The sense of Mẽbengokre nominalizations[*]

Andrés Pablo Salanova
MIT

This paper aims to provide a unified structure and interpretation for three constructions of Mẽbengokre. The constructions are "short" (arguably lexical) nominalizations, internally-headed relative clauses, and eventive complement clauses. The analysis relies on allowing determinerless noun phrases to compose directly with the predicate of which they are arguments, by means of the rule of predicate restriction (Chung and Ladusaw 2004), and on assuming that the determiners that exist in the language are unselective, and can bind both the individual-type variables introduced by the verb's arguments and the event-type variable that saturates the verb's event argument.

1 Introduction

The focus of this study are three constructions found in Mẽbengokre:[1] "short" nominalizations, exemplified in (1a), internally-headed relative clauses, as in (1b), and eventive complement clauses, shown in (1c):[2]

(1) a. piʔok-jarẽ-ɲ-dʒwʏj
writing-say-NLZR-master
'teacher'

b. mẽ kutɛ kum dʒɔrɛ jarẽ-ɲ ja
PL 3ERG 3DAT Djore say-NLZR the
'the ones that are called Djore'

[*] I would like to thank my consultants, Nory Kayapó, Bep Kamrêk Kayapó and Ikrô Kayapó, for teaching me what I know of Mẽbengokre, and Maria Amélia Reis Silva, Lisa Matthewson, and the audience at a UBC research seminar, who made valuable comments on a previous version of the present paper.

[1] Mẽbengokre is a Jê language spoken in the eastern Brazilian Amazon region by approximately 7000 people belonging to two nations, Xikrin and Kayapó. For some groundwork and references see Reis Silva and Salanova (2000) and Reis Silva (2001). The data presented here comes from original fieldwork by the author.

[2] Abbreviations used are the following: 1,2,3 – person marks; ERG – ergative; NOM – nominative; ACC – accusative; DAT – dative; POSS – possessive (with alienably possessed nouns); absolutive/genitive person marks aren't labelled for case. NFUT – non-future; FUT – future; NLZR – nominalizer; PL – plural; AP – antipassive. Morphemes that are segmentable are separated with a hyphen; those that are fused are separated by a dot.

c. ba ʌk kʌ-r ma
1NOM bird coo-NLZR hear
'I heard the bird cooing.'

These constructions have much in common, and contrast clearly with perfective[3] main clauses employing the same verbs. The points of contrast are (a) the verb form used, which displays nominalizing morphology in the constructions in question; (b) an ergative pattern, evident in pronominal agreement on the verb and in the marking of transitive subjects, and (c) the lack of higher functional structure that is present in main clauses. The first trait can be seen clearly in (1); the latter are exemplified below:

(2) Ergative split

a. Nominative pattern in main clauses:

i. ba a-pumũ
1NOM 2-see
'I see you.'

ii. ba nõ
1NOM lie
'I lie down.'

b. Ergative pattern in the constructions in question:[4]

i. ijɛ a-pumũ-ɲ
1ERG 2-see-NLZR
'I don't see you.'

ii. ɨ-nõ-r
1-lie-NLZR
'I don't lie down.'

(3) Lack of higher functional structure:

a. Focus, tense/mood, and a higher subject position in main clauses:

kukrɯt nẽ ba arɤm ku-bĩ
tapir (FOC) NFUT 1NOM already 3ACC-kill
'I killed *tapir*.'

b. Not available in the constructions in question:

(*kukrɯt) (*nẽ) (*ijɛ) arɤm ijɛ bĩ-n
tapir (FOC) NFUT 1ERG already 1ERG 3.kill-NLZR

In this paper we will propose an analysis where all the constructions in (1) share essentially the same underlying structure, and various interpretations are picked out by determiners with different properties. The approach is initially motivated as an extension of what is required to deal with the ambiguity of Mẽbengokre relative clauses. We therefore turn to this construction first.

---

[3] Non-perfective tenses often involve complex constructions, and for this reason are excluded from the characterization of canonical main clauses.

[4] Of relevance for the discussion in §5 is the fact that, in ergative constructions, the external argument can be dropped; in such situations the agent is interpreted "generically", as in the English passive.

2 Internally-headed relative clauses

In restrictive relative clauses in Mẽbengokre the head, i.e., the noun that is modified by the relative clause, appears inside it, in the position it would occupy as an argument. No special marking appears on the head noun phrase. As in other languages described in the literature (e.g., Lakhota; cf. Williamson 1987), any indefinite noun phrase inside the relative clause can be taken to be the head; this results in ambiguity when there is more than one indefinite noun phrase in the clause:[5]

(4) kubẽ kutɛ mẽ i-mʌ̃ mẽkrĩdʒʌ ɲõ-r nẽ jã
barbarian 3ERG PL 1-DAT chair give-NLZR NFUT this
'These are the chairs that a/the white man gave us', or
'This is the white man that gave us some/the chairs.'

There is no restriction as to the grammatical function within the relative clause of the noun phrase that serves as head: noun phrases in adjunct roles are freely allowed, as attested in (5). Null third-person pronominals can also be interpreted as heads, in a construction that could be considered the equivalent of a free relative (cf. 6):

(5) kubẽ kot i-tẽ-m nẽ ijɛ aŋro bĩ-n nẽ jã
barbarian with 1-go-NLZR and 1ERG peccary kill-NLZR NFUT this
'This is the white man with whom I went and killed peccaries.'

(6) mẽ tũm kutɛ arẽ-ɲ nẽ jã
PL old 3ERG 3.say-NLZR NFUT this
'This is what the ancients told.'

Note, though, that the relative clause cannot be headed by an adjunct that is not overtly present:

(7) a. kot i-tẽ-m ja
with 1-go-NLZR the
'the one who I went with'

b. * i-tẽ-m ja
(impossible with the same translation)

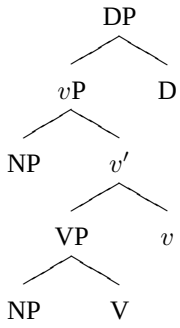
We take this to mean that the heads of relative clauses in Mẽbengokre can only be noun phrases, i.e., there are no relative clauses headed by *how*, *when*, etc. If a head (normally either the main predicate or a postposition) governing a (potentially null) noun phrase participant is not present, then that participant cannot head the relative clause.

[5] Note that there are no overt indefinite determiners in Mẽbengokre. Definite determiners *ja* (proximal) and *wa* (distal) are optional, and are just destressed (and consequently denasalized) versions of the demonstratives *jã* 'this' and *wã* 'that'.

3 The basic semantic analysis

We will assume that internally-headed relative clauses such as (1b) consist of the following structure:

(8)



That is: (a) the relative clause itself consists of a verb and all of its adjuncts and arguments,[6] but without projections in the TP or CP domain;[7] this accords with the facts presented in (3); (b) this structure is selected directly by a determiner, and (c) arguments of V (and of v) can be plain NPs, though we won't exclude the possibility that they be DPs. The sense of the latter two assumptions will be made clear below.

Let us work our way through one example. Consider the following:

(9) kubẽ kutɛ rɔpkrɔri bĩ-n ja
barbarian 3ERG jaguar kill-NLZR the

a. 'The white man who killed the/a jaguar.'
b. 'The jaguar that the/a white man killed.'

The following denotations are straightforward:[8]

(10) a. $[\![\text{bĩ}]\!] = \lambda x_e.\lambda e_l.\text{kill}'(e, x)$
b. $[\![\text{rɔpkrɔri}]\!] = \lambda x_e.\text{jaguar}'(x)$
c. $[\![\text{kubẽ}]\!] = \lambda x_e.\text{barbarian}'(x)$

We need to show how these parts come together to give the correct denotation to the subconstituent *rɔpkrɔri bĩ*. Clearly they cannot combine by Functional

[6] Following Kratzer (1996), we take the external argument to be introduced by a Voice head, which we will call v here for purely typographical reasons.

[7] In this paper, we will skirt completely the issue of the temporal and aspectual interpretation of such clauses. For discussion, see Salanova (in progress).

[8] The semantic types used are: individuals (e), eventualities (l), which can be considered just a special type of the former, and truth-values (t). For other particulars on the semantic notation used here, see Heim and Kratzer (1998).

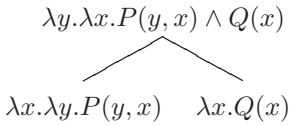
Application unless something intervenes to turn one of the terms into a type that can be argument of the other term.

Determiners are usually thought to be the functions that, taking noun denotations as arguments, yield a constituent of the appropriate type to combine with a verb: $\langle et, e\rangle$ (i.e., *the*) or $\langle et, \langle et, t\rangle\rangle$ (quantifiers).

Heim (1982) has argued that indefinite noun phrases lack quantificational force of their own; indefinite determiners are thus either of the former type, or semantically vacuous. Simplistically, the first alternative corresponds to the choice-functional approach to indefinites (see Matthewson 1999). The alternative we favor here pays heed to the fact that indefinite noun phrases in Mẽbengokre have no overt determiners: we will consider noun phrases in this language to be determinerless NPs, of type $\langle e, t\rangle$.

How do these determinerless NPs compose with main predicates that are also of type $\langle e, t\rangle$? We will follow Chung and Ladusaw (2004) and introduce a new compositional rule of Predicate Restriction:

(11) Predicate Restriction (op. cit., p. 5)



If we apply this composition rule to our example, we get $[\![$rɔpkrɔri bĩ$]\!] =$ $\lambda e.\lambda x.$kill$'(e, x) \wedge$ jaguar$'(x)$.

The head introducing the external argument is composed into the structure by another rule, Event Identification (EI),[9] proposed by Kratzer (1996). This rule is also at work for introducing all other non-core arguments and adjuncts. We do not need to go into this topic in any detail.

We will assume that *kutɛ* is vacuous. So the denotation we get for the core of the relative clause (9) is:[10]

(12) $[\![$kubẽ kutɛ rɔpkrɔri bĩ-n$]\!] =$

$\lambda y.\lambda e.\lambda x.$kill$'(e, x) \wedge$ jaguar$'(x) \wedge$ barbarian$'(y) \wedge$ Agent(e, y)

[9] In reality, we need to extend Kratzer's EI rule somewhat, since we are adding external arguments to unsaturated predicates, but this extension is trivial.

[10] Note that we have surreptitiously introduced the nominalizer *-n*, which arguably sits under v. For the purposes of this paper we will assume that nominalizing morphology is semantically vacuous. The absence of nominalizing morphology when the construction is interpreted as a main clause, forces an eventive interpretation in which all variables of type e are existentially closed. We cannot devote much space to the workings of existential closure. It is widely assumed that the free variables of indefinite noun phrases have to be closed (i.e. bound) at some point in the tree, and we can assume (with precedents in Chung and Ladusaw 2004) that this point is at the event level, once all the arguments of the main predicate have been merged. See Salanova (in progress) for more discussion.

What we are claiming, counterintuitively, is that whenever the syntactic arguments are indefinite (i.e., determinerless) noun phrases, a verbal projection is as unsaturated semantically as just a verb by itself. The denotation of such vPs is an n-place property of individuals. This vP combines with T to form matrix clauses (a topic which we won't discuss here; cf. fn. 10) and with D to form internally-headed relative clauses.

What is the determiner *ja*, then? We will claim that it is an unselective binder. This means that it binds a variable contained in its sister constituent, but which variable is bound (if the constituent contains more than one) is not determined by structure. Any one variable is bound by *ja*, while all other variables that are free at this point in the structure are bound by existential closure:[11]

(13) $[\![\text{ja}]\!] = \lambda P_{et}.\iota x P(x)$[12]

(14) kubẽ kutɛ rɔpkrɔri bĩ-n ja
barbarian 3ERG jaguar kill-NLZR the

a. 'the jaguar that a white man killed'
$\iota x \exists e \exists y$: kill$'(e, x) \wedge$ jaguar$'(x) \wedge$ barbarian$'(y) \wedge$ Agent(e, y)

b. 'the white man that killed a jaguar'
$\iota y \exists e \exists x$: kill$'(e, x) \wedge$ jaguar$'(x) \wedge$ barbarian$'(y) \wedge$ Agent(e, y)

This is pretty straightforward. Note that in addition to these readings, there are readings of this sentence where the non-head noun phrase is definite, since, as we mentioned in footnote 5, overt definite determiners are optional.

4 Eventive readings

We mentioned in the introduction that complement clauses that were formally identical to internally-headed relative clauses could get eventive interpretations. This happens at least when these clauses are used in direct perception constructions (15a), and when they are complements of manner predicates (15b):

(15) a. ba ʌk kʌ-r ma
1NOM fowl coo-NLZR hear
'I heard the bird calling.'

b. a-dʒu-jarẽ-ɲ mɛj
2-AP-say-NLZR good
'You spoke well.' (lit.: 'Your saying was good.')

[11] Existential closure and the binding by the unselective definite operator happen at the same point in the structure, and are dependent on each other. For simplicity, the denotation of *ja* is given as if it were combining with a predicate where all but one of the variables have been existentially closed.

[12] Of course, it's an empirical question whether the semantics of this phonetically null determiner in Mẽbengokre really is that of the English definite article.

In fact, the semantics that we have developed for relative clauses extends without significant modification to get the senses in (15). Remember that an event variable is free in the denotation of a nominalized clause, just as there are individual variables for each of the arguments. If it is this variable that gets bound by the unselective binder, the reading we get is eventive:[13]

(16) ba bẽnjadʒwɤrɤrɛ kutɛ bẽn dʒi-r ma
1NOM chief 3ERG speech place-NLZR hear

a. '[I heard] a chief reciting a ritual speech'
$\iota e.\exists x.\exists y.\text{recite}'(e,y,x) \wedge \text{chief}'(y) \wedge \text{speech}'(x)$

b. '[I heard] the ritual speech that a chief recited'
$\iota x.\exists e.\exists y.\text{recite}'(e,y,x) \wedge \text{chief}'(y) \wedge \text{speech}'(x)$

It seems that the determiners *ja* and *wa*, can only bind variables of type e, while the phonetically empty definite determiner can bind variables of type e and of type l (eventuality).[14] Constructions with the former are unambiguously interpreted as internally-headed relative clauses, while constructions with the latter are the only ones ambiguous between IHRC and eventive readings.

5 Short nominalizations

Mẽbengokre has two morphemes, *dʒʌ* and *dʒwɤj*, that are used to create a large repertoire of what could be intuitively called "lexical" nominalizations, such as the one given in (1a), and the following:

(17) i-dʒʌ-ku-r-dʒʌ
1-AP-eat-NLZR-container
'My eating utensils', but also: 'my eating place', 'my food', etc.

In the literature on other Jê languages (cf., e.g., de Oliveira 2005), these have been considered to be an instrument and an agent nominalizer, respectively. Our contention is that what the "nominalizers" attach to is already nominal (i.e., it's an eventive complement clause, as described above), and they themselves are no more than the semantically bleached nouns *dʒʌ* 'container' and *dʒwɤj* 'master'.

What is the relation between these nouns and the nominalized clause? The nouns cannot be external heads: as we saw above, what is interpreted as the head of an internally-headed relative clause has to be a null pronoun or a determinerless noun phrase in a governed position. This is not the case in either

[13] It's fair to ask what this means, but attempting to answer this would take us very far afield. For our present purposes, note that (15b) cannot mean 'It was good that you spoke', and more generally such complement clauses can be paraphrased as 'the event of...', but not as 'the fact that...' For further discussion, see Zucchi (1993) and Arregui and Matthewson (2001).

[14] So there is no "optionality" in the presence of *ja* and *wa*, but rather three distinct determiners, one of which is null.

(1a) or (17). In addition, *dʒʌ* and *dʒwɤj* are compatible with both an overt internal head,[15] and an external head, which in any case appears to the left of the relative clause.

Instead, we propose that the structure of these "short nominalizations" is just what the morphology leads us to believe: they are full nominalized clauses that are (genitive) complements to the bleached nouns. How exactly "the master of saying writing" comes to mean "teacher" will have to be worked out on another occasion, but the path to follow should be clear.

References

Arregui, Ana, and Lisa Matthewson. 2001. A cross-linguistic perspective on the expression of manner. In *Proceedings of the 36th ICSNL*, 1–24. UBCWPL.

Chung, Sandra, and William Ladusaw. 2004. *Restriction and saturation*. Cambridge, MA: MIT Press.

Heim, Irene. 1982. The semantics of definite and indefinite noun phrases. Doctoral Dissertation, University of Massachusetts, Amherst.

Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*. Oxford, UK: Blackwell.

Kratzer, Angelika. 1996. Severing the external argument from its verb. In *Phrase structure and the lexicon*, ed. J. Rooryck and L. Zaring, 109–137. Dordrecht: Kluwer.

Matthewson, Lisa. 1999. On the interpretation of wide-scope indefinites. *Natural language semantics* 7:79–134.

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

Reis Silva, Maria Amélia. 2001. Pronomes, ordem e ergatividade em Mẽbengokre. Master's thesis, State University of Campinas, SP, Brazil.

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 Simoen van de Kerke. Leiden, Netherlands: Research School of Asian, African, and Amerindian Studies (CNWS), Leiden University.

Salanova, Andrés Pablo. in progress. The syntax and semantics of Mẽbengokre nominalizations. Doctoral Dissertation, MIT.

Williamson, Janis S. 1987. An indefiniteness restriction for relative clauses in Lakhota. In *The representation of (in)definiteness*, ed. Eric Reuland and Alice ter Meulen, 168–190. Cambridge, MA: MIT Press.

Zucchi, Alessandro. 1993. *The language of propositions and events*. Dordrecht: Kluwer.

[15] For instance, *dʒwɤj* can co-occur with an overt agentive subject, though the sense of these constructions is not clear to us at this point.