Reduplication and verbal number in Mẽbengokre*

Andrés Pablo Salanova

February 27, 2011

Abstract

Reduplication is often found cross-linguistically as a marker of plurality and emphasis. Also quite common are cases where reduplicative morphemes encode lexical aspectual distinctions such as repeated action. In several Jê languages, though not very productively, reduplication is one among several morphemes indicating the plural of verbs, or inherently repetitive eventualities.

In this paper we present data from Mẽbengokre, and show how reduplicative morphemes are treated in the language. Two separate processes are identified: one (relatively uninteresting) process associated to onomatopoeia, found in several names for plant and animal species, and another more grammatical one (though in somewhat of an archaeological sense), where reduplication is used with some verbs to indicate repeated action and possibly transitivity. This process may be related to a more productive process found in some other Jê languages to indicate verbal plurality. We establish that all the aspectual meanings introduced by reduplicative morphology essentially boil down to simple pluralization of events, in a way that is consistent to what other pluralizing morphemes do.

1 Introduction

Mẽbengokre is a Jê language spoken in north-central Brazil by two indigenous nations, the Xikrin and the Kayapó, numbering over ten thousand individuals in total. It is closely related to Apinayé (whose phonology is described by Oliveira 2005, Burgess and Ham 1968 and Callow 1962), Suyá (described by Santos 1997), Timbira (whose phonology is described by Popjes and Popjes 1986, and Alves 1999),

*Our warmest thanks to our main language consultants, Bep Kamrêk Kayapó and Ikrô Kayapó, for their patience in teaching us their language. We would also like to thank María Luisa Rivero, the editors of the volume and an anonymous reviewer for helpful comments on a previous draft of this paper.

and Panará (described by Dourado 2001). Other languages in the family include Xavante, Xerente, Kaingang (described in Wiesemann 1972 and Cavalcante 1987), and Xokleng. Though these are quite distant from Mẽbengokre, all languages in the family are strikingly similar in their overall morphosyntactic structure.

There are several non-concatenative morphological phenomena in Mẽbengokre, most notably initial and final truncation. We will refer to these only briefly in this paper, in section 3; the interested reader should consult Salanova (2004, 2009). Reduplication itself is somewhat limited in scope and productivity in Mẽbengokre. Nevertheless, we will show that reduplication has a well-delimited linguistic function in the language, and that interesting comparative observations may be made with other languages of the family. More solid conclusions await further research. Even though there now exist preliminary grammatical sketches for most of the languages of the family, the lexicon has practically not been studied at all in any of them.

2 Reduplicative phenomena in Mẽbengokre

The data for this study come from our own field notes, collected on several trips to Kayapó and Xikrin villages over a period of 15 years, and from the examination of a draft Mẽbengokre-Portuguese lexicon.[1] In order to have an unbiased preliminary overview of reduplicative phenomena in Mẽbengokre, we have set out by looking at all lexical items where at least a CV sequence is repeated, whether there is additional evidence for morphology or not, and even if a base is not independently attested in our notes. As will be seen, a decent case can be made for the derived nature of the reduplicated forms that express verbal number, even if reduplication in this domain is not as robust as elsewhere in the language family.

[1]This consists of a Toolbox database gathering all published and unpublished lexical data on Mẽbengokre, some of it corrected and augmented by the author and by Maria Amélia Reis Silva in the field. The data include materials by Pe. Antonio Sala, Horace Banner, Earl Trapp, Terence Turner, Joan Bamberger, Pe. Renato Trevisan, Ruth Thomson, Isabelle Giannini, Darrell Posey, and Vanessa Lea, and were collected and organized by the latter, with whom the author and Reis Silva collaborated between 1998 and 2000.

In terms of form, we are able to identify C(C)V, C(C)VC, and CVCV reduplicants:[2,3]

(1)	*gogo*	'to shake, make noise'
	kaprêprêk	'to give a beating'
	karõrõ	'to snore, roar'
	keket	'to laugh'
	rôrôk	'to fall (plural)'
	tatak	'to clap, type'
	mrômrôtire	'tadpole'
	ràràr	'orange (color)'
	rara	'box-shaped woven basket with lid'
	'ã'ãre	'seagull (?)'

[2]In this paper, contrary to our previous practice, we follow the most common orthographic conventions used to write Mẽbengokre, rather than a broad phonetic transcription. The symbols stand for their usual IPA equivalents, except for r = /ɾ/, *'* = /ʔ/, *nh* = /ɲ/, *ng* = /ŋ/, *x* = /tʃ/, and *dj* = /dʒ/; *ê* = /e/, *e* = /ɛ/, *ô* = /o/, *o* = /ɔ/, *y* = /ɯ/, *à* = /ʌ/, *ỳ* = /ɤ/, *ỹ* = /ɯ̃/, and *ã* = /ʌ̃, ã/. Stress is in the final syllable. Syllable-final *r* is normally followed by a stressless epenthetic vowel which is not represented in the orthography. In addition, the suffix *-re* 'diminutive', which appears in several names of animals, is stressless. For discussion of the segmental phonology of Mẽbengokre, see Salanova (2001).

[3]The following abbreviations are used in the examples: 1, 2, 3, 12 – affixal markers for first, second, third and first person inclusive; when followed by ERG, NOM or DAT, they denote the free-standing person pronouns; 3AC is the (accusative) third person marker used with some transitive verbs and adpositions; 2>3 is a third person object marker used when the subject is second person; REFL – reflexive; NFUT – non-future; NEG – negation; ANTICAUS – anticausative; ANTIPASS – antipassive; V, N – verbal (finite) and nominal (non-finite) form of the verb, respectively; SG, PL – singular, plural (in verbs); IDEOPH – ideophone; RED – reduplicant; CL – verbal "classifier", explained in the text when it comes up. Affixes are separated by hyphens (-), while clitics are separated by an equals sign (=). Infixes are surrounded by bullets (•). Morpheme boundaries are not indicated where they are irrelevant.

Mẽbengokre makes no distinction of gender in its pronouns or agreement. In the translations, we use 'he', 'she' or 'it' indistinctly. Likewise, no morphological distinction is made between the present and the past, and our translations are often erratic in this regard.

(2) *prõrprõt* ‘to float up and down’

tertet ‘to tremble from being cold’

dujduj ‘pied lapwing (*Hoploxypterus cayanus*)’

porpot ‘pauraque (*Nyctidromus*)’

wãnhwãnh ‘show-off’

(3) *kôtêkôtê* ‘blue-crowned motmot (*Momotus momota*)’

kurakura ‘acará-boi (fish of the *Chichlidae* sp.)’

muremure ‘grey-winged trumpeter (*Psophia crepitans*)’

The last set is decidedly marginal, has an unusual trochaic prosody, and always designates animal species, strongly suggesting onomatopoeia in a very narrow sense.

One can establish two broad semantic classes in the data, with just a few words left out: live species, and verbs of repeated action. The former seem to have peculiar phonology even when the reduplicant is smaller: *’ã’ãre* ‘seagull’, for instance, can be pronounced with an independent stress on each *’ã*, something which can never befall a verb such as *’ĩ’ĩk* ‘to hang (pl., intr.)’.

Furthermore, one can never identify a base in the lexicon from which the reduplicated forms in the first class can be formed. While in reduplicants denoting repeated action this is not always possible, there is certainly a large class of them where reduplication is paradigmatic. For this reason, we make a division between two processes of reduplication: imitative reduplication, which we will not discuss further, and what we may call “reiterative” reduplication, which will be addressed in what remains of the paper.[4]

The latter group may be further subdivided into those forms for which a base

[4]By naming these two classes in this way we do not wish to state *a priori* that one of them is more onomatopoetic or sound-symbolic than the other. The labels were chosen for convenience.

We will exclude from consideration the handful of words that are in neither of these semantic categories: *ràràr* ‘orange-colored’, *wãnhwãnh* ‘show-off’, *rara* ‘box-shaped basket’. We believe this exclusion to be justified, given that, like for the imitative reduplicants, no base for the reduplicated form can be identified in the lexicon. Nothing of substance hinges on this exclusion. Below we will address the more serious question of whether some of the ‘verbs of repeated action’ shouldn’t be themselves considered onomatopoeias.

may be readily identified, (4), and those for which it can't, (5). The reader may rest assured that, as the paper progresses, we will address the other things that happen to these forms, or assert that they are irrelevant.

(4)	*totyktyk*	'to strike repeatedly'	*totyk*	'to strike'
	kyjkyj	'to make many scratches'	*kyj*	'a scratch or cut'
	krãkrãk	'to swallow'	*tokrãk*	'to swallow at once'
	tatak	'to clap, type'	*tak*	'to hit'

(5)	*kypkyp*	'to chew'
	kryjkryj	'to scratch'
	gogo	'to shake, make noise'
	ngàrngàt	'make hissing sound'

We contend that no thick line should be drawn *a priori* between these two sets, despite the fact that in the second set there is no reduplication in the technical sense.[5,6] As we will show later, this type of reduplication in Mẽbengokre is the expression of a derivational process, rather than of inflection. As in other cases of derivation, productivity is not complete, and morphophonological and semantic idiosyncrasy can often be observed.

As for the phonology, once we have excluded obvious onomatopoeia, we may state that what we have left (i.e., "reiterative reduplication") fits a C(C)V template or C(C)VC template. Furthermore, if one considers the "reiterative" cases in (1), it is apparent that the reduplicant is a prefix or an infix inserted before the last, stressed, syllable (cf. *kaprêprêk*, *karõrõ*).

In fact, if one permits infixation, it's possible to derive reduplicated forms such as *tatak* and *gogo* with a reduplicative infix or suffix after the last vowel the word, something which might in principle seem more straightforward, since one would not have prefixation side by side with infixation, but rather a single process of infixation.

[5]See, e.g., the definition assumed by the Graz Reduplication Project, among others.

[6]In section 7 we do draw a line. We hope, dear reader, that you will concur with us that it is justified *a posteriori*.

A comparison of cases of C(C)VC and C(C)V reduplication shows that the bigger picture is simpler, however, if one assumes that the reduplicant is a prefix or infix appearing before the last syllable:

(6) C(C)V reduplication

Putative base		Infixation before main syllable	Infixation at the end
go	→	*go-go*	*go-go*
ket	→	*ke-ket*	*ke•ke•t*
kaprêk	→	*ka•prê•prêk*	*kaprê•prê•k*

(7) C(C)VC reduplication

Putative base		Infixation before main syllable	Infixation at the end
kyj	→	*kyj-kyj*	**ky•ky(j)•j*
totyk	→	*to•tyk•tyk*	**toty•ty(k)•k*

That is, to get the right results, one would have to say that C(C)VC reduplicants are prefixes or pre-main-syllable infixes, while C(C)V reduplicants are post-stressed-vowel infixes. It's confusing, right? That's why we chose the other analysis.

Further down, we will argue that considering all the reduplicants to be prefixes or pre-main-syllable infixes fits in with what we know about Mẽbengokre verbal morphology more generally.

Yet another possibility is that reduplication always follows a C(C)VC template, and coda consonants get dropped in certain environments. In this view, coda consonants get dropped in non-final syllables only, regardless of whether the non-final syllable is the reduplicant (i.e., when the reduplicant is a prefix) or the stem itself (i.e., when the reduplicant is a suffix):

(8) Fixes are necessary

Putative base		Reduplication		Fixes
ket	→	*ket-ket*	→	*ke-ket*
kaprêk	→	*ka-prêk-prêk*	→	*ka-prê-prêk*

Fixes are not necessary

go	→	*go-go*
kyj	→	*kyj-kyj*
totyk	→	*to-tyk-tyk*

This approach is also interesting to deal with cases such as *tertet* or *prõrprõt*, where one part of the reduplicant ends in *-t* while the other ends in *-r* (*-r* is in all these cases followed by an epenthetic vowel: [tɛɾɛˈtɛt], [pɾɔ̃ɾɔ̃ˈpɾɔ̃t], [pɾɛɾɛˈpɾɛt], but this is due to a relatively superficial phonological rule). /*r*/ and /*t*/ alternate in certain contexts in Mẽbengokre, and it might be argued that in these cases there is full C(C)VC reduplication, with the coda C becoming /*r*/ or /*t*/ as an effect of a context-sensitive phonological rule.[7]

Though intuitively pleasing, the approach where reduplicants are uniformly C(C)VC and get pared down by the phonology is liable to a serious objection, which might have already been apparent in examining the second data set in (8). It is impossible to predict from the phonological environment just when a consonant will be dropped or not. Compare *kypkyp* with *tatak*. In both of these, C(C)VC reduplication would juxtapose two voiceless stops. In one case, the heterosyllabic stop sequence is fixed by dropping the first stop (*taktak* → *tatak*). This simplification is attested elsewhere in the language: the masculine personal name formative *Bep*, for instance, becomes *Be* in names such as *Bekĩnhre*. In the case of the reduplicated *kypkyp*, however, the simplification doesn't take place, and there's nothing

[7] As an aside, note that the alternation between /*r*/ and /*t*/ in reduplication seems to work in exactly the opposite way that it works elsewhere: other than in reduplicated forms, /*r*/ becomes /*t*/ in pre-obstruent position (cf. Xikrin *par* 'foot' + *kà* 'skin' = [patkʌ] 'shoe'). Word-finally, both /*r*/ and /*t*/ occur contrastively. These problems are outside of the purview of the present article; for more information, see Salanova (2001).

to explain this other than lexical idiosyncrasy.[8]

Regardless of the solution we choose, therefore, it has to be specified for each item of the lexicon that undergoes reduplication whether the reduplicant will be of the form C(C)V or C(C)VC, whether it's done by saying that there are two different affixes, or by a phonological rule that is applied to a lexically-conditioned subset of forms. This contrast between stems that get reduplicated with C(C)V and those that get reduplicated with C(C)VC could in principle form the basis of a subdivision of reduplicated forms into internally coherent classes,[9] but this is a question that will have to be the object of further research.

So we have a broad preliminary characterization of one type of reduplication found in Mẽbengokre: a C(C)V- or C(C)VC- derivational prefix appearing in verbs, vaguely denoting repeated action. Several objections hang over us: (1) this is not a fully productive process of reduplication, (2) we might not have been fully persuasive in arguing that it is not onomatopoeia, and (3) the semantics of "repeated action" is somewhat elusive.

These are valid objections that we will attempt to address (or at least heartfeltly bemoan) below. Before we may begin with that discussion, we need to put reduplication in the broader context of verbal morphology in Mẽbengokre. The next three sections, and particularly section 5, give us the necessary background for an analysis.

[8]It is true that the dropping of /*p*/ is also unsystematic elsewhere: next to the proper name *Bekĩnhre*, there is also the proper name *Bepkra*. But just as we would say to the child that tells a lie, the fact that other people do it doesn't make it right.

[9]For example, if we happened to find that there is a correlation between reduplicating with a C(C)VC template and greater productivity or semantic transparence, we could in good conscience talk about two separate reduplication rules, one of which might not be synchronically active.

3 The inflectional morphology of the Mẽbengokre verb

At first blush, Mẽbengokre verbal morphology seems extremely limited. A contrast between finite and non-finite forms exists in the core verbal lexicon; person marking is expressed by means of a prefix. Finiteness and person are the only properly inflectional categories of verbs. Derivational processes, to be dealt with in section 4, are limited to two morphological valency-changing operations: an anticausative and an antipassive with two allomorphs. Causation and applicativization are expressed through a strategy that seems to straddle morphology and syntax.[10]

The contrast between a finite and a non-finite form was used in Salanova (2007, 2008) as one of the main criteria to identify certain lexemes as being verbs, given that most lexical stems, whether nominal or verbal, may be used as matrix clause predicates without any other distinguishing morphological mark. It is an imperfect criterion, however, since many verbs are defective, including the majority of the verbs with reduplication, i.e., they do not have distinct finite and nonfinite forms. Finiteness can nevertheless be seen in the clausal syntax, even when the verbal morphology does not explicitly signal it: while alignment in clauses headed by non-finite verbs is ergative, (9), alignment in those headed by finite verbs is accusative, (10). Alignment in clauses with nominal predicates is always ergative-like, with transitive subjects (i.e., subjects of experiencer predicates, which are systematically nominal) being in some oblique case, while all other participants are marked by (absolutive) inflection on the verb, (11).[11]

(9) a. *ije* *a-mar*
1ERG 2-hear.N

[10] In the examples presented hereafter, we follow the practice of not segmenting the finiteness ending, as it is generally fused with the stem. The finite (or verbal) form of a verb is glossed V, and the non-finite or nominal form is glossed N. We give both the finite and the non-finite form of a verb when this is needed for clarity, as finiteness triggers allomorphy in several places of the word, often rendering the stems unrecognizable.

[11] For a discussion of where and for what purpose finite, non-finite and nominal predicates are used, see Salanova (2007, 2008). The translations of these examples are very approximative. For the purposes of the present paper finite and non-finite forms of a verb may be considered to be synonymous, or even contextually conditioned allomorphs.

'I have heard you.'

b. *ije* *mar*
1ERG 3.hear.N
'I have heard it/him/her.'

c. *i-rwỳk*
1-go.down.N
'I have gone down.'

(10) a. *ba* *a-ma*
1NOM 2-hear.V
'I heard you.'

b. *ba* *ku-ma*
1NOM 3AC-hear.V
'I heard it/him/her.'

c. *ba* *ruw*
1NOM go.down.V
'I went down.'

(11) a. *i-mã* *a-kĩnh*
1-DAT 2-please
'I like you.'

b. *i-mã* *kĩnh*
1-DAT 3.please
'I like it/him/her.'

c. *i-ngryk*
1-angry
'I'm angry.'

As can be seen in the above examples, the form that person inflection on the verb takes and the argument that it represents depend on the verb's finiteness. Finite verbs only have person inflection for the direct object (for this reason, we call this "accusative inflection"). Finite intransitive verbs don't get any inflection; their subjects are expressed by means of a free nominative pronoun. Non-finite verbs

always have person inflection for the absolutive argument, whether it's a direct object or the subject of an intransitive verb.

In addition, finiteness determines the form taken by the valence-changing affixes of the verb. This will be seen in the next section. It is for these reasons that we consider finiteness to be the most important morphological category on verbs, whether it's explicitly displayed or not.

Person inflection distinguishes first exclusive, first inclusive, second and third persons. Number marking is not part of the inflectional affix, though first person inclusive is necessarily at least two people, as opposed to the other persons which are singular if no explicit number marking is given. Third person is zero except when it marks direct objects of finite verbs. A peculiarity of third person accusative inflection is that it marginally agrees with the subject, in addition to marking the object, as can be seen in the following paradigm:

(12) *Agreement paradigm in finite transitive verbs*

ba ami-ma 1NOM REFL-hear.V 'I hear myself.'	ga i-ma 2NOM 1-hear.V 'You hear me.'	gu i-ma 12NOM 1-hear.V 'We (incl.) hear me.'	(ta) i-ma 3NOM 1-hear.V 'He hears me.'
ba a-ma 1NOM 2-hear.V 'I hear you.'	ga ami-ma 2NOM REFL-hear.V 'You hear yourself.'	gu a-ma 12NOM 2-hear.V 'We (incl.) hear you.'	(ta) a-ma 3NOM 2-hear.V 'He hears you.'
ba ba-ma 1NOM 12-hear.V 'I hear us (incl.).'	ga ba-ma 2NOM 12-hear.V 'You hear us (incl.).'	gu ami-ma 12NOM REFL-hear.V 'We (incl.) hear us.'	(ta) ba-ma 3NOM 12-hear.V 'He hears us (incl.).'
ba ku-ma 1NOM 3AC-hear.V 'I hear it.'	**ga a-ma** **2nom 2>3-hear.v** 'You hear it.'	gu ku-ma 12NOM 3AC-hear.V 'We (incl.) hear it.'	(ta) ku-ma 3NOM 3AC-hear.V 'S/he hears it.'

Accusative person inflection will be important again when we discuss the verbal "classifier" prefixes in section 5.

A possible question about person inflection is whether it should be considered a pronoun, agreement, or something else. This question is first formulated in Reis Silva (2001). For our purposes, we will say that person inflection is always referential,

which we take to be synonymous with it being "pronominal".[12] If one examines the third person, however, it is clear that it can't be deictic, like English *he* or *she*. Mẽbengokre person inflection is pronominal in the sense that its antecedent may be given by the common ground, rather than having to be syntactically present in a particular position of the clause, as would happen if it were agreement. For concreteness, the reader might think of Mẽbengokre person inflection as identical to the French object clitics *le*, *la*.[13]

Person inflection triggers allomorphy in certain stems. This allomorphy is of at least two types: with verb stems that begin in *ku*,[14] the absence of person inflection is accompanied by the dropping of the first vowel. Sometimes this might be the only morphological indication that person inflection is absent, given that third person is normally ∅:

(13) a. *krã̃ knõ̃*
head paint.red.V
"Paint his/her head read."

b. *kunõ̃*
3.paint.red.V
"Paint him red."

c. *i-kunõ̃*
1-paint.red.V
"Paint me red."

[12]The reason why this is a nontrivial assertion is that, while noun phrase objects are necessarily in complementary distribution with third person inflection unless topicalized, noun phrase subjects, including the nominative forms of pronouns, may co-occur with person inflection in non-finite intransitive clauses. Since nothing in this article hinges on this, we will just assume that when an overt subject co-occurs with person inflection, it's because the subject is topicalized.

[13]Here one could say Spanish, Italian, European Portuguese, etc., but with the caveat that some dialects of Romance (such as River Plate Spanish exemplified below) allow "clitic doubling", which makes these clitics somewhat different from what we have in mind:

Lo vi a tu hermano.
3AC.MASC see to your brother.
'I saw your brother (lit., I saw him, your brother).'

[14]This is not the *ku-* (3AC) that we examined above, but rather a stem formative which will be introduced below.

Another type of allomorphy found specifically with third person inflection is the dropping of certain stem-initial consonants:

(14) a. *i-jamỳ*
1-hug.V
'Hug me.'

b. *i-kra jamỳ*
1-child hug.V
'Hug my son/daughter.'

c. *amỳ*
3.hug.V
'Hug him/her.'

(15) a. *i-pyma*
1-fear
'Fear me.'

b. *bri pyma*
frog fear
'Fear frogs.'

c. *uma*
3.fear
'Fear him/her/it.'

Of course, it could be asked, given that these alternations are the only visible manifestation that a verb is inflected for third person, whether they shouldn't be considered a non-concatenative exponent of third person inflection. This, rather than being a real empirical question that these data can settle, seems to us to be about which theoretical device is least unseemly to a particular analyst. Those that disdain morphological zeroes might be happier with an approach where third person has several exponents, some concatenative and some processual, whereas those that only accept processual rules as a last resort might feel that there is sufficient reason to throw in some zeroes and compute the initial truncation as a morphophonological

rule. Data from related languages seem to lend support to the latter idea as a diachronic explanation (cf. Salanova 2004).

4 Derivational morphology

Derivational morphology in Mẽbengokre verbs includes an anticausative and an antipassive. Both of these are prefixes, and both have different allomorphs depending on whether the stem that they attach to is finite or non-finite. The following examples illustrate the allomorphy of the anticausative morpheme:

(16) a. *aj-kamẽ*
ANTICAUS-move.over.V
'He/she moved over.' (cf. kamẽ '(he/she) moved (it) over')

b. *bi-kamẽnh*
ANTICAUS-move.over.N
'He/she/someone has moved over.'
(cf. kamẽnh '(he/she/someone) has moved (it) over')

Anticausativized verbs function as intransitives, i.e., they inflect for the person of the subject when they are in the non-finite form, and they are invariable when they are finite. They can be retransitivized by means of the applicative clitic *o=*.

The following examples illustrate the allomorphy of the antipassive morpheme:

(17) a. *i-dju-mar*
1-ANTIPASS-hear.N
'I'm listening.' (cf. ba ku-ma 1NOM 3AC-hear.V 'I hear it.')

b. *i-djà-kur*
1-ANTIPASS-eat.PL.N
'I eat.' (cf. ba ku-ku 1NOM 3AC-eat.PL.V 'I eat it.')

c. *ba a-ku*
1NOM ANTIPASS-eat.PL.V
'I eat.'

The choice between *djà-* and *dju-* is lexically determined. *a-* is uniformly used when a finite verb form is antipassivized, but a certain number of antipassivized verbs are never found in the finite form (**a-ma*, which would be the finite form of *dju-mar*, for instance), for reasons that are unknown to us. Antipassivized verbs may be retransitivized by means of the applicative clitic *'ã=*.

As an example of retransitivization of anticausatives and antipassives, we offer the following, where the idiomatic meaning introduced by these valence-altering processes is evident:

(18) a. *ba mũm kẽn kamẽ*
1NOM thither stone move.V
'I move the stone thither.'

b. *ba mũm kẽn o=aj-kamẽ*
1NOM thither stone with=ANTICAUS-move.V
'I move the stone thither (I move with it).'

(19) a. *ba i-djàpêj kanga*
1NOM 1-work abandon
'I abandon my work.'

b. *ba i-djàpêj 'ã=i-dju-kanga*
1NOM 1-work at=1-ANTIPASS-abandon
'I'm prone to abandoning my work.'

The reader interested in these processes should consult Salanova (to appear) and Salanova and Sandalo (to appear).

Even though with this we've exhausted the list of what is clearly segmentable in Mẽbengokre verbal morphology, there are many other important generalizations to make about the shape of verbs.

5 Quasi-morphology

We begin by pointing out that all Mẽbengokre verbs, except for obviously complex fixed expressions, are either mono- or bisyllabic. Furthermore, the distribution of

monosyllables versus bisyllables is not random: while most underived intransitive verbs are monosyllables, the majority of transitives are bisyllables:

(20) a. *to* 'dance', *tẽ* 'go', *rê* 'swim', *ngre* 'sing', *bôx* 'arrive', *ty* 'die', *tỹm* 'fall',...

b. *jadjà* 'put on', *nhĩpêx* 'make', *kunõ* 'paint red', *kabi* 'taste', *pynê* 'catch',...

In fact, the pre-stressed syllable of transitive verbs is chosen among a small set of syllables: *ja-*, *ka-*, *ku-*,[15] *nhĩ-* and *py-*, with a handful of others being marginally present. This fact was observed in the earliest descriptions of Mẽbengokre verbs (Trapp n.d.). On the other hand, the final, stressed syllable of a transitive verb can be whatever the maximum syllable permitted in the language is, i.e., CCCVC, though the final C is normally dropped in the finite form.[16]

Let's call this template that describes the verb stem a sesquisyllable, borrowing a term that is commonly used in the description of some southeast asian languages (cf. Matisoff 1973: 86). We can make a preliminary generalization that underived intransitive verbs in Mẽbengokre are monosyllables, while underived transitive verbs are sesquisyllables.

Not so fast. Note that even in the examples in (10) above, we gave an example of a monosyllabic transitive verb. In fact, some of the most common transitive verbs are of this sort:[17]

(21)

bĩ	*bĩn*	'kill (sg.)'
pa	*par*	'kill (pl.), finish'
ma	*mar*	'hear'
pe	*pej*	'drink'
ga	*djàr*	'roast on coals'

What then of our generalization? To understand what's going on with transitive verbs, we need to reexamine some of the descriptive statements made in section 3.

[15] This *ku-* should not be confused with the third person marker *ku-* '3AC'; cf. fn. 14.

[16] For information on the syllable structure of Mẽbengokre, see Salanova (2001).

[17] As promised, we give both the finite and non-finite forms of these verbs, respectively. This is purely for completeness.

There we said that finite transitive verbs had a special marker *ku-* (3AC) for third person objects. The truth is that only some transitive verbs function like this:[18]

(22) a. *ba* *ku-ma*
1NOM 3AC-hear.V
'I hear it/him/her.'

b. *ije* *∅-mar*
1ERG 3-hear.N
'I have heard it/him/her.'

(23) a. *ba* *∅-kamũ*
1NOM 3-examine.V
'I examine it/him/her.'

b. *ije* *∅-kamũnh*
1ERG 3-examine.N
'I have examined it/him/her.'

That is, while in some transitive verbs a third person direct object is expressed by means of the morpheme *ku-* when they are finite, in some others the third person is consistently ∅. In the latter group, there is furthermore no agreement with the person of the subject (cf. ex. (12)). The verbs in the first group are precisely the transitive verbs that have monosyllabic stems.

In one analysis, suggested by Oliveira (2005), the reason why some verbs have *ku-* and others ∅ is to be found in the phonology: verb stems may be underlyingly mono- or bisyllabic, but verbs, at least those that are transitive, need to become bisyllabic on the surface; the *ku-* allomorph of the third person marker is chosen whenever an inflected stem ends up being less than bisyllabic, because of some word minimality constraint.

There are several problems with this approach.[19] One is that *ku-* is a different beast from ∅: it alternates with *a-* according to the person of the subject. ∅ never

[18]For clarity, we indicate third person as a zero in these examples, which are largely repeated from above.

[19]That this is an analysis which we ultimately discard does not take away merit from Oliveira's initial observation that all of these verbs have monosyllabic stems.

does that. In addition, *ku-* only appears on finite verbs; the same verbs, when in their non-finite form, inflect for third person in ∅.[20] Finally, *ku-* also appears to mark the object of certain postpositions; the set of postpositions that do this, as opposed to those that mark a third person object with ∅, cannot be characterized phonologically (though perhaps they can syntactically): *-bê* 'malefactive' (also used to mark the subject of equative predicates; cf. Salanova 2007:112), *-mã* 'dative', and *-te* 'ergative'.

Prepositions (like verbs) are known to govern particular cases in their complement noun phrases.[21] The picture that slowly emerges is one where *ku-* is not phonologically selected, but rather represents a distinct accusative person inflection, which contrasts with absolutive only in the third person. This is not unlike what happens, e.g., with the accusative vs. dative contrast in Romance clitics: while one has a contrast between accusative *le, la* and dative *lui* in the third person, the first and second person forms *me* and *te* are used to pronominalize both direct (accusative) and indirect (dative) objects.[22] Like in Romance, in Mẽbengokre there is no overt case marking on noun phrases, just on the pronominal prefixes.

However, there still is a link between the shape of a transitive verb and the case that it governs. We could state this as follows:

[20] A phonological way out of this would be to say that the non-finite forms of these verbs, though also monosyllabic, consist of a heavy syllable, while all the finite stems in question consist of a light syllable. This is an interesting possibility, though syllable weight seems to play no role elsewhere in the phonology of Mẽbengokre. Also, there is one exception to this generalization in the verb *kwỹnh*:

a. *ba* *ku-kwỹr*
1NOM 3AC-break.V
'I broke it.'

b. *ije* *∅-kwỹnh* *kêt*
1ERG 3ABS-break.N NEG
'I didn't break it.'

Here, whether third person inflection is with *ku-* or with ∅ depends solely on whether the verbal form is finite (V) or non-finite (N), as both stems consist of a heavy syllable.

[21] Cf., for example, German, where directional prepositions govern the accusative, while locative prepositions govern the dative.

[22] Again, we have used French for convenience.

(24) **Canonical form of Mẽbengokre verbs (provisional):** In Mẽbengokre,

a. underived intransitive verbs are monosyllabic;

b. transitive verbs are mostly bisyllabic;

c. all monosyllabic transitive verbs govern the accusative case, even though this is only seen in the form of the third person direct object prefix.

Verbs do have, after all, the sovereign right to govern whatever case they please. But we should still be curious about any link, even if indirect, between the phonological shape of the verb and the type of inflection that it gets.

We contend that bisyllabic transitive verbs are morphologically complex. Their direct objects are not objects of the verb root itself but of a prefix which is analogous to an incorporated adposition, hence the absolutive (the case governed by all but a few adpositions).[23] Most simplex (i.e. monosyllabic) verb roots are intransitive, but a handful are true transitives, and hence exhibit accusative person inflection.[24]

Is there any evidence for this? Though we have speculated before on a semantic basis for the distinction between transitives that assign accusative and those that assign absolutive (cf. Reis Silva and Salanova 2000), those speculations are not directly translatable into arguments for considering the latter to be made up of a root plus a transitivizer. We should therefore examine our evidence for the latter. Before we do so, however, we need to simultaneously give credit to and mark a difference with Oliveira's (2005:116-128) extensive discussion of this issue in Apinayé, which is to a great extent parallel to what we are contemplating here. Oliveira talks about morpho-semantic classes of verbs that share morphological formatives, i.e., repeated sequences which are not productive or straightforwardly segmentable

[23] I thank Maria Amélia Reis Silva for insightful discussion of this point.

[24] The reader might ask whether this means that transitive verbs are a closed class in Mẽbengokre. The answer is *yes*, but this might not be too telling as it might be the case that *all verbs* are a closed class, with verbal loanwords often entering the language as noun plus light verb constructions, or in any case as invariable words that mark all their objects through adpositions (i.e., with neither accusative nor absolutive).

morphemes. For Oliveira, these formatives have a possible diachronic origin in incorporated nouns.

The crucial difference between our approach and Oliveira's is that we state that, regardless of their diachronic origin, these prefixes are functioning synchronically as introducers of absolutive direct objects for a verb stem that is not inherently transitive. We don't go so far as to call them applicatives because there is no productive alternation between applicativized and non-applicativized verbs; rather, the prefixes are required by all transitive verbs.

So let's look for any evidence that bears on whether we can consider these formatives to be prefixes. The best type of evidence would be to find pairs of intransitive and transitive verbs that differ only by the presence of the formatives in the latter. Second best to finding an unprefixed base would be to find sets of verbs that share a base and differ in the prefix that they use. In addition, we would want to look for confirmation that these prefixes correlate with transitivity, and to seek hints as to any further semantic contribution they might make.

Of the 'short syllables' that make up the transitive verb's sesquisyllabic template (cf. the discussion under (20)), *ka-* is the best candidate for a transitivizer or marker of transitivity:

(25)	*tur*	'urinate'	*katur*	'urinate on'
	kwỳr	'defecate'	*kakwỳr*	'defecate on'
	djàkôr	'breathe'	*kakôr*	'blow on'
	djàkêj	'scratch (intr.)'	*kakêj*	'scratch (tr.)'

This prefix is likely a cognate to the postposition *kãm* 'inessive',[25] and can either be attached to simplex intransitive stems, or replace the initial 'short syllable' of an apparently complex intransitive.

[25]There is also a 'nominal applicative' *ka-* that forms relational (i.e., inalienably possessed) nouns out of unpossessed ones, as in *kangô* 'juice' (out of *ngô* 'water'), *kakô* 'wooden labret for the lip or earlobe' (out of *kô* 'polished wood'), etc.

For the last two examples, an unprefixed base is not attested. Recall that *djà-* was considered to be one of the allomorphs of the anitipassive in ex. (17). In fact, it seems that *djà-* is often in paradigmatic opposition to *ka-* and other prefixes found in transitive verbs, rather than being prefixed to a transitive stem, as we would expect of an antipassive morpheme. So perhaps what we have is a prefix *djà-* that marks intransitivity, just as *ka-* and others mark transitive verbs:

(26) a. *ropre nẽ akê*
dog NFUT scratch.V
'The dog scratches (at stuff).'

b. *ropre djàkêj kêt*
dog scratch.N NEG
'The dog doesn't scratch (at stuff).'

(27) a. *ropre nẽ pyka kakê*
dog NFUT earth scratch.V
'The dog scratches the earth.'

b. *ropre kute pyka kakêj kêt*
dog 3ERG earth scratch.N NEG
'The dog doesn't scratch the earth.'

(28) a. *ba i-djàpêj*
1NOM 1-work.N
'I work.'

b. *ba mỳja kupêj*
1NOM something work
'I work on something' (Port.: 'mexo com alguma coisa')

However, it is not always the case that the intransitive prefix is in paradigmatic opposition with one of the transitivizers, as next to (26) there are cases such as the following:

(29) a. *ba apnã*
1NOM stomp.V
'I stomp on the ground.'

b. *i-djàpnãr kêt*
1-stomp.N NEG
'I don't stomp on the ground.'

(30) a. *ba pry pynã*
1NOM path walk.V
'I walk the path.'

b. *ije pry pynãr kêt*
1ERG path walk.N NEG
'I don't walk the path.'

Here the antipassive prefix *djà-* is clearly outside of *py-*, which we identified above as one of the putative transitivity prefixes. The *-y-* is dropped as a result of a morphophonological process that is not exclusive to this context.[26] We will conclude that the prefix *djà-* is ambiguous between, on the one hand, an antipassive that attaches to stems that may already contain a transitivizing prefix, and, on the other, a marker of (unergative) intransitive verbs, in paradigmatic opposition to the transitivizing prefixes.[27,28] This will eventually necessitate that we reformulate our statement about the canonical form of verbs in Mẽbengokre.

As we mentioned above, something that sets the antipassive prefixes *djà-* and *dju-* apart from the transitive prefixes is that they interact with finiteness: whereas *djà-* and *dju-* are the allomorphs employed with non-finite forms of verbs, the form used with finite forms of verbs is *a-*.[29] This alternation, however, does not seem to distinguish between *djà-* or *dju-* used outside of a transitive prefix, such as in

[26] In the Xikrin dialect, this *-y-* is dropped also when the direct object is present; i.e., (30a) would be said *ba pry pnã*. This is possibly an instance of the process illustrated in (13), though the conditioning environments are somewhat different.

[27] If the reader would like to see us struggle with the ambiguity of the intransitive prefix *djà-* without reaching a satisfactory conclusion, we remit her or him to Salanova 2001, ch. 3.

[28] We will say nothing in this paper about the opposition between unergative and unaccusative intransitive verbs. What we have in mind is mostly semantic: unergatives are activities, while unaccusatives are changes of state or states. This generally matches up with the morphosyntax, but we make no claims about it here.

[29] This behavior is also displayed by the anticausative, which is *aj-* or *a-* with finite forms of verbs, and *bi-* with non-finite forms.

example (29), and a *djà-* or *dju-* used in paradigmatic opposition with a transitivity prefix:[30]

(31) a. *ba a-kô*
1NOM CL-$\sqrt{\text{BREATHE}}$.V
'I breathe/smoke.'

b. *i-djà-kôr*
1-CL-$\sqrt{\text{BREATHE}}$.N
'I breathe/smoke.'

(32) a. *ba karinhô ja-kô*
1NOM tobacco CL-$\sqrt{\text{BREATHE}}$.V
'I smoke tobacco.'

b. *ije karinhô ja-kôr*
1ERG tobacco CL-$\sqrt{\text{BREATHE}}$.N
'I smoke tobacco.'

To sum up this section and tie it in with the previous one we say the following: Mẽbengokre has at least one transitivity marking prefix *ka-* that attaches to monosyllabic roots, and at least one prefix *djà-* that marks unergative intransitives. The latter sometimes acts like it's in paradigmatic opposition to the transitivity prefix. We will gloss all of these prefixes as CL, for (verbal) classifier, which we believe to be a neutral term. Though the intransitive CLs are homophonous to antipassive morphemes, we consider the latter as a separate entity, as they appear outside of other CLs. Anticausative prefixes are clearly distinct from CLs, as they never appear in paradigmatic opposition with other CLs, but rather always outside of them.

6 More transitivity prefixes and synopsis

Though we haven't said it explicitly, we have already offered some examples where the transitivity prefix present in a stem is not *ka-* but some other one. We have

[30] In this example, we separate the transitive/intransitive prefixes from the verb root for the first time, and gloss them CL. Since the verb root can't be said in isolation in most cases, we add a $\sqrt{}$ to the gloss, indicating that it's a root (though not necessarily square).

even shown that those other prefixes can be replaced by *djà-*. In this section, we look at the prefixes *ja-*, *ku-*, *nhĩ-*, *py-* and the occasional other prefix found in pre-root position. We have two goals: first, to show that they are like *ka-*; second, to start speculating about what these prefixes might do in addition to indicating the transitivity of a verb, given that there are so many of them performing the same general function.

We established that *ka-* marks transitivity in at least a few verbs that have intransitive counterparts. There are examples that show that the other prefixes may fulfill the same function:[31]

(33)				
	ngrà	'dry' (adj.)	*kungrà*	'to wipe, scrub'
	djêk	'with scratches'	*kudjêk*	'to scratch' (tr. v.)
	tom	'with beeswax'	*kutom*	'to apply beeswax to the head'
	prãr	'to be empty handed'	*kuprãr*	'to empty out, raze'
	djàpêj	'to work'	*nhĩpêx*	'to make'
	djàkôr	'to breathe'	*jakôr*	'to blow at'
	djàptôr	'to spit'	*kutôr*	'to spit on'

Most of the examples of this are not easy to identify, as the meanings of the members of the pair differ in idiosyncratic ways, and one or both might appear only in fixed expressions. It is nevertheless clear that there are families of verbs with related meanings that share a root and differ in whether a prefix is present or not. In fact, several different prefixes, and not only the intransitive *djà-* exemplified above, may be in paradigmatic opposition with the same root. What one observes in these cases is that even though this prefix plausibly harks back to a more productive applicative, and is therefore primarily marking transitivity, its synchronic function is wider than this, as may be seen in the following examples:

[31] From now on, we will gloss over the morphophonological irregularities found in the data (cf., for instance, the /*p*/ in *djàptôr* below. These are very idiosyncratic, and we have very little to say about them. They might be revealing, but we haven't yet had the epiphany.

(34)	*djàpêj*	'to work'	*kupêx*	'to work on'
			nhĩpêx	'to make'
	nhãnh	'to bite'	*kamjãnh*	'to chew'
			djumjãnh	'to chew on grains'
	bjêr	'to drag along'	*jabjêr*	'to bring back game'
			djupjêr	'to carry on the back'
	pôt	'open'	*japôk*	'to make a hole' (tr. v.)
			nhõpôk	'to hollow out'
			kapôj	'to make an incision'
	(base unattested)		*kungrĩ*	'to arrange in a pile'
			kangrĩ	'to put together, pack up'
			jangrĩ	'to tie up a bundle'
			nhỹmngrĩ	'to cover a load with leaves'
	(base unattested)		*djàmjỳr*	'to poke the ground with a hoe'
			kamjỳr	'to poke with a needle'
			pymjỳr	'to stick into the ground'

We do not have the intimacy with the Mẽbengokre lexicon that is needed to understand this properly, but we may tentatively characterize the prefixes as encoding certain properties of the action, the object, and of how the latter is affected. Very broadly, we could say the following: *ku-* is associated with a single object, that is affected directly; *ja-* is associated with a more diffuse object, that is affected less directly; *ka-* is associated with an object that is touched on the surface; *djà-* is for intransitive verbs that denote an activity carried out with an implicit object; about *nhĩ-*, *py-*, and prefixless stems, we don't yet know what to say.[32]

We do wish to point out one semantic feature that is closely related to the transitivity-affectedness complex, however: verbal number. The distinction between

[32]Oliveira (2005) offers slightly different semantics to some of the formatives of Apinayé: *ka-* 'physical contact or manipulation', *ku-* 'movement or direct contact against a surface', etc. This discrepancy serves to illustrate the ellusiveness of this category's contribution to meaning.

pairs of morphologically related verbs often reflects the number of the object:

(35) a. *ku-ôj*
CL-$\sqrt{\text{BURN}}$.N
'light a fire (i.e., a bonfire)'

b. *ja-ôj*
CL-$\sqrt{\text{BURN}}$.N
'light many small fires (i.e., people's cigarettes)'

(36) a. *nhôr*
hang.N
'hang (i.e., a large piece of meat)'

b. *ja-nhôr*
CL-$\sqrt{\text{HANG}}$.N
'hang (i.e., beads on a string)'

Verbal plurality is so important for our analysis that it gets its own section, coming up ahead. Just to tie it in to the applicative function of the transitivity prefixes, we note that verbal plural is required whenever the object is plural, though it may also be used with a singular object to indicate a repeated action.

Summing up our picture of Mẽbengokre verbal morphology, we have the following:

(37)

①		②		③		④		⑤
Person	—	ANTIPASS/ANTICAUS	—	CL	—	Root	—	N/V

Verbal number does not have a slot of its own, but, as we said, is often encoded in position ③, or by suppletion of the whole stem (i.e., encompassing positions ②, ③, ④ and ⑤).[33,34] It should be clear by now that verbal number is not an inflectional

[33]Indeed, in some suppletive pairs, such as *nhỹr* 'to sit (sg.)' vs. *krĩn* 'to sit (pl.)', one of the members is an active verb, and thus has both a finite and a non-finite form, and the other is stative, having only a nominal or non-finite form. In addition, in some pairs one of the members is an intransitivized transitive stem, whereas the other member is basic: *rwỳk* 'to go down (sg.)' vs. *bixadjwỳr* 'to go down (pl.)' (= *bi-ja-djwỳr* ANTICAUS-CL-$\sqrt{\text{PUT.DOWN.PL}}$.N).

[34]For an analogy with this dissociation between a position-class analysis and the semantic features that make up the verbal word, think of Germanic particle verbs, such as English *eat up*, *screw*

category in Mẽbengokre, but rather an aspect of the semantics of stems that interacts in a particularly strong way with the syntax. To the extent that it can be encoded in the prefixes, we could consider it to be part of the derivational morphology. One should keep in mind, however, that the number of suppletive pairs of verbs that differ in number is possibly larger than the set of verbs that indicate number on the prefix, and that the vast majority of verbs don't display any contrast between singular and plural forms.

Let's look a little more closely at how verbal number works in Mẽbengokre.

7 Verbal number in Mẽbengokre

Consider the following:[35]

(38) a. *ngra nẽ ba saku kam djàr o mõ*
agouti NFUT 1NOM bag in put.SG.N with go.PL.V
'I was putting the agouti into the bag.'

b. *ngra nẽ ba saku kam ngjêx o mõ*
agouti NFUT 1NOM bag in put.PL.N with go.PL.V
i. 'I was putting the agoutis into the bag.'
ii. 'I kept on putting the (same) agouti into the bag.'

(39) a. *rop nẽ ba tak*
dog NFUT 1NOM strike
'I struck the dog.'

b. *rop nẽ ba titik*
dog NFUT 1NOM beat

around, *burn out*, etc. From a formal point of view, the 'particles' are easily identified as a distinct position class (even distinct words), yet it is not trivial to locate any specific grammatical meaning in them. Notions such as 'directionality', 'aspect', etc. are often expressed in that position, but may also be attached to the verb stem. Furthermore, like in the Mẽbengokre case that we are examining here, the meaning of the whole is seldom related to the meaning of the parts in a straighforward way.

[35]The plural auxiliary *mõ* in (38) should be ignored, as in these examples it simply signals that the action is in the progressive. For information about the Mẽbengokre progressive construction, please see Rivero et al. (to appear).

i. ‘I fought off the dogs.’

ii. ‘I thrashed the dog.’

These examples show the two typical uses of the verbal plural: to indicate the plurality of the object, and to indicate an action that is repeated. Example (38) has already been identified as a singular-plural pair in Trapp’s (n.d.) list. To our knowledge, it has never been argued that (39) could be instantiating the same opposition.[36]

Let’s set aside the (i) readings. Non-human noun phrases in Mẽbengokre do not indicate number, and verbal number in these examples might be making them be interpreted as plural in a roundabout way that we don’t care to find out about here. Rather, we want to face the task of showing that ‘to beat up, thrash’ could be considered to be the plural of ‘to strike’.

A little bit of light-hearted theoretical discussion is in order here. In the case of (38), the predicate could be said to denote a complex event which involves a change of state mediated by an activity. These event types are normally called “accomplishments” (cf. Vendler 1967, Dowty 1979, Smith 1997). Pluralizing such a predicate implies the pluralization of the whole complex eventuality. In the case of (39), on the other hand, the event is simpler: it is a punctual eventuality that does not have a before-state and an after-state except in an extralinguistic sense (the patient might be distracted before s/he is hit, and resentful afterwards, but this can’t be diagnosed linguistically). These event types are called “achievements” by Vendler, though Smith makes a further subclassification into “achievements” and “semelfactives”, and the example at hand seems to belong to the latter category. Pluralizing this type of event implies describing a multiplicity of punctual sub-eventualities. It is a fact of English that this plurality of punctual eventualities is often treated as a single eventuality. For instance, consider the following:

[36]For a discussion of some of the semantic complexities of verbal number, see Salanova (2007: 100 et ss.).

(40) a. When he entered the room, I sneezed (once).

b. I sneezed for an hour after I smelled your perfume.

Even though there are many individual instances of sneezing in the second case, no particular adjustments need to be made on the verb to accomodate this. For all intents and purposes, the predicate in (40b) is treated as a single durative (rather than punctual) eventuality. On the other hand, if the pluralized event is complex, one can normally not pull off something like this in English:

(41) a. When I saw him come, I put the agouti into the bag.

b. # While I was out hunting, I put the agouti into the bag for an hour.

c. I *kept putting* the agouti into the bag (and it kept coming out) for an hour.

While this is true in English, it seems not to be the case universally. According to Cusic (1981), who studied pluractionality in several native languages of North America, the pluralization of complex events, as in (41c), and the pluralization of punctual events that can be seen as forming a single complex eventuality, as in (40b), are both typically subsumed under pluractional (i.e., verbal plural) morphology.

In fact, there seems to be a continuum between these cases and cases where the verbal plural indicates plurality of the object: in *hit the dogs*, for instance, we could say that the verbal plural has scope over each subevent of hitting one dog, even if the whole is conceived as a single event.

We feel that we have good grounds to conclude that both (38) and (39) are instantiations of a single category of verbal plural, though with different scope in each case: a whole complex eventuality in one case, and a punctual subeventuality in the other.

The same conclusion may be reached for the following pairs, all of which denote repeated punctual eventualities:

(42)	*prêk*	'to startle, snap'	*kaprêprêk*	'to slap repeatedly'
	tãk	'to jump up'	*tãtãk*	'to beat (i.e., the heart), be sore'
	tak	'to strike'	*tatak*	'to clap, type'
	tok	'to startle'	*totok*	'to make repeated noise'
	mok	'to open'	*momok*	'to crack open repeatedly'

As we said above, 'repeated action' is part of the meaning of one set of reduplicated forms identified in (1). The problem with many of these is that we do not readily find a base out of which the reduplicant is formed, but we believe this to be an artifact of historical erosion:

(43)	*krakrak*	'to break up, munch grains'
	titik	'to beat repeatedly, type'
	pepek	'to droop repeatedly'

Yet another set for which we cannot readily identify a base seems to include the sense of plural of one of the participants, rather than necessarily of repeated events. In fact, several of these have stative meaning, and thus could not be iterative:

(44)	*rôrôk*	'to fall (pl.)'	
	ngrengrer	'to raze, destroy'	(possible base *ngrẽk* 'to stir')
	prõrprõt	'to float around'	(possible base *prõt* 'to run, flow, boil')
	krôkrôk	'to shake on branches'	

The limiting case are some examples that refer to noise-generating activities, and which are transparently onomatopoetic.[37] These bring us to an important point, to be made momentarily:

[37]Yet another oddball set are the following statives:

mjômjôp	'to itch, scratch'
rerek	'weak, soft'
prerpret	'to burn in the mouth (from being spicy)'

We don't know how to classify these.

(45)

'à'àk	'to shake a rattle'
gogo	'to make noise'
kangãngã	'to moan, complain, buzz'
kangêngê	'to wave a rattle'
karõrõ	'to snore, roar'
keket	'to laugh'
krikrit	'to make noise (peccary, snake)'
ngjengjek	'to buzz, squeak'
rãrãk	'to roar, thunder'
tertet	'to clatter'
xorxot	'to vibrate, oscillate'

Examples such as those in the last set raise the spectrum that the forms that we have been considering, rather than cases of true morphological reduplication, are paralinguistic ideophones or onomatopoeias. It would be truly far-fetched to argue that 'to laugh' or 'to tremble' is understood by Mẽbengokre speakers as a repeated semelfactive event; this would imply a granularity that seems completely alien to the way we perceive such events. Rather, it seems that *keket* imitates the sound of laughter, and *tertet* the sound of clattering teeth.

To avoid this serious objection, we need to draw a line between those cases where a base for the reduplicated form can be identified and those cases where it can't. Though this choice puts examples like *rôrôk*, *titik* and *pepek*, which all have an obvious plural sense, on the wrong side of the line, most of the plurals have a base, and most of the cases where a base can't be identified do not involve a repeated action in any generally acceptable sense. The base of reduplicated words such as *rôrôk* and *titik* would be, in our account, the victims of historical attrition, and might be found to exist in related languages. Words such as *keket* and *tertet*, on the other hand, possibly never had a base.

We have, after all these pages devoted to fascinating but apparently unrelated topics, identified the true reduplicated forms of Mẽbengokre. The reader may see that these forms, though arguably also iconic, are qualitatively different from superficially similar onomatopoeias.

Now we are ready for more. And it wouldn't be misleading to say that the best is yet to come.

The question to address is the following: what is the relation between the reduplicants that we have introduced as a manifestation of verbal number in this section, and the verbal template in (37)? Clearly reduplication should be identified with verbal number, about which we said that it could be all over the place. However, the reader might remember that we made a case for number being preferentially associated to the CL slot. We will now argue for that, and finish our analysis by comparing the reduplicated forms with the other prefixed verbal forms in the language.

8 Some evidence from ideophones

The term ideophone refers to a linguistic expression that evokes a sensory impression in an iconic way. A priori, this makes ideophones indistinguishable from onomatopoeia, though in common usage onomatopoeia refers to the general phenomenon of language that imitates the outside world, and one might thus find onomatopoeia in many places in a language, while ideophones are often grammaticalized into a coherent open class.

Ideophones belong to different grammatical categories in different languages. In Mẽbengokre, they are verbal main predicates. As far as following the morphological template of verbs given in (37), they are like some defective verbs in not distinguishing finite and non-finite forms. They do seem to require a prefix indicating (in)transitivity, but the prefixes that they employ are exclusive to them: *to-* if they are transitive verbs, or *e-* if they are intransitive verbs:[38]

[38] One could perhaps identify two emergent subclasses of ideophones in Mẽbengokre. While some

(46) a. *toyk*

'to slide (tr.)'

b. *eyk*

'to slip (intr.)'

(47) a. *toy*

'to drag along the floor (tr.)'

b. *ey*

'to drag oneself along the floor (intr.)'

(48) a. *totyk*

'to snap (tr.)'

b. *etyk*

'to snap (intr.)'

The relevance of ideophones to our discussion is that they often reduplicate.[39] When they do, the reduplicant prefix normally supplants the CL prefix, and the verb is transitive by default:

(49) a. *kyj*
scratch

'a scratch (N)'

are indistinguishable from other verbs except for the prefix *to-*/*e-*, others differ in that the root is clearly imitative, is prosodically separated from the prefix, contains sounds that are not part of the phonological inventory of the language, and is often, though not necessarily, followed by the verb *anẽ*, *anhỹr* 'to do thus; say'. The reader should assume that our examples, insofar as they may be identified as belonging to one or the other subclass, are of the first, "lexicalized", kind.

[39]This might be another criterion to distinguish the subclasses identified in footnote 38: only the "lexicalized" ideophones reduplicate in the way that we describe here. Imitative ideophones may show repetition, as in the following example, but with a clearly imitative intent; this can be seen both in the fact that the prefix is not omitted, that there are pauses between the repetitions, and that the ideophone root is repeated many times, rather than just once as occurs in reduplication:

õkôt bê to tuk tuk tuk anẽ
chest on CL IDEOPH IDEOPH IDEOPH do.so.V
'He (hit) him on the chest, tuk tuk tuk.'

b. *to-kyj*
CL-scratch
'to make a scratch on a surface (tr. V)'

c. *kyj-kyj*
RED-scratch
'to make many scratches on a surface (tr. V)'

(50) a. *to-krãk*
CL-$\sqrt{\text{SWALLOW}}$
'to swallow (tr. V)'

b. *krã-krãk*
RED-$\sqrt{\text{SWALLOW}}$
'to swallow repeatedly or lengthily (tr. V)'

We have given ad hoc glosses to the previous examples, but it should be clear that RED and CL are in paradigmatic opposition in these words. We contend most resolutely that that is equivalent to stating that the former is just a subtype of the latter.[40]

We would however not hide from the reader that the data that we have on ideophones is something that we chanced upon only in our last visit to the field. This means that it is still not as complete or well-understood as other data presented in this paper. It remains to be established to what extent the ideophones are lexicalized as verbs (see the two footnotes in this section), and whether the prefixes *to-* and *e-* of ideophonic verbs are truly comparable to the CL prefixes of other verbs. At this point, all we can say is that this is the default hypothesis.[41]

[40]The reader might ask about the form *totyktyk*, presented in (4). This, like *kaprêprêk*, also presented earlier, is one of a handful of cases where reduplication seems to occur concurrently with the CL prefixes. We don't know what to say about these at this point.

[41]Note that, like the CL prefix *ka-*, *to-* seems to be related to a postposition, in this case the comitative or instrumental postposition *o*. Like *ka-*, and unlike any true postposition, it is inseparable from the verb.

9 Reduplication as verbal number

We are now in a position to summarize our thoughts about verbal reduplication in Mẽbengokre.

Semantically, reduplication expresses iteration of an event or plurality of one of the participants. Formally, reduplication is a C(C)V- prefix attached to verbal roots. Like Marantz (1982) and much subsequent work in prosodic morphology, we mean this quite literally: the affix is a phonological skeleton, and its exact form is determined by *phonological* copying: there is no distinct *morphological* process of reduplication, only affixation. This is useful here as it allows us to identify the reduplicative morpheme with the other verbal prefixes: the reduplicant both occupies the same slot as *ka-*, *ja-*, etc., and has a similar CV shape. Furthermore, the other prefixes often also encode event iteration or participant plurality, though they don't do so exclusively, and are associated with transitive or unergative intransitive verbs, features that are also characteristic of reduplication.

Admittedly, all the examples that we have of a reduplicant morpheme being in paradigmatic opposition with another prefix all come from the ideophonic verbs. More examples of this sort would be the ultimate demonstration that reduplicants occupy the same slot as other prefixes. The reader should remember that it was also quite difficult to find examples to show that the prefixes were in fact in paradigmatic opposition (i.e., as in ex. (34)). Given the spotty character of our data, this might be an accidental lacuna.

To conclude, let us restate our generalization about the canonical form of verbs.

We had stated that transitive verbs were mostly sesquisyllabic, because they were made up of a root plus a prefix, and some basic transitives were monosyllabic but showed special properties; underived intransitives were monosyllabic, but this was because we considered that *djà-* was always an antipassive prefix. Toward the end of section 5 we concluded that *djà-* could also be a verbal prefix on par with

the prefixes found on transitive verbs. In addition, the reduplicated verb forms turn out to have the same shape as all other sesquisyllables. Thus:

(51) **Canonical form of Mẽbengokre verbs (definitive):**

a. Mẽbengokre verbs are either simple roots (monosyllables) or roots plus a CL prefix (sesquisyllables).

b. The CL prefixes are:

i. *ja-*, *ka-*, *ku-*, *nhĩ-*, *py-*, *to-*, *C(C)V-* and *C(C)VC-* for transitives;

ii. *djà-*, *dju-*, *e-*, *C(C)V-* and *C(C)VC-* for unergative intransitives.

iii. Meanwhile, *to-* and *e-* are used exclusively with ideophonic verbs.

c. The CL prefixes are associated with the function marking transitivity, number and object properties on verbs.

d. Unprefixed verbs include unaccusative intransitives and true transitives, which can be identified by the use of a special form of the third person object marker, *ku-* (3AC).

There are very few exceptions to this generalization. Two that we can think of are *tak* 'hit' and *'ôk* 'paint', which are transitive but bear no prefix, and inflect for third person with ∅ rather than *ku-*. There are also a few other possible CL prefixes that appear on a small number of transitive verbs. We didn't list these above because they don't appear on enough verbs for us to make any meaningful claims, though two of these examples were already given in (34):

(52) a. *dju-pjêr*
CL-$\sqrt{\text{CARRY}}$.N
'carry on the shoulders'

b. *djwa-pre*
CL-$\sqrt{\text{TIE}}$.N
'tie someone down'

(53) a. *nhỹm-ngrĩnh*
CL-$\sqrt{\text{BUNDLE}}$.N
'cover a load with leaves'

b. *nhỹm-kjêr*
CL-$\sqrt{\text{SPREAD.APART}}$.N
'look for something all over the place'

In the first example we have the CL prefix *dju-*, normally used with unergative intransitives, being used with a transitive verb. The prefix *djwa-* might be related to transitive *dju-*, or might have a completely separate origin. In the second set of examples we have the prefix *nhỹm-* occupying the spot of the CL prefixes.

These isolated facts should serve as leads into newer inquiries. As for us, we wish to end our paper with a comparison of the Mẽbengokre facts to those found in another Jê language, Kaingang, as we believe that the function of the CL prefixes of Mẽbengokre will only be fully understood when examined comparatively.[42]

10 Reduplication and number in Kaingang

Cavalcante (1987: 58 et ss.) describes the formation of verbal plurals in Kaingang. This morphological category is claimed to indicate the number of the absolutive argument (cf. also Urban 1985), though D'Angelis (2004) shows that it may also indicate repeated action with singular participants. The majority of verbs in Kaingang do not contrast a singular and a plural form. Of the approximately 200 that do in Cavalcante's data, just over half do so by reduplication. In all the Kaingang examples, the first column gives the singular form of the verb, while the second column has the plural:[43]

[42]This is not to say that a lot couldn't be concluded from looking at a single language synchronically, even if the morphology is not very productive or regular. As an example of what can be achieved in this way, one only needs to look at the vast generative literature on the aspectual prefixes of the Slavic languages (for a recent examination of a small subset of these prefixes, see Žaučer 2009). It will be impossible to achieve such depth of understanding in the Jê family as long as the descriptions of the lexicon remain as superficial as they are at present.

[43]We have converted Cavalcante's transcriptions to the orthography for simplicity, as nothing

(54)	*kõm*	*kõmkõm*	'to dig'
	mrãn	*mrãnmrãn*	'to wet'
	nẽm	*nẽmnẽm*	'to cut with scissors'
	gàg	*gàggàg*	'to roast'
	gunh	*gunhgunh*	'to stick into the ground'
	jành	*jànhjành*	'to urinate'
(55)	*kanhin*	*kanhinnhin*	'to play'
	kãmun	*kãmunmun*	'to measure'
	jẽmĩ	*jẽmĩmĩ*	'to feel with the hands'
	wĩrĩn	*wĩrĩrĩn*	'to turn around, to surround'
(56)	*gren*	*gringren*	'to dance'
	mranh	*mrynhmranh*	'to break'
(57)	*kar*	*kankar*	'ready'
	mrô	*mrôgmrô*	'to bathe'
	nã	*nãgnã*	'to lie down'
	nãn	*nãgnãn*	'to float'
	nê	*nêgnê*	'to hide in the earth'
	nĩ	*nĩgnĩ*	'to sit down'
	non	*nugnon*	'to open'
	nũr	*nũgnũr*	'to sleep'

The basic pattern is of a CVC reduplicant that is either prefixed (cf. 54) or infixed before the stressed syllable of the base (cf. 55). In some cases (cf. 56) there is vowel raising on the reduplicant, something which incidentally serves to show that the process in question is in fact prefixation rather than suffixation. Yet another variant of the reduplicant has fixed segmentism, either replacing a segment of the

hinges on the phonological detail. The letters *m*, *n*, *nh* and *g* represent partly or fully nasalized voiced stops (i.e., *g* → [ŋ, ᵑŋ, ŋᵍ, ᵑŋᵍ], according to context); other symbols have approximately the same value as in Mẽbengokre.

base (as in *kankar*, *nugnon*) or filling in in cases where the base does not have a coda (as in *nĩgnĩ*, *mrôgmrô*).

Reduplication sits side by side with other morphological devices as a marker of the verbal plural. With monosyllabic stems, the most common device other than reduplication is prefixation, of either a CV or a CVC prefix (cf. 58). With stems that are already bisyllabic, one finds various forms of substitution of the first syllable, with the most common being the augmentation of a CV first syllable to CVC, by the infixation of *g* (cf. 59):

(58)	*fã*	*kygfã*	'to cry'
	rũm	*yogrũm*	'to shake, move'
	grĩn	*kugrĩn*, *tugrĩn*	'to wind up'
	tu	*pãgtu*	'to carry (long object)'
	sĩn	*kãsĩn*	'to make small'
	ter	*kãgten*	'to die'
(59)	*kafãn*	*kygfãn*	'to husk corn'
	fãnãn	*fãgnãn*	'to spend almost all'
	kagje	*kygje*	'to tie a knot'
	peju	*pigju*	'to steal'
	pafãm	*pygfãm*	'to nurse'
	kãgmĩ	*kugmĩ*	'to grab, hold'
	tãnfyn	*jãnfyn*	'to raise'

To sum up, reduplication and prefixation together account for the expression of plural in all but 10 of the monosyllabic stems that mark it; prefix "substitution" accounts for the expression of the plural in three quarters of bisyllabic stems that mark it, while most of the rest reduplicate and become trisyllabic. Prefixes that indicate the plural come from a closed set (in Cavalcante's analysis, *ky-*, *kã-*, *ko-*, *ku-*, *pã-*, *tu-* and *jo-*, but the elements in this set may also be augmented by a *-g-*).

This is as far into Kaingang as we need to go for our present purposes. This

quick incursion into Kaingang only intends to show that it would be plausible to advance the hypothesis that "reiterative" reduplication in Mẽbengokre verbs is a reflex of the process that may be seen, perhaps more pervasively or robustly, in the verbal lexicon of Kaingang. Like in Mẽbengokre, verbs in Kaingang are either monosyllabic or sesquisyllabic. Also like in Mẽbengokre, but perhaps even more clearly and productively, number affects the short syllable, either by replacing it, or by adding one where there is none.[44] In both languages, the morpheme that is added may or may not be reduplicative. We have not attempted a systematic comparison in search of cognates or more quirky structural similarities, but we encourage our readers to probe futher in that direction.

11 Conclusions

Even though the data on which we can base our conclusions is incomplete, we hope to have shown the reader that there is a well delimited subclass of reduplicative phenomena in Mẽbengokre that fits in with the morphological template of the Mẽbengokre verb. We also hope to have shown more generally that Mẽbengokre verbs have structure beyond what is easily segmentable.

Further in support of this analysis, we noted that it makes number marking in Mẽbengokre and Kaingang virtually identical: both are expressed through prefixation, reduplication, and substitution of affixes.

A research program stems from our overview, which we hope will be taken up by other researchers working in the language family: namely, to understand the interesting morphological sub-regularities found in the verbal lexicon of the Jê languages, and to investigate the subtleties in verbal meaning that lurk under the deceivingly poor morphology of these languages.

[44] Kaingang also has a small set of verbs where the number opposition is marked by suppletion of the stem.

References

Alves, Flávia de Castro. 1999. Aspectos fonológicos do Apãnjekra (Jê). Master's thesis, Universidade de São Paulo.

Burgess, Eunice, and Patricia Ham. 1968. Multilevel conditioning of phoneme variants in Apinayé. *Linguistics* 41:5–18.

Callow, John. 1962. The Apinayé language. Doctoral Dissertation, London University.

Cavalcante, Marita Pôrto. 1987. Fonologia e morfologia da língua Kaingang: o dialeto de São Paulo comparado com o do Paraná. Doctoral Dissertation, Universidade Estadual de Campinas.

Cusic, David Dowell. 1981. Verbal plurality and aspect. Doctoral Dissertation, Stanford University.

D'Angelis, Wilmar da Rocha. 2004. Concordância verbal de número em Kaingáng: algumas pistas. *LIAMES* 4:71–81.

Dourado, Luciana. 2001. Aspectos morfossintáticos da língua Panará (Jê). Doctoral Dissertation, Universidade Estadual de Campinas, Campinas.

Dowty, David. 1979. *Word meaning and Montague grammar*. Dordrecht: Kluwer.

Marantz, Alec. 1982. Re: Reduplication. *Linguistic Inquiry* 13:435–482.

Matisoff, James. 1973. Tonogenesis in Southeast Asia. In *Consonant types and tone*, Southern California Occasional Papers in Linguistics, 71–95. Linguistics department, University of Southern California.

Oliveira, Christiane Cunha de. 2005. The language of the Apinajé people of Central Brazil. Doctoral Dissertation, University of Oregon.

Popjes, Jack, and Jo Popjes. 1986. Canela-Krahô. In *Handbook of Amazonian languages*, ed. D. C. Deryshire and G. Pullum, volume 1, 128–199. Berlin: Mounton de Gruyter.

Reis Silva, Maria Amélia. 2001. Pronomes, ordem e ergatividade em Mẽbengokre. Master's thesis, Universidade Estadual de Campinas.

Reis Silva, Maria Amélia, and Andrés Pablo Salanova. 2000. Verbo y ergatividad escindida en Mẽbengokre. In *Indigenous languages of lowland South America*, ed. Hein van der Voort and Simon van de Kerke. Leiden, Netherlands: Research School of Asian, African, and Amerindian Studies (CNWS), Leiden University.

Rivero, María Luisa, Ana Arregui, and Andrés Pablo Salanova. to appear. The syntax and semantics of viewpoint aspect in Mẽbengokre. In *Proceedings of the 13th Workshop on the Structure and Constituency of the Languages of the Americas*. UBCWPL.

Salanova, Andrés Pablo. 2001. A nasalidade em Mẽbengokre e Apinayé: o limite do vozeamento soante. Master's thesis, Universidade Estadual de Campinas.

Salanova, Andrés Pablo. 2004. Subtractive truncation in Mẽbengokre. MIT ms.; updated version is under review for *Anthropological Linguistics*.

Salanova, Andrés Pablo. 2007. Nominalizations and aspect. Doctoral Dissertation, Massachusetts Institute of Technology.

Salanova, Andrés Pablo. 2008. Uma análise unificada das construções ergativas do Mẽbengokre. *Amérindia* 32:109–134.

Salanova, Andrés Pablo. 2009. Não existem prefixos relacionais nas línguas Jê. In *Línguas e culturas Macro-Jê*, ed. S. Martins Filho and S. Braggio. Goiânia: Universidade Federal de Goiás.

Salanova, Andrés Pablo. to appear. Semántica causativa, sintaxis aplicativa. In *Actas del coloquio Amazónicas III*. Bogotá: Universidad Nacional.

Salanova, Andrés Pablo, and Filomena Sandalo. to appear. Antipassive. In *Proceedings of the 13th Workshop on the Structure and Constituency of the Languages of the Americas*. UBCWPL.

Santos, Ludoviko dos. 1997. Descrição da morfossintaxe da língua Suyá/Kĩsêdjê (Jê). Doctoral Dissertation, Universidade Federal de Santa Catarina, Florianópolis.

Smith, Carlota. 1997. *The parameter of aspect*. Dordrecht: Kluwer, 2nd edition.

Trapp, Earl. n.d. List of Kayapó verbs. Unpublished ms.

Urban, Greg. 1985. Ergativity and accusativity in Shokleng (Gê). *International Journal of American Linguistics* 51:164–87.

Vendler, Zeno. 1967. *Linguistics in philosophy*. Ithaca, NY: Cornell University Press.

Wiesemann, Ursula. 1972. *Die phonologische und grammatische Struktur der Kaingang-Sprache*. Janua Linguarum, Series Practica. The Hague: Mouton.

Žaučer, Rok. 2009. A VP-internal/resultative analysis of 4 "VP-external" uses of Slavic verbal prefixes. Doctoral Dissertation, University of Ottawa.