorical ranking of the markedness constraints would still hold for these alternative approaches.
With all else being equal, two markedness constraints would have to be crucially ranked in order to derive a heavy syllable reduplicant with the second mora being realized as either a long vowel or as a geminated consonant. For this purpose we can propose such constraints as *LONG$_V$ and *LONG$_C$, defined in (11) and (12), respectively:

(11)  *LONG$_V$ - Long vowels are prohibited.

(12)  *LONG$_C$ - Long consonants are prohibited.

A ranking of *LONG$_V$ >> *LONG$_C$ would yield a reduplicant with gemination (as in 13), whereas the opposite ranking would yield a reduplicant with a long vowel (as in 14):

(13)  *LONG$_V$ >> *LONG$_C$

/
ban/ + RED2  RED2 = $\sigma_{sy}$  *LONG$_V$  *LONG$_C$

<table>
<thead>
<tr>
<th></th>
<th>a.</th>
<th>b.</th>
<th>c.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ba-ban</td>
<td>baa-ban</td>
<td>bab-ban</td>
</tr>
<tr>
<td></td>
<td>!</td>
<td>!</td>
<td></td>
</tr>
</tbody>
</table>

(14)  *LONG$_C$ >> *LONG$_V$

/
ban/ + RED2  RED2 = $\sigma_{sy}$  *LONG$_C$  *LONG$_V$

<table>
<thead>
<tr>
<th></th>
<th>a.</th>
<th>b.</th>
<th>c.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ba-ban</td>
<td>baa-ban</td>
<td>bab-ban</td>
</tr>
<tr>
<td></td>
<td>!</td>
<td></td>
<td>!</td>
</tr>
</tbody>
</table>

Since both the (b) and (c) candidates can surface in O’odham (depending on the meaning of the RED morpheme), it is not clear how we should rank the markedness constraints *LONG$_C$ and *LONG$_V$. If all else is equal and if one of these constraints is categorically ranked higher than the other, then a heavy syllable template should yield either long vowel reduplicants or geminating reduplicants.

This synchronic puzzle only arises given one particular synchronic theory of phonology and morphology interaction, but since my focus here is on the diachronic development of these patterns I will raise but not answer the following synchronic question: how do we account for a long vowel heavy syllable reduplicant in a grammar that prefers to create reduplicants by means of gemination? (Because the geminating heavy syllable appears for multiple reduplicative functions, i.e. distributive and repetitive, and since the long vowel reduplicant is limited to a relatively small sub-group of culturally-important nouns in this language, I take it to be the case that gemination is the default, i.e. productive, way to create a heavy syllable reduplicant in Tohono O’odham). See Haugen (2008) for a proposed solution to this synchronic problem within Correspondence Theory.

Our focus in the remainder of this section will be on the diachronic explanation for the rise of the long vowel plural reduplicant in Tohono O’odham. Given the comparative Uto-Aztecan evidence to be presented below I will argue that this long vowel plural was an innovation. In brief, my claim is that Tohono O’odham is the daughter of a proto-language that had a heavy syllable distributive reduplicant (as I will argue in section 3.2), as well as at least two classes of plural nouns (§3.3). A
reconstruction of the stages of the development of the long vowel plural will be detailed in section 3.4.

3.2. A Heavy Syllable Reduplicant as 'Distributive' in Proto-Uto-Aztecan

I argue that a heavy syllable reduplicant can be reconstructed as a Proto-Uto-Aztecan morpheme marking 'distributive'. Heath (1977, 1978) has previously argued for two patterns of reduplication for 'distributive' in earlier stages of Uto-Aztecan. These were a *CV- prefixal reduplicant concurrent with a thematic suffixal morpheme *-tai- for the Northern Uto-Aztecan (NUA) languages, and a *CV?- prefixal reduplicant (without a thematic suffix) for the Southern Uto-Aztecan (SUA) languages.

The evidence for a difference between the prefixal distributive reduplicants in the two major sub-groups is not compelling; in fact, to the contrary, the evidence suggests that they were the same, either a geminating heavy syllable or a CVL-heavy syllable, where “L” is a laryngeal consonant ([ʔ] or [h]). (I will not be addressing the presence or absence of a co-occurring thematic suffix here).

The division of Uto-Aztecan into the major sub-groups of "Northern" and "Southern" is traditional, but there is dispute as to whether these indicate descent from a particular proto-language after the split of PUA or whether they should be regarded as simple areal designations with no cladistic implications; see the Appendix for a UA family tree with points of contention highlighted. Some scholars hold that NUA was derived from a "Proto-Northern-Uto-Aztecan" (PNUA) (e.g. Manaster Ramer 1992, based on a shared phonological sound change of intervocalic *-c- > *-y-), while others maintain that the SUA languages were derived from a "Proto-Southern-Uto-Aztecan" (e.g. Miller 1984, based on shared cognate density). Both camps deny the likelihood of both PNUA and PSUA as having been post-PUA daughter languages. In any event, heavy syllable reduplicants containing both laryngeals and geminates exist in both of these sub-groupings, and Occam’s razor suggests reconstructing such reduplicants for PUA, a position which would render one’s opinion on later UA sub-grouping irrelevant with respect to this issue.

Haugen (2005, 2008) reconstructs the PUA 'distributive' as a heavy syllable reduplicant, where the second mora was underdetermined by the grammar. It could have been realized as gemination, a laryngeal (via consonant epenthesis), or a long vowel, depending on the relative rankings of markedness constraints in a given language. Each of these possibilities is attested in extant UA languages. One

---

6 The superscript /'/ here indicates that this suffix would have triggered ablaut in the stem to which it attached.

7 Two anonymous reviewers object to the employment of Optimality Theoretic notions here and in the following discussion. We can remain agnostic on important questions in modern phonological theory while still maintaining that OT machinery proves to be exceedingly useful for our purposes below. The central concern of OT, i.e. the delineation of a universal set of markedness constraints which must be ranked relative to one another in the grammar of any
advantage of this approach is that it allows us to reconstruct cognate reduplicant morphemes even without identity of surface realization. For example, the 'distributive' is realized as a geminating heavy syllable in Tohono O'odham, as in (8) above, but it can also be realized as a CVL- reduplicant in Guarijio numerals (15) and Nahuatl verbs (16) and numerals (17):

(15) Guarijio heavy syllable distributive reduplication (Miller 1996)
   a. naó ‘cuatro/four’ naʔ-náó ‘de cuatro en cuatro’/‘four by four’
   b. mariki ‘cinco/five’ maʔ-márigi ‘de cinco en cinco’/‘five by five’

(16) Nahuatl reduplication for distributive action (Canger 1981: 37)
   a. ni-k-te:-teki in tlaskalli b. ni-k-teʔ-teki in tlaskalli
      I-it-RED-cut the bread I-it-RED-cut the bread
      ‘I cut/slice the bread’ ‘I cut the bread to pieces’

(17) Xalatzala (Central Guerrero) Nahuatl distributive reduplication (Canger 1981)*
   a. oh-ome ‘two at a time’
   b. yeh-yeyi ‘three at a time’
   c. nah-na:wi ‘four at a time’
   d. ma:-ma:kʷili ‘five at a time’

Although the CV?- pattern is only used for distributivity on numerals in Guarijio, the CVL- reduplicant is completely productive in Nahuatl, and Canger (1981) reconstructs a *CV?-. pattern for 'distributive' in Proto-Aztec.

Recall that Heath reconstructs *CV?-. pattern for PSUA. This reconstruction does not directly account for the geminating heavy syllable reduplicant form found in O'odham distributives. However, both patterns can be accounted for if we reconstruct a heavy syllable reduplicant for the proto-language, and then allow the language-specific ranking of markedness constraints to determine the output realization of the second mora. For example, let us suppose that this heavy syllable distributive reduplicant was realized via gemination in PUA. This would imply that the constraint given language, was originally designed to capture typological generalizations in synchronic grammar(s). This concern can be straightforwardly applied to questions of diachrony in closely related languages through re-ranking certain of the universal constraints. Outstanding problematic issues that have been raised for OT, such as what the nature of the universal set of constraints might be or whether a single input-output mapping is sufficient or whether several mappings occur in different levels (as in, for example, Kiparsky 2000’s Stratal Optimality Theory) are completely orthogonal to our discussion here. The basic approach of employing relative rankings of universal constraints is maintained in any version of OT, and it has even been adopted in other modern frameworks (e.g. Morphological Doubling Theory as proposed by Inkelas and Zoll 2005). For a more extended defense of this approach to comparative Uto-Aztecan morphophonology see Haugen (2008).

* In some varieties of Nahuatl, the laryngeal consonant is realized as /h/. I assume that the laryngeal is epenthetic in the environment of heavy syllable reduplication for both dialect types (i.e. whether the laryngeal is glottal stop or /h/).
*LONGC would have been ranked lower than *LONGv or DEPIO, the constraint prohibiting epenthesis of a laryngeal into the coda position. The effect of this constraint ranking is shown in (18), with the hypothetical stem noka:

\[
\text{REF2} \quad \sigma_{IJ} = \begin{array}{c|c|c|c|c|}
\hline
/noka/ + \text{REF2} & \text{REF2} = \sigma_{IJ} & *\text{LONGv} & *\text{DEPIO} & *\text{LONGC} \\
\hline
a. no-noka & * & * & * & * \\
b. noo-noka & * & * & * & * \\
c. non-noka & * & * & * & * \\
d. no?-noka & * & * & * & * \\
\hline
\end{array}
\]

(18a) does not create a heavy syllable and is ruled out by the REF2 template constraint. The relative ranking of *LONGv and DEPIO is not determined, but candidates with a violation of either, (18b) and (18d), are less optimal than the winning candidate, (18c), which violates neither. Thus, the crucial ranking here is that of *LONGC, which must be lowly ranked so that a single violation of it is acceptable in the optimal surface candidate. (Underlying long vowels, of course, would be allowed to surface in stems given a sufficiently high ranking for MAXIO).

If correct, this hypothetical constraint ranking would have been retained from PUA by the UA languages that continue to use gemination as the default for creating heavy syllable reduplicants. This is in fact the most common way for Uto-Aztecan languages to create heavy syllable reduplicants; such extant languages with this pattern include Tohono O'odham, Yaqui, and Mayo (in the accented class of words). Now let us suppose that some languages, such as Guarijio, have promoted the constraint *LONGC higher in the overall constraint hierarchy, thus creating a dispreference for geminates. If *LONGC were to be re-ranked higher than DEPIO, then the insertion of a laryngeal would become the preferred manner of creating the heavy syllable reduplicant. This new state of affairs is illustrated in (19):

\[
\text{REF2} \quad \sigma_{IJ} = \begin{array}{c|c|c|c|}
\hline
/noka/ + \text{REF2} & \text{REF2} = \sigma_{IJ} & *\text{LONGv} & *\text{LONGC} \\
\hline
a. no-noka & * & * & * \\
b. noo-noka & * & * & * \\
c. non-noka & * & * & * \\
d. no?-noka & * & * & * \\
\hline
\end{array}
\]

The heavy syllable template constraint and *LONGv are still highly ranked, so candidates that violate those (19a and b) are eliminated. Now that *LONGC has been promoted, the violating candidate (19c) is now sub-optimal, and it is eliminated. Candidate d now only violates the lowest-ranked constraint and thus becomes the optimal candidate.

The advantage of this approach over other conceivable alternatives is that the ranking of constraints is not unique to the process of reduplication, but apply to the phonology of the language as a whole. We find positive evidence in favor of a constraint ranking like (19) holding in Guarijio by noting that there is a general ban
on geminates across the Guarijio language (Miller 1996). This can be clearly
demonstrated by comparing Guarijio roots with those in Mayo, where the latter have
a geminate consonant but the former has a laryngeal:

(20) Guarijio Roots with a laryngeal in correspondence to Mayo geminates

<table>
<thead>
<tr>
<th>Guarijio</th>
<th>Mayo</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ehté</td>
<td>étte</td>
<td>'louse'</td>
</tr>
<tr>
<td>b. ehka-ní</td>
<td>hékka</td>
<td>'shade'</td>
</tr>
<tr>
<td>c. kahtí</td>
<td>káttek</td>
<td>'to be seated'</td>
</tr>
<tr>
<td>d. mahta</td>
<td>máatta</td>
<td>'metate'</td>
</tr>
<tr>
<td>e. tehté</td>
<td>tétte</td>
<td>'stone'</td>
</tr>
</tbody>
</table>

The discrepancy between the geminates in Mayo and the [h] in Guarijio can be
straightforwardly explained by stating that debuccalization occurred in Guarijio. The
development of a general ban on geminates in Guarijio explains why the distributive
reduplication in (15) above (and the plural nouns in Guarijio that will be discussed
below) also involves the pattern of CV?-, rather than the normally expected pattern
of Uto-Aztecan heavy syllable reduplication triggering gemination.

Under the hypothesis that a heavy syllable distributive is reconstructable for
PUA, it is clear that this reduplicative morpheme must have developed in different
ways in different UA languages. In some languages, e.g. Tohono O'odham, its form
and function have remained largely intact. In other languages, it retains the form but
has taken on a new function (or functions), as in Yaqui and Mayo (1-5 above). In yet
other UA languages, this reduplicant has maintained its form but taken on a new
function and lost its productivity, as in some non-productive forms for plurals and
duals in Tümpisa Shoshone (NUA, Central Numic) (21):

(21) Tümpisa Shoshone (Central Numic) number marking (Dayley 1989)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Pa.pi</td>
<td>'older brother'</td>
<td>pa.pi.am.mü 'older bro. (pl.)'</td>
</tr>
<tr>
<td>b. pat.si</td>
<td>'older sister'</td>
<td>pap.pat.si.am.mü 'older sister (pl)'</td>
</tr>
<tr>
<td>c. pe.tü</td>
<td>'daughter'</td>
<td>pep.pe.tüm.mü 'daughter (pl)'</td>
</tr>
<tr>
<td>d. tangum.mü</td>
<td>'man'</td>
<td>tat.tangung.ku 'man (dl)'</td>
</tr>
<tr>
<td>e. tokkwapü</td>
<td>'aunt'</td>
<td>tot.tok.kwapüammü 'aunt (pl.)'</td>
</tr>
<tr>
<td>f. tua</td>
<td>'son'</td>
<td>tut.tu.am.mu 'son (pl.)'</td>
</tr>
</tbody>
</table>

In the case of Guarijio and Nahuatl, the form has changed (with debuccalization of
the geminate leading to a laryngeal consonant), but the function remains the same
(15-17). Another language which seems to have a CVL distributive is Comanche
(NUA, Numic), although it does not seem to be productive in this language. CV
reduplication is used for plural in Comanche, although sometimes this reduplicant
also has a laryngeal ([h]), and in some cases this seems to yield a distributive
meaning (22d):

(22) Comanche number marking (Charney 1993)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a. nam.mi</td>
<td>'sister'</td>
<td>na.na.na.mi 'sisters'</td>
</tr>
<tr>
<td>b. pi.a</td>
<td>'big'</td>
<td>pi.pi.a [pi.via] 'big group'</td>
</tr>
</tbody>
</table>
In sum, the claim here is that Tohono O'odham retains an old UA pattern for creating the distributive morpheme: i.e. the use of a heavy syllable reduplicant triggering gemination. We now turn to a consideration of the system for marking plurals in PUA.

### 3.3. Two Classes of Plurals in PUA: Marked and Unmarked

Special morphological marking for plurals in two different noun classes can be reconstructed for PUA, although the daughter languages differ with respect to how this marking occurs, and in which nouns fall into the relevant classes (Hill and Hill 2000). Hill and Hill refer to these classes as “marked” and “unmarked”, respectively, and the difference typically involves some kind of distinction along the lines of animate/inanimate, human/non-human, etc.

According to Hill and Hill, UA languages typically use suffixation for the unmarked class, and some process of reduplication for the marked class. For example, unmarked nouns in Cupeno (NUA, Takic) take the suffix -m, as with kaxa-l ‘quail’ > kaxa-l-im ‘quails’ (J. Hill 2005). Human nouns and certain nominals that have been given importance through semantic extension, on the other hand, typically show reduplication in addition to this suffixation. For example, puu-l ‘doctor’ > pu- vu-l-im ‘doctors’. A few Cupeno nouns are attested with reduplication only. Reduplication co-occurring with suffixation is also displayed in the human nouns in the examples from Tümpisa Shoshone and Comanche in (21) and (22) above.

Similarly, in Guarijio most nouns are not marked at all for the plural (Miller 1996). However, human nouns are reduplicated, and some of these are reduplicated with a glottal stop, and some are optionally reduplicated with a glottal stop. Some examples are given in (23):

### (23) Guarijio (Taracahitic) “marked heavy syllable” = “marked plural”

<table>
<thead>
<tr>
<th>Stem</th>
<th>RED w/ out</th>
<th>RED w/</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ku.ci.ta</td>
<td>—</td>
<td>ku?.ku.ci</td>
<td>‘son, daughter’</td>
</tr>
<tr>
<td>b. ma.la.la</td>
<td>ma.ma.la.la</td>
<td>ma?.ma.la.la</td>
<td>‘daughter’</td>
</tr>
<tr>
<td>c. se.pu.ri</td>
<td>se.se.pu.ri</td>
<td>se?.se.pu.ri</td>
<td>‘uncle, aunt’</td>
</tr>
<tr>
<td>d. pa.mi.la</td>
<td>—</td>
<td>pa?.pa.mi.la</td>
<td>‘boss, governor’</td>
</tr>
<tr>
<td>e. no.la</td>
<td>no.no.la</td>
<td>no?.no.la</td>
<td>‘son’</td>
</tr>
</tbody>
</table>

As mentioned above, Tohono O’odham indicates a contrast between marked and unmarked plurals with a vowel-length distinction in the prefixal reduplicant: unmarked plurals have a short vowel reduplicant and marked plurals have a long vowel reduplicant.
Following the discussion of Hill and Hill (2000) as well as previous work in PUA morphology like Langacker (1977), we can reconstruct at least one plural morpheme along the lines of "-m for unmarked plurals in PUA." In the next section I will propose a series of stages which led to the marked plurals having long vowel reduplicants in Tohono O'odham. Crucially, this reconstruction will involve the enhancement of the short vowel reduplicant with an extra mora; this is akin to the process of mora augmentation discussed by Davis (2001).

3.4. An Historical Account of the Long Vowel Plurals in Tohono O'odham

The preceding sections have supplied the backdrop that we need in order to return to our central issue of enhancement of form as a mechanism for the development of new reduplicative morphemes. I suggest that the long vowel reduplicant for the Tohono O'odham marked plurals was developed by the enhancement of a phonological contrast in a particular reduplicative morpheme, a light syllable plural reduplicant, motivated by the maintenance of the inherited semantic contrast of "marked" vs. "unmarked" plural. This development plausibly occurred in the following stages.

First, Early Tohono O'odham would have inherited a marked/unmarked plural distinction from PUA. By hypothesis, this could have been much like the extant system of Cupeno and other NUA languages, where unmarked plurals were created with a suffix and the marked plurals were created with reduplication in addition to that suffix. If PUA also had a heavy syllable reduplicant for distributive (as was argued to be the case in section 3.2 above), this plural reduplicant was probably a light syllable, as is currently typical across the family in those languages that have plural reduplication. This first stage was presumably paralleled by Guarijio, another SUA language:

\[
\begin{array}{c|c|c|c}
\text{Stage 1: Inheritance of marked/unmarked distinction from early Uto-Aztecan} & \text{Singular} & \text{Plural} & \text{Marked Plural} \\
\hline
\text{Early O'odham:} & \text{noun.stem} & \text{noun.stem-SUFFIX} & \text{RED-noun.stem-SUFFIX} \\
\text{Early Guarijio:} & \text{noun.stem} & \text{noun.stem-SUFFIX} & \text{RED-noun.stem-SUFFIX} \\
\end{array}
\]

Over time there was a tendency toward the loss of the plural suffix in some SUA languages, including O'odham and Guarijio. In this second stage, Middle Guarijio would have lost its plural suffix in both the unmarked and marked plurals. This led to the lack of contrast between unmarked plural and singular nouns that is retained in the modern language:

\[\text{K. Hill (2001) reconstructs two plural suffixes for PUA: } **-mi \text{ and } **-ti. A detailed reconstruction and historical account of the plural and its development in the various subgroups of UA has not been worked out, and goes beyond the scope of the present paper.}\]
Stage 2: Tendency toward loss of plural suffix in some Southern Uto-Aztecan languages

| Stage 2: Tendency toward loss of plural suffix in some Southern Uto-Aztecan languages |
|---------------------------------|-----------------|-----------------|
| Early Middle                   | Singular        | Plural          | Marked Plural   |
| O'odham:                       | noun.stem       | noun.stem-SUFFIX | RED-noun.stem-SUFFIX |
| Middle                         | noun.stem       | noun.stem-SUFFIX | RED-noun.stem-SUFFIX |
| Guarijio:                      |                 |                 |                 |

For O'odham, the tendency toward the loss of the plural suffix might have led all the way to a loss of the singular/plural distinction, as it did in Guarijio, or, as was the case for Yaqui, the loss of the marked/unmarked distinction itself. However, it seems that speakers of O'odham made sure to maintain both the singular/plural distinction, on the one hand, and the marked/unmarked plural noun distinction on the other. This might have led to a conflict like that depicted below as Stage 3:

Stage 3: Conflict between reduplicated prefixes for unmarked/marked plurals

| Stage 3: Conflict between reduplicated prefixes for unmarked/marked plurals |
|---------------------------------|-----------------|-----------------|
| Later Middle                   | Singular        | Plural          | Marked Plural   |
| O'odham:                       | noun.stem       | RED-noun.stem-SUFFIX | RED-noun.stem-SUFFIX |

At this hypothetical Stage 3 there would have been a tension between either using the inherited suffix or reduplication to indicate the unmarked plural. Phonological reduction is a general feature of the Tepiman languages, as seen in such borrowed stems as Tohono O'odham čı̱kan ( < Nahuatl tekipanoa 'work'), and in their containing the productive morphological process of truncation (for the perfective).

However, just reduplicating a CV, as normal for marked plurals, would have lost the marked/unmarked distinction. Therefore, I claim, speakers employed the strategy of mora augmentation (Davis 2001), a kind of morpho-phonological enhancement, to create a new pattern of reduplication. The usual pattern of mora augmentation in this language would have led to gemination, but reduplication with gemination would have conflicted with the distributive and repetitive reduplicative morphemes that were inherited from the proto-language (cf. §3.2). So, contrast was maintained for the marked plurals via the alternative strategy of mora augmentation through vowel-lengthening, as in (27):

Stage 4: development of prefixal reduplicant to maintain contrast between SG/PL in O'odham, enhancement of animate prefix with an extra mora

| Stage 4: development of prefixal reduplicant to maintain contrast between SG/PL in O'odham, enhancement of animate prefix with an extra mora |
|---------------------------------|-----------------|-----------------|
| Late O'odham:                  | Singular        | Plural          | Marked Plural   |
|                                | noun.stem       | RED-noun.stem   | REED-noun.stem  |

The shift to use of CV reduplication in the plural would have led to paradigmatic pressure for the marked plural to be enhanced somehow in order to maintain the semantic distinction of marked/unmarked.

---

10 Yaqui kept the -m plural suffix and lost nominal reduplication and the marked class.
Under the scenario sketched here, the process of vowel-lengthening would have served as such a phonological enhancement strategy to perform two morphological functions. First, it maintained an inherited contrast between marked and unmarked plurals, but second, it also maintained the contrast between the marked plurals and the distributives. If the surface contrast had not been enhanced (i.e. changed in some way, e.g. by lengthening the vowel through mora augmentation), these nouns would sound just like any other plural noun and they would thus have lost their “marked” status. If the otherwise default mechanism of gemination was used, though, the semantic notion of plurality would have been overridden by the already expected notion of ‘distributivity’.

This historical development suggests that enhancement of form is a viable mechanism for creating a new reduplicative morpheme. It also suggests that homophony avoidance was a motivating factor in the phonological enhancement of the long vowel reduplicant in Tohono O’odham.\footnote{The historical development of long vowel reduplicants is proposed here for Tohono O’odham (and possibly some other Tepiman languages, cf. footnote 4) only. There are some other Uto-Aztecan languages that also contain long vowel reduplicants (e.g. Hopi and Nahuatl), but, under the assumption that long vowel reduplicants have not been retained from PUA, the historical development of those particular morphemes will have to be left to future research. What is important here is that the comparative evidence suggests that the long vowel reduplicants for marked plurals in O’odham (and perhaps Tepiman more generally) were a unique innovation, and it is that innovation that the account sketched above has been proposed to explain.}

4. Conclusion

In conclusion, this paper has proposed two different possible diachronic paths of development for reduplicative morphemes that are distinguishable from the more common path countenanced in the grammaticalization approach (i.e. reduction of form with or without semantic bleaching).

The first, functional reduction, involves the simplification of reduplicative systems through the loss of semantic contrasts concurrent with the maintenance of differences in reduplicative forms, thus resulting in free variation where different forms can serve the same function.

The second involves the complexification of reduplicative systems, where new reduplicants are created through enhancement of form in order to mark a new semantic contrast. Urbanczyk (2005) attempts to extend recent approaches to phonological “contrast enhancement” to the domain of morphology by discussing instances where contrasts are enhanced in the morphological process of reduplication. This paper thus supports Urbanczyk’s claims by arguing that enhancement of form is a possible path in the development of new reduplicative morphemes. Central to the enhancement strategy reconstructed above was homophony avoidance (i.e. the maintenance of paradigmatic contrasts).
5. References


6. Appendix: The Uto-Aztecan languages & major sub-groups

<table>
<thead>
<tr>
<th>Northern Uto-Aztecan (?)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Numic</strong></td>
</tr>
<tr>
<td>a. Western Numic: Mono, Northern Paiute</td>
</tr>
<tr>
<td>b. Central Numic: Panamint, Shoshone, Comanche</td>
</tr>
<tr>
<td>c. Southern Numic: Kawaiisu, Ute (Chemehuevi, Southern Paiute, Ute)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tübatulabal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tübatulabal</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Takic</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Serrano-Gabrielino</td>
</tr>
<tr>
<td>(1) Serranoan: Serrano, *Kitanemuk</td>
</tr>
<tr>
<td>(2) *Gabrielino (Gabrielino, Fernandino)</td>
</tr>
<tr>
<td>b. Cupan</td>
</tr>
<tr>
<td>(1) Cupeño, Cahuilla</td>
</tr>
<tr>
<td>(2) Luiseño</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hopi</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hopi</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Southern Uto-Aztecan (?)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sonoran (?)</strong></td>
</tr>
<tr>
<td>a. Tepiman: Upper Piman (Tohono O'odham, Akimel O'odham, *Nevome), Pima Bajo, Northern Tepehuan, Southern Tepehuan</td>
</tr>
<tr>
<td>b. Taracahitan</td>
</tr>
<tr>
<td>(1) Tarahumaran: Rarámuri (Tarahumara), Guarijío</td>
</tr>
<tr>
<td>(2) Opatan: *Opata, *Eudeve</td>
</tr>
<tr>
<td>(3) Cahitan: Yaqui, Mayo</td>
</tr>
<tr>
<td>c. Tubar</td>
</tr>
<tr>
<td>Tubar</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Corachol-Aztecan</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Corachol: Cora, Huichol</td>
</tr>
<tr>
<td>b. Aztecan</td>
</tr>
<tr>
<td>(1) *Pochutec</td>
</tr>
<tr>
<td>(2) General Aztec: Pipil, Aztec (many varieties)</td>
</tr>
</tbody>
</table>

* = an extinct language

(adapted from Miller 1984)