Bole suffix doubling as morphotactic extension

Kevin M. Ryan
Harvard University
kevinryan@fas.harvard.edu

Abstract

Bole (West Chadic) has innovated a pattern of semantically vacuous multiple exponence whereby certain verbal suffixes (including all subject agreement and the ventive) are realized on both sides of certain other suffixes (including all object agreement, the additive, and totality), as in [ɗɔ̀pp-**án**-tá-**ŋ**-gó] (follow-PlS-FemSgO-PlS-Perf) 'they followed her' and [ɗɔ̀pp-**át**-tá-**k**-kó] (follow-23FsgS-FemSgO-23FsgS-Perf) 'she followed her.' The pattern is pervasive and regular, but absent from closely related languages. I propose a synchronic and diachronic analysis in terms of adjacency bigram morphotactics: Analogical pressures (which Ryan 2010 terms 'morphotactic extension') favor the realization of the suffix in multiple positions of the suffix string. Languages can resolve this tension between positions either by doubling (as in Bole) or by permitting free variation (as in Tagalog). This analysis of Bole doubling is argued to be superior to one in terms of affix-stem alignment or weakness-driven augmentation, two proposals for similar patterns in other languages, based in part on the fact that Bole doubling is conditioned by both the morpheme preceding and following the second copy.

1 Introduction

Suffix doubling in Bole is a type of multiple exponence, in the sense that it involves multiple realizations of a single morphological feature within a single prosodic/morphological word (Caballero and Harris 2012, Caballero and Inkelas 2013, Harris 2017). More specifically, it is a case of semantically vacuous affix repetition, whereby the same morpheme appears multiple times within the word without independent justification from syntax, semantics, or phonology. This type of repetition is to be distinguished from cases of multiple affixation in which each instance contributes independently to the meaning. In 1, for instance, the Tagalog causative prefix *pa* can be repeated indefinitely, but each repetition introduces another causative shell (Maclachlan 1989). Thus, there is no multiple exponence.

(1)
a. **pa**-kuló? 'to boil something'
b. **pa-pa**-kuló? 'to make someone boil something'
c. **pa-pa-pa**-kuló? 'to cause someone to make someone boil something'

Furthermore, multiple exponence is usually taken to exclude compound-like structures with multiple inflection that can be analyzed as agreement. For example, Vedic Sanskrit has a type of compound called a "double dual" in which each member is inflected for the dual, even though each member is semantically singular (Ryan 2006, Kiparsky 2010). In other words, the inflections reflect the total number of the compound, not the number of each member. *Mātárāpitárā* in 2 is an example.

(2) maːtár **-aː** =pitár **-aː**
mother -Dual =father -Dual
'mother and father' (both semantically singular)

A case of non-vacuous repetition from Bole is given in 3 (Gimba 2000). The same pronominal suffix *mu* 'we' is repeated in two places in the clitic group, but this is not a case of multiple exponence, as the

In the spring of 2010, as a graduate student at UCLA, I had the pleasure of working with Russ on his Yobe Languages Research Project, an NSF grant (BCS-0553222, P.I. Russell G. Schuh). On April 22nd, 2010, we co-delivered the annual joint UCLA-USC phonology seminar, entitled "Suffix doubling and suffix deletion in Bole" (Ryan and Schuh 2010). This paper is based on my half of that talk, which draws on empirical data from Russ's half, Gimba (2000), and Gimba and Schuh (2014).

two *mu*s arguably have different uses: The first indicates the subject, while the second modifies 'body,' serving as an "intransitive copy pronoun" (see §3). (Incidentally, the PL-1PL sequence in 3 *is* a case of multiple exponence, but not one involving affix repetition of the type that is the focus of this paper.)[1]

(3) 'yòr -á꞉ **-mú** =jì꞉ **-mú**
stand -PL -1PL =body -1PL
'(let's) stand!'

Finally, because the Bole verbs analyzed below involve vacuous repetition of the same morpheme, they are unlike cases of multiple exponence involving redundant specification of a feature using distinct morphemes. Examples of this type from Choguita Rarámuri are provided in 4 (Caballero 2008).

(4) a. pá **-s** **-ki** -ma
throw -APPL -APPL -FUT.SG
'he/she/it will throw for someone'
b. sú **-n** **-ki** -ma
sew -APPL -APPL -FUT.SG
'he/she/it will sew for someone'

Three bona fide cases of affix repetition qua multiple exponence (all three treated by Caballero and Inkelas 2013) are now reviewed before turning to the Bole data. First, in Jita, the causative suffix *y* is realized two (or more) times in certain verbs, such as 5 (Downing 2005). Note that 5 is not a double causative semantically.

(5) oku= gus **-y** -á꞉n **-y** -a
INF= buy -CAUS -REC -CAUS -FV
'to sell to each other'

Second, consider the Chichewa verb in 6 (Hyman 2003). The reciprocal is doubled, but, once again, the verb is not semantically a double reciprocal. The rationale for doubling is "morphocentric," as Hyman (2003) puts it (see also Ryan 2010).

(6) a- ku- máŋ -íts **-an** -ir **-a꞉n** -a
3PL- PROG- buy -CAUS -REC -APPL -REC -FV
'to make each other tie for'

Third, Breton plural diminutives require marking the plural both before and after the diminutive, as in 7 (Stump 1991).

(7) bag **-où** -ig **-où**
boat -PL -DIM -PL
'(little) boats'

In agreement with Caballero and Inkelas (2013) and others, such cases cannot be analyzed as reduplication or assimilation. For one thing, vacuous repetitions of the same affix almost never surface adjacent to each other cross-linguistically, in contradistinction to reduplication, where adjacency is the norm (though not without exception). This lack of adjacency also makes it difficult if not impossible for a hypothetical reduplication analysis to define the part of the base that gets copied in phonological terms, since the duplicated material is not aligned with one of the edges of the base, and may have varying prosodic profiles depending on its immediate context. In Bole, for instance, the reduplicant, if taken to be the second copy, would have to reach into its base to copy the penultimate as opposed to adjacent VC. Moreover, semantically vacuous affix repetition is unlike reduplication in that RED usually contributes independent semantic content. It is true that some cases of reduplication do not involve a contentful RED (e.g. copy epenthesis, Stanton and Zukoff forthcoming; cf. also Zuraw 2002). But the

[1] I employ the following abbreviations: 1, 2, 3 (first, second, and third person), ADD 'additive,' APPL 'applicative,' CAUS 'causative,' F 'feminine,' FUT 'future,' FV 'final vowel,' ICP 'intransitive copy pronoun,' INF 'infinitive,' M 'masculine,' NULLO 'null object,' O 'object,' PL 'plural,' PROG 'progressive,' R 'root,' REC 'reciprocal,' S 'subject,' SG 'singular,' TOT 'totality,' and VENT 'ventive.' In some cases in Bole, the same object suffix varies in its interpretation as direct or indirect depending on extraneous factors; these are glossed appropriately, but without comment.

present cases are not like copy epenthesis, since they are not phonologically motivated. Finally, even if it were possible to analyze doubling as RED, it would not simplify the morphological analysis below, since one would still need to account for the conditions under which the double appears, which do not reduce to phonology.

The remainder of the paper is organized as follows. The contexts in which doubling is found in Bole are characterized in §2. These contexts include outside-in conditioning, by which doubling is conditioned not only by the availability of applicable doublers and interveners, but also by the morpheme immediately following the locus of the second copy, as treated further in §3. Some comparative notes on West Chadic are then provided in §4, suggesting that doubling may have arisen *in situ* in the suffix string in Bole, though this premise is not critical for the analysis that follows. An analysis in terms of bigram morphotactics is developed in §5 and §6, the latter focusing on morphotactic extension (i.e. analogy in affix ordering). Finally, §7 concludes by considering other possible approaches to doubling in Bole and some outstanding questions.

2 Contexts for doubling in Bole

Bole, spoken in northeastern Nigeria (Yobe and Gombe States), is part of the Bole-Tangale subgroup of West Chadic (Afro-Asiatic). The focus here is the analysis of the rampant multiple exponence found in Bole's verbal system. Aspects of the verbal morphology that are not critical to this discussion are put aside; see Gimba (2000), Gimba and Schuh (2014), and references therein for paradigms and further background.[2] Multiple exponence in Bole involves the semantically gratuitous doubling of a suffix across another single suffix, as with plural subject agreement (PLS, boldface) in 8. In 8a–b, subject agreement /an/ doubles across object agreement /to/. In 8c, /an/ doubles across the totality extension (TOT), which might be glossed 'up' or 'well.' Note that the second copy of the doubled suffix precedes TOT in 8b but follows it in 8c.[3]

(8) a. [ŋgòr**án**táŋgó]
ŋgor **-an** -to **-an** -ko
tie -PLS -3FSGO -PLS -PERF
'they tied her'

b. [ŋgòr**án**tá**n**tì]
ŋgor **-an** -to **-an** -ti
tie -PLS -3FSGO -PLS -TOT
'they tied her up'

c. [ŋgòr**án**tù**ŋ**gó]
ŋgor **-an** -ti **-an** -ko
tie -PLS -TOT -PLS -PERF
'they tied up'

The allomorphy in 8 and elsewhere in this article is phonologically conditioned, reflecting local processes such as assimilation, ablaut, resolution, and tone rules (Schuh 2001). As such, it is not important for the treatment of doubling. In every case, both copies of the doubled morpheme are reflexes of the same underlying form (e.g. /an/ in the case of PLS); there is no suppletion, portmanteaux, etc.[4] Subject agreement is confined to the perfective, but doubling not involving subject agreement is also found in other aspects.[5] The perfective is marked by /ko/ = [ko ∼ go ∼ wo] (various tones and lengths), though it is not realized segmentally in all perfective forms (§3). I usually employ the Class A1 root *ŋgor(u)-* 'tie' in examples here.

[2]Gimba's (2000) paradigms give surface forms, such that the perfective suffix is not made explicit when it is segmentally null. As discussed in §3, the perfective conditions doubling even when segmentally null, so this point requires caution.

[3]In glossed forms from this point on, I give a surface phonetic form first, with tone, followed by a morphological parse (roughly, underlying form), without tone. This is not to imply that morphemes lack tones underlyingly; but their underlying tones, which are sometimes difficult to establish, are largely irrelevant for present purposes.

[4]The ventive suffix, for its part, comes in three suppletive allomorphs depending on aspect, but it is always the same version that doubles within a word. This concurrence is expected on the present approach. For example, the ventive that subcategorizes for the perfective could hardly be inserted in, say, a subjunctive form, regardless of how many times it is inserted.

[5]Limited subject agreement occurs also in the imperative, but does not furnish a context for doubling.

I term suffixes that can double DOUBLERS, suffixes that can intervene between doubles INTERVENERS, and suffixes that can immediately follow the second copy FOLLOWERS. All possible doublers, interveners, and followers are enumerated in Table 1. The doublers include all subject agreement, which comprises only two suffixes, namely, plural (any person) and feminine singular (2nd or 3rd person). The ventive, also known as *Entfernungserweiterung* ("indicates event initiated at a distance with effect at point of reference," Schuh p.c.), is also a doubler in both of its suppletive variants /in/ and /it/.[6] These variants are synonymous but the former is found in the perfective and the latter in the subjunctive (in Bole, Ngamo, etc.; probably reconstructable as such in Proto-Bole-Tangale).[7] The interveners include all object agreement as well as two extensions, namely, totality (TOT) and additive (ADD; "sort of 'pro-adjunct'; can indicate repetition," *id.*). Schuh (p.c.) suggests that TOT and ADD are "probably reconstructable" as **ti* and **di*, respectively, "but there are many shifts." In Bole, they are realized as such word-finally, but as [tu(ː)] and [du(ː)] nonfinally. TOT and ADD cannot cooccur within a verb, but either can combine with VENT. Finally, possible followers include TOT and PERF (including its segmentally null realization). As mentioned above, TOT can be both an intervener and a follower.

Table 1: Suffixal doublers, interveners, and followers in Bole

Doublers	/ak/ /in/	23FSGS VENT	/an/ /it/	PLS VENT
Interveners	/no/ /ʃi/ /ku/ /to/ /di/	1SGO 2FSGO 2PLO 3FSGO ADD	/mu/ /ko/ /su/ /ni/ /ti/	1PLO 2MSGO 3PLO 3MSGO TOT
Followers	/ko/	PERF	/ti/	TOT

A schematic representation of verb forms with doubling is provided in 9. The first copy of the doubler immediately follows either the verbal root or a post-root suffix such as ventive /aːkoː/ or a theme vowel, which do not participate in doubling.

(9)

			Doubler	**Intervener**		**Follower**
Root	(Suffix)	{	Subject Agr. Ventive	} {	Object Agr. Additive Totality	} Copy of Doubler {	Perfective Totality	}

A verb with doubling is built up progressively in 10. In 10a–b, NULLO indicates that the transitive verb is employed without an explicit object. When no subject agreement is explicit, the verb is interpreted as M3SG, as in 10a and 10c. Plural subject agreement doubles in 10d (across object agreement), but not in 10b (across aspect). Note that /to/ is realized with a low vowel nonfinally, as in (c–d).

(10) a. [ŋgórwòːyí]
ŋgor -ko -yi
tie -PERF -NULLO
'he tied it'

b. [ŋgòráŋgòːyí]
ŋgor -an -ko -yi
tie -PLS -PERF -NULLO
'they tied it'

c. [ŋgórtáːwó]
ŋgor -to -ko
tie -3FSGO -PERF
'he tied her'

[6]Suffixes given as /VC/ here might also be analyzed as /C/. This question is not critical for present purposes.

[7]A third suppletive allomorph of the ventive, namely, /aːkoː/, occurs in the future/progressive and habitual. This version is not involved with doubling.

d. [ŋgòr**án**tá**ŋ**gó]
ŋgor **-an** -to **-an** -ko
tie -PLS -3FSGO -PLS -PERF
'they tied her'

An example of doubling in the subjunctive (as opposed to perfective) is given in 11. In this case, in order to set up a context for doubling, all of the examples involve the totality extension, which is always final in 11. The ventive doubles in 11d (over object agreement), but not in 11b (over totality). Although TOT can be an intervener, an appropriate follower is not present in 11b to condition doubling. The subject is glossed X because subjunctives do not exhibit subject agreement. But they do exhibit object agreement, as 11c–d exemplify. The object must be interpreted as an indirect object in this context, though it is the same /to/ as above (see footnote 1; Gimba 2000:220).

(11) a. [ŋgòrtí]
ŋgor -ti
tie -TOT
'that X tie (it) up'

b. [ŋgòríttí]
ŋgor -it -ti
tie -VENT -TOT
'that X tie (it) up and bring it'

c. [ŋgòrtá**:**tì]
ŋgor -to -ti
tie -3FSGO -TOT
'that X tie (it) up for her'

d. [ŋgòr**ít**tá**t**tì]
ŋgor **-it** -to **-it** -ti
tie -VENT -3FSGO -VENT -TOT
'that X tie (it) up for her and bring (it)'

3 Outside-in conditioning

As mentioned in §2, doubling only occurs before an appropriate "follower" suffix, a case of OUTSIDE-IN CONDITIONING. Two followers are possible, namely, the totality extension /ti/ and the perfective aspect /ko/. The latter can be realized in any of its allomorphs, including segmentally null, as discussed in this section. Both followers were illustrated in §2. Non-doubling due to the lack of an appropriate follower is now illustrated in 12. In 12a, ventive /it/ cannot double before NULLO, despite the eligible intervener. In 12b–c, it cannot double when no suffix follows, again despite the interveners.

(12) a. [ŋgòríttà**:**yí]
ŋgor -it -to -yi
tie -VENT -3FSGO -NULLO
'that X tie her and bring her'

b. [ŋgòríttó]
ŋgor -it -to
tie -VENT -3FSGO
'that X tie (it) for her and bring (it)'

c. [ŋgòríttí témʃí]
ŋgor -it -ti temʃi
tie -VENT -TOT sheep
'that X tie up the sheep and bring it'

The lack of doubling in forms like 12 cannot be explained by phonology. In general, Bole words can end with plosives and nasals, as exemplified in 13a–b (Gimba and Schuh 2014). This is equally true for verbs, as illustrated by 13c–d. 13d furnishes a near-minimal pair vis-à-vis 12c. In 13d, subject agreement — /ak/, realized as [(a)t] due to assimilation — doubles across totality, yielding word-final [t]. Doubling is conditioned in 13d by the null perfective, which is absent from 12c, a subjunctive. There is thus a morphological difference between 12c and 13d, but the point here is that there is nothing wrong phonologically with word-final [t] arising from doubling. The explanation for the lack of doubling in 12c must be morphological.

(13) a. [dài ʃít]
'bright red'

b. [ájàp]
'amazement'

c. [ŋgòrát tèmʃí]
ŋgor -ak -ko temʃi
tie -23FSgS -Perf sheep
'she tied the sheep'

d. [ŋgòr**átt**ù**t** tèmʃí]
ŋgor **-ak** -ti **-ak** -ko temʃi
tie -23FSgS -Tot -23FSgS -Perf sheep
'she tied up the sheep'

Similarly, consider the "intransitive copy pronoun" (ICP), a subject agreement enclitic (Newman 1971, Schuh 1983, Gimba 2000:150–3). Gimba (2000) suggests that the ICP is the intransitive counterpart of totality, hence my translation '(completely).' Morphologically, the ICP comprises two parts, namely, [jìː] 'body' followed by a normal pronoun (see Table 1), as in 14. PlS does not double in 14, as it does not contain an eligible intervener.

(14) [daːndé 'yòrán jìːsú]
daːnde 'yor -an -ko =ji -su
children stop -PlS -Perf =body -3PlO
'the children (completely) stopped'

Doubling can apply before the ICP, but only when the (null) perfective is there to condition it, as in 15a–b. In other words, the ICP itself is not a follower, and does not condition doubling. This is clear from 15c, which has the ICP but no perfective suffix, and hence no doubling. Once again, as far as the phonology is concerned, doubling would be licit in 15c. As 15b shows, [jj] (as would occur with doubling) is distinct from [j]. The lack of doubling in 15c is due to the lack of an eligible follower, not phonology.

(15) a. [daːndé 'yòr**án**dù**n** jìːsú]
daːnde 'yor **-an** -di **-an** -ko =ji -su
children stop -PlS -Add -PlS -Perf =body -3PlO
'the children (completely) stopped again'

b. [ítá 'yòr**ád**dù**j** jìːtó]
ita 'yor **-ak** -di **-ak** -ko =ji -to
she stop -23FSgS -Add -23FSgS -Perf =body -3FSgO
'she (completely) stopped again'

c. ['yòríddí jìːnì]
'yor -it -di =ji -ni
stop -Vent -Add =body -3MSgO
'that he stop here (completely) again'

At least five considerations favor the underlying presence of perfective /ko/ in forms like 15a–b even when it is not realized segmentally on the surface. First, deletion occurs in a predictable context, namely, when /ko/ immediately precedes another word. Compare 16a, without deletion, to 16b, which adds 'yesterday' to the same verb, triggering deletion.

(16) a. [ŋgòr**án**táŋgó]
ŋgor **-an** -to **-an** -ko
tie -PLS -3FSGO -PLS -PERF
'they tied her'

b. [ŋgòr**án**tá**n** nzònó]
ŋgor **-an** -to **-an** -ko nzono
tie -PLS -3FSGO -PLS -PERF yesterday
'they tied her yesterday'

Second, the perfective suffix has effects on tone even when it is segmentally null. In particular, it blocks an otherwise general process of high tone spreading to a syllable not beginning with a voiced obstruent (Gimba 2000:140). For example, 'sheep' is /tèmʃí/. After a high-final verb, its low becomes high, as in 17a. But spreading is blocked in perfective forms such as 17b. This blocking makes sense if the perfective leaves a floating low tone even when it is segmentally unrealized. In this sense, the perfective is overtly realized even when segmentally null.

(17) a. [ŋgòrtí témʃí]
ŋgor -ti temʃi
tie -TOT sheep
'that X tie up the sheep'

b. [ŋgòrát tèmʃí]
ŋgor -ak -ko temʃi
tie -23FSGS -PERF sheep
'she tied the sheep'

Third, as seen in passing above, certain suffixes have final and nonfinal allomorphs. Such suffixes appear in their nonfinal forms before segmentally unrealized /ko/, again suggesting that /ko/ is overt. For instance, totality /ti/ is [tu(ː)] nonfinally, as in 18a. In 18b it remains [tuː] on the surface even though it is ostensibly final in its word, thanks to underlying /ko/.

(18) a. [ŋgórtùwó]
ŋgor -ti -ko
tie -TOT -PERF
'he tied (it) up'

b. [ŋgórtùː tèmʃí]
ŋgor -ti -ko temʃi
tie -TOT -PERF sheep
'he tied up the sheep'

Fourth, perfective /ko/ is stably realized on the surface in related languages where it is inferred in Bole (Schuh p.c.). Finally, without a uniform perfective suffix in the perfective (whether realized or not), it is more difficult to explain the distribution of doubling. For instance, doubling applies across TOT in the perfective (13d) but not in the subjunctive (12c). Schuh and I explain this contrast by invoking null /ko/ in the former. Without /ko/, one might imagine stipulating that doubling is constructionally limited to the perfective. But this would not work, since doubling occurs in the subjunctive under other circumstances. Moreover, it would miss that even in the perfective, doubling is limited to certain followers; the generalizations are positional (i.e. morphotactically conditioned), not constructional (i.e. featurally conditioned).

In sum, doubling is crucially conditioned by the suffix following the second copy. Two suffixes, the perfective and totality, trigger doubling as followers. The perfective triggers doubling even if it is segmentally unrealized, in which case other evidence, such as tone, confirms that it is still overt.

4 Comparative Bole-Tangale notes

Schuh (in his half of Ryan and Schuh 2010) highlights four innovations of Bole verbal morphology relative its West Chadic relatives: (1) the 23FSgS subject agreement suffix /ak/, (2) the null object suffix /yi/, (3) the elision of /ko/ (and /yi/) just discussed in §3, and (4) the suffix doubling that is the topic of this paper. Some Bole-Tangale cognates are provided in Tables 2 and 3, which cover Bole, Karekare, and two dialects of Ngamo. I adapt Schuh's transcription slightly for Bole to bring it into line with the rest of this paper (e.g. showing nasal assimilation), but do not do so for the other languages.

Table 2: Cognates in four Bole-Tangale languages: perfective verbs with a third-person plural subject. Subject agreement doubles in Bole.

	Karekare	Ngamo (Gudi)	Ngamo (Yaya)	Bole
	às-án-kò	ngàr-àn-kô	ngàr-án-kò	ŋgòr-áŋ-gòː-yí
+3FSgO	às-ân-tó	ngàr-àn-tò	ngàr-án-tŏ	ŋgòr-**án**-tá-**ŋ**-gó
+Tot	às-án-sì-kó	ngàr-án-kò	ngàr-án-tù-kó	ŋgòr-**án**-tù-**ŋ**-gó
+Tot +3FSgO	às-ân-tá-sì	ngàr-àn-tòː-tî	ngàr-án-tóː-tì	ŋgòr-**án**-tá-**n**-tì
+Vent	às-àː-néː-kò	ngàr-àː-nô	ngàr-à-nô	ŋgòrú-ŋ-gòː-yí
+Vent +3FSgO	às-àː-nê-tó	ngàrì-n-tò	ngàr-á-n-tŏ	ŋgòr-**ín**-tá-**ŋ**-gó

Table 3: Cognates in four Bole-Tangale languages: perfective verbs with a singular subject. The ventive doubles in Bole.

	Karekare	Ngamo (Gudi)	Ngamo (Yaya)	Bole
+Vent +3FSgO	às-nê-tó	ngàrì-n-tò	ngàrí-n-tŏ	ŋgòr-**ín**-tá-**ŋ**-gó
+Vent +Tot	às-néː-sí-kò	ngàrí-n-kò	ngàr-nó'ò	ŋgòrú-**n**-tù-**ŋ**-gó
+Vent +Tot +3FSgO	às-nê-tà-sí	ngàrì-n-tòː-tî	ngàrí-n-tóː-tì	ŋgòr-**ín**-tá-**n**-tì

These cognate sets suggest that the emergence of multiple exponence in Bole was not due to the univerbation of formerly separate words expressing agreement, such as light verb constructions akin to 'they tied it, they did.' Rather, the whole suffix string was likely already intact when Bole innovated doubling. Compare Yaya Ngamo [ngàr-án-tóː-tì] to Bole [ŋgòr-**án**-tá-**n**-tì] 'they tied up for her.' Note that the lengthened [óː] in Yaya corresponding to Bole [á-n] is presumably driven by open-syllable lengthening/closed-syllable shortening, an otherwise general process in Bole-Ngamo, and need not reflect compensatory lengthening.[8] Moreover, a light verb origin is implausible on the grounds that there were never light verbs to begin with. The perfective and totality markers never require supporting auxiliaries in Bole-Tangale; they are added directly to the root verbal complex. That said, these historical considerations are not critical for the synchronic analysis that follows, and are provided mainly by way of background. They are, however, consistent with my proposal, in that I argue that doubling is motivated by local analogical pressures in the form of bigram morphotactics.

5 Morphotactic analysis

Ryan (2010) argues that arbitrary affix ordering restrictions — which cannot be motivated by independent semantic, syntactic, or phonological considerations — are grammatically encoded as adjacency bigram constraints, as in 19, in which X and Y are (classes of) morphemes.

(19) **X-Y**: Penalize a candidate lacking X-Y.

For instance, such constraints can motivate counterscopal ordering, in which two mutually scoping affixes, such as the causative and reciprocal, are fixed in a certain order regardless of which scopes over the other. X-Y $\gg$ Y-X ensures that X and Y are realized in that order regardless of scope and other considerations. (Additional constraints penalize deleting or duplicating morphemes.) Ryan (2010)

[8] Reconstructing the Bole doubling pattern to Proto-Bole-Tangale would be less parsimonious because doubling would then need to be lost independently in multiple branches, including at least Karekare and Ngamo (separately in the latter case, assuming that Bole and Ngamo form a subgroup to the exclusion of Karekare).

argues that adjacency bigrams are superior to other proposals for arbitrary ordering including precedence bigrams ("X must precede Y"; cf. Paster 2006, Caballero 2008), affix alignment (cf. Trommer 2003), affix movement (cf. Embick and Noyer 2001), and a monolithic template of more than two position classes (cf. Hyman 2003). Adjacency bigrams are argued to better capture nontransitive ordering restrictions, gradient variation in ordering (including predicting possible vs. impossible types of variation), analogical extension in ordering (as discussed presently), learnability (e.g. how are language-specific movement rules or monolithic templates inferred if not from surface adjacency relations?), and context-sensitivity in ordering (e.g. only X-Y is permitted unless Z immediately follows, in which case only Y-X is permitted).

To exemplify two such cases before returning to Bole, first, Chumbivilcas Quechua (Muysken 1988) exhibits nontransitivity in ordering (Ryan 2010). Consider the three verbal suffixes *ri* 'inchoative,' *schi* 'assistive,' and *na* 'reciprocal.' In a doubly derived verb, *ri* can only precede *schi* and *schi* can only precede *na*. If transitivity held, this would entail that *ri* precede *na* when the two cooccur. But in fact *na-ri* is the only acceptable order. This system is captured by adjacency bigrams in 20, which contains three tableaux. They cannot be motivated by a template or position class system.

(20)

			ri-schi	schi-na	na-ri
a.	☞	-ri-schi-		*	*
b.		-schi-ri-	*!	*	*
a.	☞	-schi-na-	*		*
b.		-na-schi-	*	*!	*
a.	☞	-na-ri-	*	*	
b.		-ri-na-	*	*	*!

Second, consider context-sensitive reorderability. In Tagalog, for instance, the "contemplated aspect" reduplicant RED is free to occur either immediately before or immediately after the prefix *ka* ("telic") when the root ("R") follows: RED-*ka*-R-an ∼ *ka*-RED-R-an (Schachter and Otanes 1972, Ryan 2010:766). But in verbs of the form *ka-pag*-R, RED cannot immediately follow *ka*. Schematically, X-Y ∼ Y-X, except before Z. This case is akin to outside-in conditioning, in that it cannot be motivated by precedence relations alone.[9] Adjacency conditions must be invoked, as in 21, which is a simplified sketch (and not the only possible bigram analysis of this fragment). See Ryan (2010) for a full analysis of Tagalog and for other cases of context-sensitive reorderability.

(21)

			ka-pag
a.	☞	RED-ka-R-an	*
b.	☞	ka-RED-R-an	*
a.	☞	RED-ka-pag-R	
b.		ka-RED-pag-R	*!
c.	☞	ka-pag-RED-R	

Returning to Bole, I employ the representative data set in 22, which abstracts away from allomorphy. These data include doublers both doubling (d, i, k) and failing to double (b, f, g, j), such that all of the conditioning variables discussed in §2 are instantiated. They also include cases of totality as both an intervener (k) and a follower (i). R is the root, *ko* PERF, *yi* NULLO, *an* PLS, *to* 3FSGO, *ti* TOT, and *it* VENT. Forms with *ko* are perfective, and those without it are subjunctive. Because allomorphy is not analyzed here, the segmental (but not tonal) elision of suffixes is moot.

(22)

a.	{R, ko, yi}	R-ko-yi
b.	{R, an, ko, yi}	R-an-ko-yi
c.	{R, to, ko}	R-to-ko
d.	{R, an, to, ko}	R-**an**-to-**an**-ko
e.	{R, ti}	R-ti
f.	{R, it, ti}	R-it-ti
g.	{R, it, to}	R-it-to
h.	{R, to, ti}	R-to-ti
i.	{R, it, to, ti}	R-**it**-to-**it**-ti
j.	{R, it, to, yi}	R-it-to-yi
k.	{R, an, ti, ko}	R-**an**-ti-**an**-ko

[9]For just the second tableau in 21, one could invoke two precedence constraints to the effect that "RED must precede *ka*" and "RED must follow *pag*," freely ranked with each other to generate the variation. However, (b) in the first tableau is then not generated. There is no system-wide solution with precedence.

The grammar is then set up as follows. Every observed bigram is encoded as a constraint. Presumably, the learner posits these constraints as it encounters pairs of affixes, effectively tracking possible transitions in the language. For example, *R-ko* and *ko-yi* are posited as constraints, while *ko-R* and *yi-ko* are not. In some cases, both orders are observed, in which case both are posited as constraints (e.g. *it-to*, *to-it*). In total, 17 bigrams are observed in 22. Additionally, I posit a constraint against doubling, say, *FeatureSplit (henceforth *Split) (Xu and Aronoff 2011), though others (including *Struc) may work equally well here. *Split penalizes each instance of multiple exponence (here, each double). An input for this simulation is taken to be an unordered set of morphemes (with no duplicates, assuming that doubling is not semantically motivated). Inputs in principle encode more information than this, but this is all of the information that is necessary for present purposes. Inputs are given on the left side in 22, in braces. The candidate set for each input includes all ordering permutations and all possible suffix duplication schemes. For example, for {R, ko, yi}, it includes *R-ko-yi*, *R-yi-ko*, *yi-ko-R*, *R-ko-ko-yi*, *R-ko-yi-ko*, *ko-R-ko-R-yi*, and so forth. In principle, candidates containing more than two repetitions are available (e.g. *R-ko-ko-ko-yi*), but for the purposes of simulations I cap doubles at two copies. More complex candidates are harmonically bounded in this case, so this simplification is not harmful.

A tableau file was generated automatically and submitted for evaluation to OT-Help2 (Staubs et al. 2010), yielding the grammar in 23. Additionally, I assume that RealizeMorpheme (Kurisu 2001) is undominated, such that a morpheme in the input cannot go unrealized altogether (segmental Max might also work, but it is often violated). Affixes also cannot be inserted if they are not input-licensed; I assume that this follows from Dep, not shown.[10]

(23)
Stratum 1:	R-an, R-it, an-ko, an-ti, an-to, it-ti, it-to, ko-yi, ti-an, to-an, to-yi
Stratum 2:	R-to, to-ko, to-ti, *Split
Stratum 3:	R-ko, R-ti, to-it

Tableau 24 illustrates doubling across object agreement before totality. Only a handful of contenders are shown. The constraints in the top stratum can only be jointly satisfied if *Split is violated, resulting in doubling.

(24)
	{R, it, to, ti}	R-an	R-it	an-ko	an-ti	an-to	it-ti	it-to	ko-yi	ti-an	to-an	to-yi	R-to	to-ko	to-ti	*Split	R-ko	R-ti	to-it
a.	☞ R-it-to-it-ti	*		*	*	*			*	*	*	*	*	*	*	*	*	*	
b.	R-it-to-ti	*		*	*	*	*!		*	*	*	*	*	*	*	*	*	*	*
c.	R-to-it-ti	*	*!	*	*	*		*	*	*	*	*		*	*		*	*	

Tableau 25 illustrates totality as an intervener rather than follower.

(25)
	{R, an, ti, ko}	R-an	R-it	an-ko	an-ti	an-to	it-ti	it-to	ko-yi	ti-an	to-an	to-yi	R-to	to-ko	to-ti	*Split	R-ko	R-ti	to-it
a.	☞ R-an-ti-an-ko		*			*	*	*	*		*	*	*	*	*	*	*	*	*
b.	R-an-ti-ko		*	*!		*	*	*	*	*	*	*	*	*	*		*	*	*
c.	R-ti-an-ko	*!	*		*	*	*	*	*		*	*	*	*	*		*		*

This grammar as it stands is purely morphotactic, and one might object that it is brute force. However, bigram morphotactics is not brute force in general; there are plenty of logically possible ordering scenarios that it cannot generate (Ryan 2010, Ryan and Schuh 2010). But to the extent that the present analysis is brute force, it can be considered a baseline or proof of concept. As one adds general principles of affix ordering to the grammar, the morphotactic component can be simplified. But as long as bigram morphotactics are available, a working analysis of Bole morphology is ensured. Moreover, given arbitrary differences between languages, morphotactic constraints cannot be whittled away entirely (*ibid.*). Compare, for instance, Bole and Yaya Ngamo in Table 2. The paradigms are close, the main difference being that Bole adds multiple exponence. The theory needs to generate both cases.

6 Morphotactic extension

Bigram morphotactics can also motivate the emergence of doubling through a process called morphotactic extension, which essentially amounts to analogy in affix order. Ryan (2010) describes morphotactic extension for Tagalog aspectual red (see above) using the following example. In relatively simple verbs like 26, red usually occupies the second position in the word.

[10]Doubling does not violate Dep, as both copies correspond to the input.

(26) a. ma-RED-ka-R
b. pag-RED-pa-R

However, in more complex cases, such as *ma-ka-pag-pa*-R, it can vary freely between the second position and a position deeper into the prefix string, as depicted by Figure 1. In this case, the relative frequencies of the two options are 75% and 25%, respectively. Morphotactic extension explains this optionality as being driven by a tension created by forms such as 26. 26a supports RED between *ma* and *ka*; 26b supports it between *pag* and *pa*. In *ma-ka-pag-pa*-R, both options are available, and indeed both are employed. Because *SPLIT is highly ranked in Tagalog, free variation emerges, not doubling.

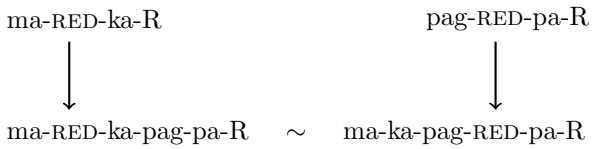

Figure 1: Morphotactic extension as a source of variation in Tagalog.

A similar process of analogy can motivate the emergence of doubling in Bole, as schematized in Figure 2. However, because Bole ranks *SPLIT low, doubling rather than free variation emerges as the optimal response to the morphotactic tension created by the simpler forms.

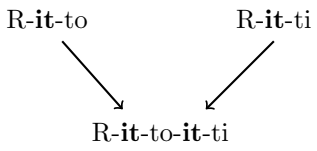

Figure 2: Morphotactic extension as a source of doubling in Bole.

To be somewhat more concrete, consider the hypothetical pre-Bole to Bole developments in 27.[11] Pre-Bole in 27 is based on Ngamo.

(27)

	Pre-Bole		Bole	Parse
a.	*R-an-ko	>	R-an-ko	R-PLS-PERF
b.	*R-ti-ko	>	R-ti-ko	R-TOT-PERF
c.	*R-an-ti-ko	>	R-**an**-ti-**an**-ko	R-PLS-TOT-(PLS)-PERF

Imagine that a hypothetical bigram-morphotactic learner is exposed to these three pre-Bole forms with relative frequencies of, say, 10, 2, and 1, respectively. The bigram learner matches its training data here, as it should (since this is essentially the stable situation in Ngamo). But in order to simulate language change, something has to give; learners must occasionally converge on a different grammar. I shall now illustrate that with bigram morphotactics, the pre-Bole learner is particularly conflicted between R-*an-ti-ko* (without doubling) and R-***an***-*ti*-***an***-*ko* (with doubling) above and beyond all other logically possible contenders. One might investigate this question in a few ways. First, one could expose the learner to fewer training data and then wug-test it. Second, one could increase the learner's smoothing factor (i.e. propensity to generalize). Third, in a framework such as maximum entropy Harmonic Grammar (maxent HG) in which candidates are assigned probabilities, one could check which candidate has the second-highest probability, and with what proportion. I pursue this last tack here.

Using maxent HG learning software by Wilson and George (2008) with a relatively strong smoothing factor of $\sigma^2 = 100$, I train the grammar on the data in 27 with the aforementioned relative frequencies. As always, all ordering and doubling permutations are included in the spreadsheet as candidates, and constraints include all observed bigrams as well as *SPLIT. For the key input {R, an, ti, ko}, R-*an-ti-ko* is by far the most probable output (98.9%), as expected. Among innovative mappings, however, R-***an***-*ti*-***an***-*ko* is the most probable by a wide margin (0.6%). The probability of R-***an***-*ti*-***an***-*ko* increases as σ^2 decreases, but under any reasonable σ^2, it is the best unfaithful mapping; that, not the precise values, is the point. To unpack why this particular unfaithful mapping is the silver medalist (and with

[11] For simplicity, I omit the NULLO *yi* from modern Bole in 27a, as it was an independent innovation, irrelevant here.

nonnegligible probability), consider that the learner sees mostly R-*an-ko*, which supports the constraints *R-an* and *an-ko*. Now the learner has to decide what to do with the less common combination {R, an, ti, ko}. It has two options. First, it can maintain a one-to-one mapping, but at a morphotactic cost: *R-an* and *an-ko* cannot be simultaneously satisfied without doubling, assuming that *ti* also has to intervene between R and *ko*. But doubling allows the learner to simultaneously satisfy both morphotactic constraints, as sketched in 28. Admittedly, as an anonymous referee points out, this discussion treats only the actuation of doubling, and not its eventual entrenchment as the only grammatical outcome. I leave intergenerational modeling to future work.

(28)

	Output	Satisfies	At The Expense Of
a.	R-an-ti-ko	R-an, *Split	an-ko
b.	R-**an**-ti-**an**-ko	R-an, an-ko	*Split

7 Other approaches to vacuous affix repetition

I now address two alternative approaches to semantically gratuitous affix repetition before concluding with some further issues. First, Downing (2005) analyzes Jita causative *y* doubling (as seen in §1) using stem alignment (cf. McCarthy and Prince 1993) and Optimal Paradigms (McCarthy 2005). The constraint motivating doubling is Align-y "align Caus to the end of the the stem" (assuming that the inner copy also surfaces faithfully). For stem alignment to work for Bole, some domain must be specified as the stem, to which the second copy aligns. But outside-in conditioning renders such a stem undefinable. For example, in 29, doubling is absent. Therefore, on the stem analysis, one would have to say that the relevant stem closes before *to* (as discussed in §3, the absence of doubling in cases like 29 cannot be motivated phonologically). But in other contexts, doubling (even of the same ventive suffix) is permitted across *to*, resulting in a contradiction. Moreover, Tot can occur both before and after the second copy. If Tot closes the stem when it is final, such that the second copy immediately precedes it, it must do so when it is nonfinal as well, predicting that doubling never crosses Tot.[12]

(29) [ŋgòríttàːyí]

ŋgor -it -to -yi
tie -Vent -3FsgO -NullO

'that X tie her and bring her'

Second, Caballero and Inkelas (2013) analyze doubling in terms of weakness-driven augmentation. In such cases, the inner morph is insufficiently salient, so it is augmented with an outer morph. For example, consider Choguita Rarámuri *sú-n-ki-ma* 'sew-Appl-Appl-Fut.sg,' as seen in §1. Caballero and Inkelas (2013) (see also Caballero 2008, 2010, 2011) observe of this form that "an inner, lexicalized marker (e.g. *-n*) weakly marks Applicative, while a second, regular, outer exponent (*-ki*) brings target meaning to minimum threshold level." They implement this insight using gradient violations of M-Faith (e.g. *-n* contributes only 0.5 to the exponence of the applicative). In general, weaker affixes are characterized by properties such as less productivity, less segmentability, more allomorphy, and smaller size (Hay and Plag 2004). This type of analysis is well motivated for many cases, but not viable for the Bole system treated here. First, note that unlike Choguita Rarámuri, the same morpheme is copied in Bole, so there is no issue of differential salience or productivity. There is also little independent evidence of weakness: All doublers are fully productive and regular, frequently appearing undoubled. For example, plural subject agreement /an/ is extremely common, and usually occurs undoubled, as in most simple verb forms. This approach would also have to contend with outside-in conditioning, in that suffixes like /an/ double only if certain other suffixes follow. But there is no look-ahead in cyclical construction of the type that Caballero and Inkelas (2013) assume.[13] Finally, even if one assumes weakness, one still needs a theory of morphotactics to determine the doubles' placement. The bigram analysis of Bole is parsimonious in that it motivates

[12] An analysis in terms of morpheme-morpheme alignment is viable, since it essentially duplicates bigram morphotactics. However, morpheme-morpheme alignment is more powerful than bigrams, and the extra power is unneeded. For a pair of morphemes X and Y, four alignment constraints are possible (left-left, left-right, right-left, right-right), vs. two bigrams (X-Y, Y-X). Furthermore, alignment introduces an unneeded dimension of gradience, since its violations scale with distance, whereas bigrams are categorical, evaluating only strict adjacency (cf. McCarthy 2003 pro categorical constraints).

[13] Cyclicity does not rule out that the suffix might be copied and then deleted in a later cycle in a kind of "Duke-of-York gambit." But in that case establishing a phonological motivation for deleting is problematic.

both the placement of the doubles and the fact of doubling with the same machinery: Doubling is driven by the same morphotactic constraints that control affix order in non-doubling contexts.

In sum, given these issues with the two aforementioned alternative approaches to multiple exponence, I pursue a morphotactic analysis of Bole here, though I do not contest the other approaches in general. The present analysis leaves some issues unresolved. First, it ignores allomorphy, allomorphy that is so extreme in some instances that it eliminates surface traces of the affix altogether, as with the deletion of perfective /ko/. While segmentally deleted /ko/ can leave tonal effects in its wake, in some forms, such tonal evidence is unavailable, meaning that /ko/ is realized wholly abstractly. The morphotactic analysis here takes /ko/ as a given in such cases, assuming that learners posit it even when null, as supported by various evidence in §3. But this is a substantial promissory note. Second, I gloss over the issue of whether the morphemes treated here are best analyzed as suffixes or enclitics, assuming them to be the former (as in Ryan and Schuh 2010, but cf. Gimba 2000). At any rate, this question is not critical for the bigram analysis, since bigrams can apply equally to suffixes and enclitics, and the issues are the same either way. Finally, the present analysis appears to miss a generalization: All doublers have the form /VC/ (or possibly just /C/ depending on one's analysis of the morphophonology),[14] while all interveners (and followers) have the form /CV/. On the present approach, this is a coincidence, but not a massive one, since only a few suffixes are doublers (viz. the ventive and two subject agreement suffixes), and the generalization may have an unrelated explanation, since doublers and interveners tend to occupy different positions of the suffix string. These questions invite further research.

References

Caballero, G. (2008). *Choguita Rarámuri (Tarahumara) phonology and morphology* (Unpublished doctoral dissertation). University of California, Berkeley.

Caballero, G. (2010). Scope, phonology and templates in an agglutinating language: Choguita Rarámuri (Tarahumara) variable suffix ordering. *Morphology*, *20*(1), 165–204.

Caballero, G. (2011). Multiple exponence and the phonology-morphology interface. *Proceedings of the North East Linguistics Society (NELS)*, *39*, 177–190.

Caballero, G., & Harris, A. (2012). A working typology of multiple exponence. In F. Kiefer, M. Ladányi, & P. Siptár (Eds.), *Current issues in morphological theory: (ir)regularity, analogy and frequency* (pp. 163–188). Amsterdam: John Benjamins.

Caballero, G., & Inkelas, S. (2013). Word construction: Tracing an optimal path through the lexicon. *Morphology*, *23*(2), 103–143.

Downing, L. (2005). Jita causative doubling provides optimal paradigms. In L. Downing, T. A. Hall, & R. Raffelsiefen (Eds.), *Paradigms in phonological theory* (pp. 122–144). Oxford University Press.

Embick, D., & Noyer, R. (2001). Movement operations after syntax. *Linguistic Inquiry*, *32*(4), 555–595.

Gimba, A. M. (2000). *Bole verb morphology* (Unpublished doctoral dissertation). University of California, Los Angeles.

Gimba, A. M., & Schuh, R. G. (2014). *Bole-English-Hausa dictionary and English-Bole wordlist* (Vol. 148). Berkeley: University of California Press.

Harris, A. (2017). *Multiple exponence*. Oxford: Oxford University Press.

Hay, J., & Plag, I. (2004). What constrains possible suffix combinations? On the interaction of grammatical and processing restrictions in derivational morphology. *Natural Language and Linguistic Theory*, *22*, 565–596.

Hyman, L. (2003). Suffix ordering in Bantu: A morphocentric approach. *Yearbook of Morphology*, *2002*, 245–281.

Kiparsky, P. (2010). Dvandvas, blocking, and the associative: The bumpy ride from phrase to word. *Language*, *86*(2), 302–331.

Kurisu, K. (2001). *The phonology of morpheme realization* (Unpublished doctoral dissertation). University of California, Santa Cruz.

[14] For example, Schuh (2001) and Ryan and Schuh (2010) takes PLS to be /an/ (parsing surface [-an-] in some cases), while Gimba (2000) takes the same [a] to be part of the stem, or a theme vowel. Similarly, Schuh (2001) gives the 23FSGS as "/(a)G/" (where G indicates the first part of a geminate), avoiding the issue.

Maclachlan, A. (1989). The morphosyntax of Tagalog verbs: The inflectional system and its interaction with derivational morphology. *McGill Working Papers in Linguistics*, *6*, 65–84.
McCarthy, J. (2003). OT constraints are categorical. *Phonology*, *20*, 75–138.
McCarthy, J. (2005). Optimal paradigms. In L. Downing, T. A. Hall, & R. Raffelsiefen (Eds.), *Paradigms in phonological theory* (pp. 170–210). Oxford University Press.
McCarthy, J., & Prince, A. (1993). Generalized alignment. In G. Booij & J. van Marle (Eds.), *Yearbook of morphology* (Vol. 1993, pp. 79–153). Dordrecht: Kluwer.
Muysken, P. (1988). Affix order and interpretation: Quechua. In M. Everaert, A. Evers, R. Huybregts, & M. Trommelen (Eds.), *Morphology and modularity.* Holland: Foris Publications.
Newman, P. (1971). Transitive and intransitive in Chadic languages. In V. Six (Ed.), *Afrikanische Sprachen und Kulturen, ein Querschnitt (Festschrift J. Lukas)* (Vol. 14). Hamburg: Hamburger Beiträge zur Afrika-Kunde.
Paster, M. (2006). Pulaar verbal extensions and phonologically driven affix order. *Yearbook of Morphology*, *2005*, 155–199.
Ryan, K. M. (2006). *The Rig-Vedic fate of the devatā dvandva (double dual).* (MS., University of California, Los Angeles)
Ryan, K. M. (2010). Variable affix order: grammar and learning. *Language*, *86*(4), 758–791.
Ryan, K. M., & Schuh, R. G. (2010). Suffix doubling and suffix deletion in Bole. (Talk handouts, Joint UCLA-USC Phonology Seminar, April 22)
Schachter, P., & Otanes, F. (1972). *Tagalog reference grammar.* Berkeley, California: University of California Press.
Schuh, R. G. (1983). The evolution of determiners in Chadic. In E. Wolff & H. Mayer-Bahlburg (Eds.), *Studies in Chadic and Afroasiatic linguistics* (pp. 157–210). Hamburg: Helmut Buske.
Schuh, R. G. (2001). Sources of gemination and gemination as a morpheme in Bole. *32nd Annual Conference on African Linguistics (ACAL 32).*
Stanton, J., & Zukoff, S. (in press). Prosodic identity in copy epenthesis: Evidence for a correspondence-based approach. *Natural Language & Linguistic Theory.*
Staubs, R., Becker, M., Potts, C., Pratt, P., McCarthy, J. J., & Pater, J. (2010). *OT-Help 2.0. Software package.* (Amherst, MA: University of Massachusetts, Amherst)
Stump, G. (1991). A paradigm-based theory of morphosemantic mismatches. *Language*, *67*, 675–725.
Trommer, J. (2003). The interaction of morphology and syntax in affix order. *Yearbook of Morphology*, *2002*, 283–324.
Wilson, C., & George, B. (2008). *Maxent grammar tool.* (Software package. University of California, Los Angeles)
Xu, Z., & Aronoff, M. (2011). A realization Optimality Theory approach to blocking and extended morphological exponence. *Journal of Linguistics*, *47*(3), 673–707.
Zuraw, K. (2002). Aggressive reduplication. *Phonology*, *19*(3), 395–439.