

## Relative Clauses are Phrase Structural in Malagasy

Edward L. Keenan

*University of California, Los Angeles*

Malagasy, spoken throughout Madagascar, is a Western Austronesian language with a characteristically rich voice system.<sup>1</sup> In this paper we show first how to derive and compositionally interpret nuclear clauses built from verbs in different voices. They are directly projected from verbal affixes, not derived by A or A' movement. Theta role assignment follows explicitly from semantic interpretation. And second, we derive and compositionally interpret relative clauses, again with no A or A' movement. Merits of our analysis compared with most mainstream approaches are:

1. An explicit compositional semantic interpretation of relative clauses, thus (partially) accounting for how speakers interpret novel expressions.
2. A new derivation of the Malagasy “only subjects extract” constraint, which strongly satisfies *Inclusiveness* (Chomsky 2001:2) in that no traces or indices are used.
3. Simplicity and learnability: deriving and interpreting nuclear Ss is localized in about 20 verbal affixes plus feature checking (Agree) under adjacency. Theta role assignment is explicitly effected in the categories built by affixing roots.
4. A cognitive trigger for interpreting extraction by variable binding operators in voice-poor languages like English. English-Malagasy differences are “essentially morphological in character” Chomsky (1996:7).

Key words: Malagasy, relative clause, voice (grammatical)

This paper is organized as follows. Section 1 reviews the case and voice marking patterns present in nuclear clauses and exemplifies our analysis of their syntax and semantics. Section 2 extends that analysis to a sample of the many voice affixes in Malagasy, and §3 studies the *circumstantial* voice, problematic in the literature. Section 4 compares our approach with mainstream ones, and §5 derives and interprets relative clauses building directly off the voice system.

---

<sup>1</sup> The Malagasy voice system is comparable to those of Tagalog (Schachter & Otnes 1972, Kroeger 1993) and Philippine languages generally, Timugon (Prentice 1971), Kimaragang Dusun (Kroeger 1988), Balinese (Wechsler & Arka 1998), Atayal (Huang 2002), and Seediq (Aldridge 2002).

## 1. Case, voice, and nuclear clause structure

### 1.1 Case (Keenan & Polinsky 1998)

Malagasy distinguishes three nominal case forms. Here are the singular pronouns in standard orthography (*o* = /u/, word final *y* = /i/).

(1) <b>Singular Pronouns</b>	gen (genitive)	dft (default)	acc (accusative) <sup>2</sup>
I	<i>-ko</i>	<i>aho</i>	<i>ahy</i>
you	<i>-nao</i>	<i>ianao</i>	<i>anao</i>
he, she, it	<i>-ny</i>	<i>izy</i>	<i>azy</i>

Genitive case is, by definition, the case of possessors of nouns. Full DP<sub>gen</sub>s are bound to their N in a morphophonemically complex way. The distribution of DP<sub>gen</sub>s is much greater in Malagasy than in English. Most objects of Prepositions are genitive, and many voice forms of verbs (below) select their Agents in the genitive. In text counts, DPs occur as genitives much more frequently than in any other case (Keenan & Manorohanta 2001). We illustrate some DP<sub>gen</sub>s built from *trano* ‘house’.

(2) ny tranoko	ny tranonao	ny tranony	ny tranon-dRabe
the house+1s.gen	the house+2s.gen	the house+3.gen	the house+Rabe.gen
‘my house’	‘your (sg) house’	‘his house’	‘Rabe’s house’

In (3a-d) are some simple Ss illustrating the three case markings of DPs. (3a-b) may both be translated *He is chasing you*, (3c-d) as *Rabe is chasing me*.

- (3) a. [p<sub>1</sub> Manenjika anao] izy  
 chase 2s.acc 3.dft
- b. [p<sub>1</sub> Enjehiny] ianao  
 chase+3.gen 2s.dft
- c. [p<sub>1</sub> Manenjika ahy] Rabe  
 chase 1s.acc Rabe.dft
- d. [p<sub>1</sub> Enjehin-dRabe] aho  
 chase+Rabe.gen 1s.dft

There is strong evidence, well agreed upon by generative grammarians (Keenan 1976, 1995, Paul 2001, Sabel 2003, Potsdam 2004, Pearson 2005, Travis 2006) who have

<sup>2</sup> Abbreviations used in this paper are as follow: 1s: 1<sup>st</sup> person singular; 2s: 2<sup>nd</sup> person singular; 3: 3rd person; dft: default; gen: genitive; pres: present; acc: accusative; art: article; tr: transitive.

worked on Malagasy, supporting a major constituent break at the right parenthesis in (3).<sup>3</sup> Theory neutrally, we call the string to its left a P1, *one place predicate*. One constituency test is that the yes-no question particle *ve* occurs at the locus of our right bracket.

- (4) a. Manenjika azy ve Rabe?  
 chase 3.acc ? Rabe  
 ‘Is Rabe chasing him?’  
 b. \*Manenjika ve azy Rabe?  
 chase ? 3.acc Rabe  
 ‘Is Rabe chasing him?’

P1 internal placement as in (4b) is ungrammatical. Further, the P1 is meaningful. It denotes a property of individuals and occurs as a constituent with that meaning in other constructions (relative clauses, nominalizations). Finally, the DP in each of (3a-d), which combines with the P1 to form an S (a P0), has the same case form in all cases. We call this case *default* (a name justified by Pearson 2005). We eschew non-structural terms like *subject*, *topic*, *pivot*, *trigger*, and *nominative*.

## 1.2 Voice

Data in (5) illustrate four voice affixes which build verbs from the root *roso* ‘serve’. The translation of (5a) suffices for (5b-d) as well. We treat voice affixes, both syntactically and (shortly) semantically, as functions which derive verbs from roots.

- (5) a. [<sub>P1</sub>  $\emptyset$ +*maN*+*róso* (mandróso) vary ny vahiny amin'ny  
 pres+*maN*(serve) rice the guest on'the  
 lovianao] izy  
 dishes+2s.gen 3.dft  
 ‘He is serving rice to the guest on your dishes.’  
 b. [<sub>P1</sub>  $\emptyset$ +*a*+*róso*+ny (arósony) ny vahiny amin'ny  
 pres+*a*(serve)+3.gen the guest on'the  
 lovianao] ny vary  
 dishes+2s.gen [the rice].dft  
 c. [<sub>P1</sub>  $\emptyset$ +*róso*+*Vna*+ny (rosóany) vary amin'ny  
 pres+*Vna*(offer)+3.gen rice on'the  
 lovianao] ny vahiny  
 dishes+2s.gen [the guest].dft

<sup>3</sup> We ignore the scope of tense marking, as nothing we say depends on it.

- d. [<sub>P1</sub> ø+[*maN*+róso]+*ana*+ny (androsóany) vary ny  
 pres+*ana*(*maN*(offer))+3.gen rice the  
 vahiny] ny lovianao  
 guest [the dishes+2s.gen].dft

Some comments on (5): (1) The initial *m* in (5a) alternates with *n* for past and *h* for future. So this *m* is structural. It is obligatorily absent in (5d), though *n* and *h* are used for past and future for all the verbs in (5). (2) *N* in *maN* denotes a homorganic nasal (a common Western Austronesian trait). The *ndr* in *mandroso* in (5a), like *n-dR* in (3d), is a single phoneme, a voiced, prenasalized post-alveolar fricative. (3) With most roots the *Vna* suffix in (5c) is realized as *-ina*, sometimes just *-na*. (4) Stress is invariant under prefixing but shifts rightward under suffixing, as indicated. (Stress is minimally phonemic in Malagasy and not usually marked in the orthography). Finally, the P1 constituency tests that apply in (3) apply in (5) as well.

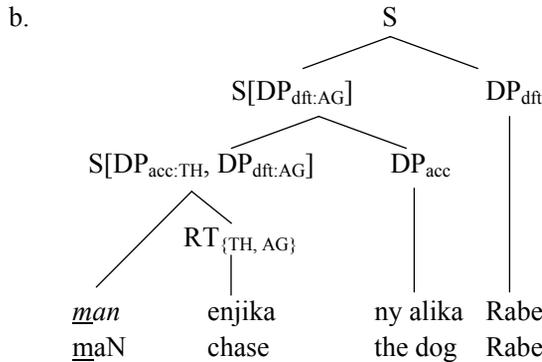
More importantly: the verbs in (5a-c) are all formed by affixing the root. But that in (5d) is formed by suffixing the *verb* in (5a), retaining the prefix less the initial *m*. Malagasy presents about nine *m* initial prefixes like *maN* which suffix *-ana* as in (5d) retaining the prefix less initial *m*. The verbs in (5b-d) pattern together in that they all select a genitive Agent phrase and form imperatives by suffixing *y*: *arosóy*, *rosóy* and *androsóy*. The verb in (5a) has no genitive argument and its imperative suffixes *a*: *mandrosóa* (retaining the initial *m*). All four verbs in (5) are atelic (Travis 2005, Rasolofo 2006), which explains their translation in the progressive. There are other affixes, noted later, which build telic verbs lacking imperative forms. Tense marking, obligatory, merges outside voice morphology and plays no role here.

### 1.3 Nuclear clauses: syntax

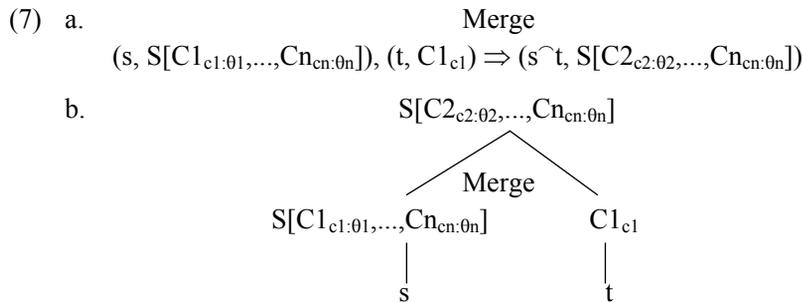
We derive *nuclear* Ss by first combining a verbal root with a voice affix forming a Pn, *n-place predicate*. Then *n* applications of Merge yield a nuclear S. (6b) illustrates our derivation of the nuclear S in (6a). Merge, which combines predicates and their arguments, is defined in (7).<sup>4</sup>

- (6) a. *maN*+enjika (manenjika) ny alika Rabe  
*maN*(chase) the dog Rabe  
 ‘Rabe is chasing the dog.’

<sup>4</sup> Formally, voice affixes are interpreted by functions whose domains include roots or other affixed verbs. We could extend Merge to combine roots and affixes, but this “generalization” is not enlightening.



*enjika* is listed in the Malagasy lexicon as a root with two theta role features: TH for Theme, and AG for Agent. (Theta roles are semantically represented, below). *maN* prefixes that root to form a P2 with two argument features, each specifying the category, case, and theta role of that argument. In the case of *manenjika* ‘chases’ the first argument feature checks (agrees with) a  $DP_{acc}$  interpreted as Theme which forms a P1 that requires a  $DP_{dft}$  interpreted as Agent to form a P0 (“zero features left to check”). Merge is an operation that combines Pns successively with n arguments to form P0s:



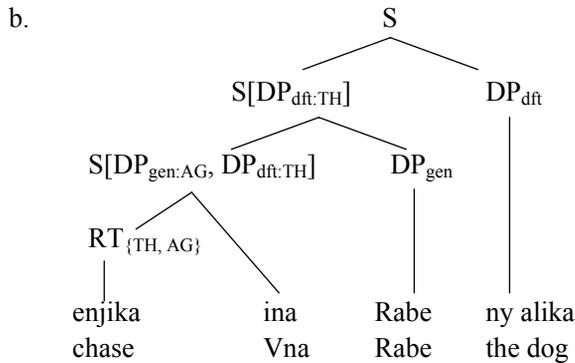
So  $\text{Dom}(\text{Merge})$ , the domain of Merge, contains pairs  $\langle u, v \rangle$  of expressions in which the category and case of  $v$  match (are checked against) the category and case required by  $u$ . Merge concatenates<sup>5</sup> a string  $s$  of category  $S[C1_{c1:01}, \dots, Cn_{cn:0n}]$  with a string  $t$  of category  $C1_{c1}$ , satisfying a category ( $C1$ ) and a case ( $c1$ ) requirement, to form  $s \hat{\ } t$  of category  $S[C2_{c2:02}, \dots, Cn_{cn:0n}]$ . Theta role  $\theta_i$  is assigned to the  $i^{\text{th}}$  argument (below). A *nuclear* S is a P0 derived by successive applications of Merge to a voice-affixed root. All its arguments

<sup>5</sup> More than concatenation is involved. Root initial consonants may drop: *maN*(sasa) = *manasa* ‘washes (tr)’ or map to the closest non-continuant: *maN*(la) = *manda* ‘refuses’. We ignore this Austronesian complexity here.

are assigned a case and interpreted with a theta role. Voice affixes number about 20. To define them we must state their domain—the expressions they combine with, and what their value is at each expression in their domain.

In (8) below we contrast the syntax of Ss built with the *maN* form of “chase” *manenjika*, with the *Vna* form, *enjehina*, built by suffixing the root (often calling an epenthetic consonant and modifying a final “weak” syllable *-ka*, *-na*, or *-tra*).<sup>6</sup>

- (8) a. enjika+*Vna*+Rabe (= enjehin-dRabe) ny alika  
 Vna(chase)+Rabe.gen the dog  
 ‘Rabe is chasing the dog.’



Note that *Vna*(enjika) = *enjehina* is a P2: its first argument is a genitive case DP interpreted as Agent, its second a default case DP interpreted as Theme.

#### 1.4 Nuclear clauses: semantics

In an arbitrary situation *s* with domain *E*—the set of (possibly abstract) objects about which we may be speaking, we semantically interpret P0s as elements of {True, False},<sup>7</sup> P1s as functions from *E* into {True, False} and in general P<sub>*n*+1</sub>s as functions from *E* into the possible P<sub>*n*</sub> denotations. To say that *John laughed* is interpreted as True in a situation *s* with domain *E<sub>s</sub>* says formally that the function *laughed* is interpreted as in *s* maps the object *John* is interpreted as in *s* to True. More generally an *n*+1 place predicate semantically maps entities in *E<sub>s</sub>* to *n*-place predicate denotations. So, for example, *is chasing John* is interpreted as a P1, one that may map Bill to True or to False in *s*. And *is chasing* maps

<sup>6</sup> Pearson (2004, 2005) treats (quite reasonably) epenthetic consonants as part of the root. They are deleted if the root is not suffixed.

<sup>7</sup> A richer system would interpret P0s as *propositions*—functions from an index set *W*, whose elements are called *possible worlds*, into {True, False}.

John to that function.

We enrich this bare (and insufficient) extensional semantics with theta roles, as follows. We take a *theta role* to be a binary relation between elements of  $E_s$  and possible root denotations—themselves n-ary relations over  $E_s$  where n is the number of theta role features (AG, TH, etc.) in the category of the root. For example (*enjika*,  $RT_{\{AG,TH\}}$ ) will denote a binary relation  $[[enjika]]$  over  $E_s$ , whereas (*roso*,  $RT_{\{AG,TH,GL\}}$ ) will denote a ternary relation (GL = GOAL). We build this analysis into the interpretation of the P2 (*manenjika*,  $S[DP_{acc:TH}, DP_{dft:AG}]$ ) as in (9), writing  $[[d]]_s$  for the interpretation of the expression d in situation s (We usually omit the subscript s).

$$(9) \quad \llbracket \underline{maN}(enjika, RT_{\{AG,TH\}}) \rrbracket(y)(x) = \text{True if and only if} \\ (x, y) \in \llbracket enjika \rrbracket \wedge \text{THEME}(y \llbracket enjika \rrbracket) \wedge \text{AGENT}(x, \llbracket enjika \rrbracket).^8$$

We may read the right hand side of (9) as “x and y are participants in a chasing event  $[[enjika]]$ , y is THEME of  $[[enjika]]$  and x is AGENT”. And we may now define:

**Definition** Two nuclear Ss are *theta equivalent* iff their arguments can be bijectively matched, each argument in one having the same theta role as its match in the other.<sup>9</sup>

*Jill opened the door* and *Ruth closed the window* are theta equivalent under the matching *Jill*  $\mapsto$  *Ruth* and *the door*  $\mapsto$  *the window*. What is important is that theta role assignment and thus theta equivalence derives explicitly from semantic interpretation. And it is the entire P2  $\underline{maN}(enjika)$ , not the root or voice affix alone, that associates arguments and theta roles. In (10) we give the semantic interpretation of the *Vna* form of “chase”.

$$(10) \quad \llbracket Vna(enjika, RT_{\{AG,TH\}}) \rrbracket(v)(u) = \text{True if and only if} \\ (v, u) \in \llbracket enjika \rrbracket \wedge (\text{AGENT}(v, \llbracket enjika \rrbracket) \wedge \text{THEME}(u \llbracket enjika \rrbracket))$$

(9) and (10) entail that  $\llbracket \underline{maN}(enjika, RT_{\{AG,TH\}}) \rrbracket(b)(a) = \llbracket Vna(enjika, RT_{\{AG,TH\}}) \rrbracket(a)(b)$ . And this guarantees that (11a-b) are theta equivalent. (The bijection matches *ny alika* ‘the dog’ with *ny saka* ‘the cat’ and *Rabe* with *Rasoa*).

<sup>8</sup> We understand here that  $\llbracket \underline{maN}(enjika, RT_{\{AG,TH\}}) \rrbracket$  is the value of the  $\llbracket \underline{maN} \rrbracket$  function at  $\llbracket (enjika, RT_{\{AG,TH\}}) \rrbracket(y)(x)$ .

<sup>9</sup> The generalized definition must also bijectively match predicates if the Ss have more than one. Then matched arguments must bear the same theta role to matched predicates. So *Nijery azy aho* ‘I watched him’ with just one verb is not theta equivalent to *Nampandihy azy aho* ‘I made-dance him’ with two verbs: (*m*)*andihy* ‘dances’ and (*m*)*ampandihy* ‘makes dance’. *azy* is Agent of ‘dance’ and Theme of ‘makes-dance’, but in the first S *azy* only has one theta role.

- (11) a. Manenjika ny alika Rabe  
b. Enjehin-dRasoana ny saka

Replacing *ny saka* with *ny alika* and *Rasoana* with *Rabe* in (11b), (9) and (10) guarantee that the resulting Ss are logically equivalent (True in the same situations).

Worth noting is that this approach to semantic equivalence is standard in formal logic. We prove that (12a-b) are logically equivalent using their compositional interpretation for each way of interpreting the formulas P and Q. There is no sense in which (12a-b) are derived from the same structure. It is the standard logical meaning of *and*, *or*, *not*, *if-then*, and *neither-nor* that guarantees the logical equivalence of the structurally non-isomorphic expressions in (12a-b).

- (12) a. If P then Q and if Q then P  
b. Either both P and Q or neither P nor Q

Thus far we have seen some examples of the derivation of some voice forms and the theta equivalence of the Ss they build. We now extend our empirical coverage and generalize our formulation.

## 2. Classifying voice forms

It is useful to classify verbs as *+m* if their present tense form is *m* initial and *-m* if not. The *m*-class of a verb is uniquely determined by the most recent voice affix to apply in its derivation so it makes sense to refer to affixes as *+m* or *-m*. And the behavior of many other morphosyntactic affixes is sensitive to the *m*-class of the verb. We have already noted that *+m* verbs form imperatives by suffixing *a*, *-m* ones by suffixing *y*. Further *+m* verbs do not subcategorize a genitive argument whereas *-m* ones always do. Here we note additional distinguishing properties and additional affixes in each class.

### 2.1 *+m* verbs

**2.1.1** *maN* has already been illustrated. Its domain is just roots, many with two theta roles, like *enjika* ‘chase’ yielding the P2 *manenjika*. Some other examples are: *halatra* ⇒ *mangalatra* ‘steals’, *haja* ⇒ *manaja* ‘respects’, *rakotra* ⇒ *mandrakotra* ‘covers’. *maN* also derives P3s from roots with three theta roles, as we saw in (5). Also *ome* ⇒ *manome* ‘gives’, *solo* ⇒ *manolo* ‘substitute’, *tolotra* ⇒ *manolotra* ‘offers’. *maN* is the only affix to build P3s from roots. Also *maN* forms two classes of P1s: one with an Agent theta role, as in *sidina* ⇒ *manidina* ‘flies (as in *Birds fly*)’, *leha* ⇒ *mandeha* ‘goes’, and *dihy* ⇒ *mandihy*

‘dances’. The other has a Theme theta role, as in *hatsiaka* ⇒ *mangatsiaka* ‘is cold’, *hetaheta* ⇒ *mangetaheta* ‘is thirsty’, and *firapiratra* ⇒ *mampiratra* ‘shines’. So the valence of a *maN* derived verb is unpredictable, though most are P2s or P3s.

**2.1.2** *mi*, like *maN* prefixes just roots, but unlike *maN* it preserves the root. It is the most frequent and most productive voice affix (Keenan & Manorohanta 2001). *mi* builds many P1s of category S[DP<sub>dft:AG</sub>]: *miasa* ‘works’, *mihomehy* ‘laughs’, *milomano* ‘swims’. It also builds some Theme argument P1s: *miely* ‘disperse’, *mianjera* ‘collapse’, *mifaly* ‘rejoice’, *migadra* ‘to be in chains’. And *mi*, like *maN*, builds P2s of category S[DP<sub>acc:TH</sub>, DP<sub>dft:AG</sub>]: *mikapoka* ‘beats’, *mividy* ‘buys’, *mitady* ‘seeks’, *mikarakara* ‘takes care of’, *mijery* ‘watches’, *miantso* ‘calls’, *misambotra* ‘captures’. *maN* applies to none of these roots, so if we know of a Pn that it interprets its DP<sub>dft</sub> as Agent we cannot predict the verbal prefix: it is *maN* with *enjika* ‘chase’ and many others, but *mi-* with *kapoka* ‘beat’ and many others.

*mi* never builds ditransitive verbs (P3s) of transmission, in distinction to *maN*. But *mi* and *maN* do have many roots in their common domain. And Rahajarizafy (1960:50) notes that in such cases the *mi* verb usually has lesser valence. Often (Travis 2000) *maN(r)* is a causative of *mi(r)*: *miverina* ‘come back’ / *mamerina* ‘give back’; *mihidy* ‘is locked’ / *manidy* ‘locks (tr)’; *milatsaka* ‘falls’ / *mandatsaka* ‘drops’, so this contrast is useful to language learners.

**2.1.3** *ma* is a non-productive prefix which derives *m*-class verbs from a few common roots: *hita* ⇒ *mahita* ‘sees’, *tory* ⇒ *matory* ‘sleeps’ and *handro* ⇒ *mahandro* ‘cooks’.

**2.1.4** ∅—About 20 vowel initial roots form the present and imperative by prefixing *m* directly. We treat them as having a zero prefix: *-ila* ‘need’, *-ino* ‘believe’, *-indrana* ‘borrow’, *-anana* ‘possess’, *-iditra* ‘enter’, *-onina* ‘reside’. Some are also in Dom(*Vna*).

- (13) a. *m*+∅+ila (míla) azy Rabe  
           *m*(need)                   3.acc Rabe.dft  
           ‘Rabe needs it.’  
       b. *ila+ina*+Rabe (iláin-dRabe) izy  
           *Vna*(need)+Rabe.gen           3.dft  
           ‘Rabe needs it.’

### 2.1.5 Common properties of *+m* verbs

By definition *+m*-verbs prefix *m* in the present tense (and retain it in the imperative). And as noted they form imperatives by suffixing *a*. In addition: they always lie in Dom(*ana*), the *circumstantial* affix, which applies to *+m* verbs deriving *-m* ones illustrated in (5d) and discussed in §3. The *+m* causative prefix (Randriamasimanana 1986, Andrianierenana 1996) *mamp* applies only to *+m* verbs, deriving *+m* ones. Ditto for the reciprocal prefix *mif* (Keenan & Razafimamonjy 2004). Lastly, *+m* verbs usually form agent nominals (English *-er*) by replacing the initial *m* with *mp*. We exemplify these prefixes below but do not study them here. Our point is that, to learn to use and understand the *ana* suffix and the causative, reciprocal, and agent nominal prefixes, speakers must learn to recognize the *m*-class verbs.

- (14) a. *mif*(*maN*(*soratra*)) = *mifanoratra* ‘write to each other’ as in  
Mifanoratra (taratasy) Rabe sy Raso  
‘Rabe and Raso write (letters) to each other.’
- b. *mamp*(*mif*(*maN*(*soratra*))) = *mampifanoratra* ‘make write to each other’  
as in  
Mampifanoratra azy ireo aho  
‘I make them write to each other.’
- c. *mp*(*maN*(*soratra*)) = *mpanoratra* ‘writer’  
*mp*(*mamp*(*mi*(*anatra*))) = *mpampianatra* ‘teacher ([cause-to-learn]-er)’

There are another four *m* initial voice prefixes (*manka*, *mana*, *miaN*, and *maha*) of which *maha* expressing indirect cause or potentiality (Phillips 2000) is very productive and has a telic interpretation (Travis 2005).

## 2.2 *-m* verbs

*-m* verbs by definition do not form present tense with *m*. They form imperatives by suffixing *y* (not *a*), and the first argument they select is always genitive.

### 2.2.1 *Vna*

Suffixed roots and *mamp-* (causative) prefixed verbs with *-ina*, *-(e)na* or *ana*, retain (or epenthesize) root final consonants. The derivation of (15b) is in (8b).

**Vna**

- (15) a. (énjika,  $RT_{\{AG,TH\}}$ )  $\mapsto$  (enjéhina,  $S[DP_{gen:AG}, DP_{dft:TH}]$ )  
 b. enjika+ina+Rabe (= enjehin-dRabe) ny alika  
 Vna(chase)+Rabe.gen the dog  
 ‘Rabe is chasing the dog.’

The verb in (15b) interprets its first (genitive) argument as Agent, its second (default) case argument as Theme. In general *Vna*(*r*) takes at least two arguments, per (16).

- (16) *Vna*(*r*) has category  $S[DP_{gen:01}, \dots, DP_{dft:0n}]$ , some  $n > 1$

Many roots lie both in  $Dom(\underline{ma}N)$  and  $Dom(Vna)$ : *haja*  $\Rightarrow$  *manaja* / *hajaina* ‘respects’, *vono*  $\Rightarrow$  *mamono* / *vonoina* ‘hits, kills’, *la*  $\Rightarrow$  *manda* / *lavina* ‘refuses’, *fonos-*  $\Rightarrow$  *mamono* / *fonosina* ‘wraps’, *vaky*  $\Rightarrow$  *mamaky* / *vakina* ‘cuts, reads’. Also many lie in both  $Dom(\underline{mi})$  and  $Dom(Vna)$ : *bata*  $\Rightarrow$  *mibata* / *bataina* ‘raises up’, *kapoka*  $\Rightarrow$  *mikapoka* / *kapohina* ‘beats’, *antso*  $\Rightarrow$  *miantso* / *antsoina* ‘calls’. Note that (9) and (10) jointly imply (17) when *r* is a root in  $Dom(\underline{ma}N)$  and  $Dom(Vna)$ . The verbs are logical converses.

- (17)  $[[\underline{ma}N(r)](y)(x)] = [[Vna(r)](x)(y)]$

**Proposition 1** For *r* of category  $RT_{\{AG,X\}}$  in the domain of both  $\underline{ma}N$  and *Vna*,

- a. nuclear Ss built from  $\underline{ma}N(r)$  and *Vna*(*r*) are theta equivalent: the  $DP_{dft}$  argument of each has the same theta role as the non- $DP_{dft}$  argument of the other.  
 b.  $\underline{mi}(r)$  and *Vna*(*r*) build theta equivalent Ss when  $\underline{mi}(r)$  is transitive (a P2).

**Proposition 2** For *r* of category  $RT_{\{AG,X\}}$ , nuclear Ss built from  $\underline{ma}N(r)$  and *Vna*(*r*), are logically equivalent when their Agents are the same individual denoting DPs and their Themes are, too. Similarly for  $\underline{mi}(r)$  and *Vna*(*r*) in **Prop 1b**.

**2.2.2 A**

Prefixes *a* to roots, as in (5b), and like *mi* and  $\underline{ma}N$  only to roots. **Prop 1** holds with *a*(*p*) replacing *Vna*(*p*). The  $DP_{dft}$  of a P2 built by *a* is an “Intermediary”: Theme with roots of movement or transmission, Instrument or Theme otherwise. See Paul (1999) for extensive discussion. (18b) illustrates its Theme usage.

- (18) a. [*maN*+*tosika* (*manosika*) *ny fiara*] Rabe  
           *maN*(push) [the car].acc Rabe.dft  
           ‘Rabe pushes/is pushing the car.’  
       b. [*a*+*tosika*+Rabe (*atosi-dRabe*)] [*ny fiara*].dft  
           a(push)+Rabe.gen the car  
           ‘Rabe pushes/is pushing the car.’

The root *tosika* is in Dom(*a*) but not Dom(*Vna*). Other examples: *tao* ‘do, make’ → *manao*, *atao*, \**ataovina*; *verina* ‘return’ → *mamerina*, *averina*, \**verenana*; *lahatra* ‘line up’ → *mandahatra*, *alahatra*, \**laharina*. There are even more roots in Dom(*Vna*) not in Dom(*a*): *enjika* ‘chase’, as in (15). Also *lalao* ‘play’ → *milalao*, *lalaovina*, \**alalao*; *la* ‘refuse’ → *manda*, *lavina*, \**ala*; and *haja* ‘respect’ → *manaja*, *hajaina*, \**ahaja*.

So merely given a Pn with DP<sub>dft</sub> Theme and its DP<sub>gen</sub> Agent we cannot predict its voice affix. It is *-ina* with *enjika* ‘chase’ and *a-* with *tosika* ‘push’. With 3-place roots of transmission, as in (5), usually both *a* and *Vna* apply. *a* assigns Theme to its DP<sub>dft</sub> and *Vna* assigns Goal. Similar roots are *tolotra* → *atolotra* / *tolorana* ‘serve’ and *solo* → *asolo* / *soloana* ‘replace, substitute’. There are also roots in which *Vna* assigns Theme to the P1 sister, and *a* Instrument (or Intermediary).

- (19) a. *maN*+*tondro* (*manondro*) *ny tranony* amin'ny *tehin*y Rabe  
           *maN*(point-out) the house-his with'the cane-his Rabe  
           ‘Rabe points out his house with his cane.’  
       b. *a*+*tondro*+Rabe (*atondron-dRabe*) *ny tranony* ny *tehin*y  
           a(point-out)+Rabe.gen the house-his the cane-his  
       c. *tondro+ina*+Rabe (*tondroin-dRabe*) amin'ny *tehin*y ny *tranony*  
           *Vna*(point-out)+Rabe.gen with'the cane-his the house-his

Other roots like *tondro* (Paul 1999) are *didy* ‘cut’, *fefy* ‘fence in’, *rakotra* ‘cover’. (20) gives the interpretations of *maN*, *a*, and *Vna* at 3-place roots *r* of category  $RT_{\{AG, GL, TH\}}$ .

- (20) a. [*maN*(*r*,  $RT_{\{AG, TH, GL\}}$ )](*z*)(*y*)(*x*) = True iff  
           <*z*,*y*,*x*> ∈ [*r*] ∧ THEME(*z*, [*r*]) ∧ GOAL(*y*, [*r*]) ∧ AGENT(*x*, [*r*])  
       b. [*a*(*r*,  $RT_{\{AG, TH, GL\}}$ )](*x*)(*y*)(*z*) = True iff  
           <*z*,*y*,*x*> ∈ [*r*] ∧ AGENT(*x*, [*r*]) ∧ GOAL(*y*, [*r*]) ∧ THEME(*z*, [*r*])  
       c. [*Vna*(*r*,  $RT_{\{AG, TH, GL\}}$ )](*x*)(*z*)(*y*) = True iff  
           <*z*,*y*,*x*> ∈ [*r*] ∧ AGENT(*x*, [*r*]) ∧ THEME(*z*, [*r*]) ∧ GOAL(*y*, [*r*])

Thus (5b-c) are theta equivalent, so this notion does not just apply to pairs we might

analogize to active vs passive in English. And **Prop 1** extends to:

### Proposition 3

- a. Nuclear Ss built from P2s *maN(r)* and *a(r)* are theta equivalent, the DP<sub>dfi</sub> argument of one bearing the same theta role as that of the non-DP<sub>dfi</sub> argument of the other.
- b. For *r*, a ditransitive root of transmission in the domains of *maN*, *a*, and *Vna*, the nuclear Ss they build from *r* are theta equivalent, as in (5a-c).

### 2.2.3 Root verbs

Malagasy has many roots which function as *-m* P2s without affixation. They usually form *+m* verbs with *maN* or *mi* and are often in Dom(*Vna*). The *Vna* form is dynamic, the root form more static (Paul 1999). Root predicates, such as adjectives, do not mark past tense and only mark future with *ho* even when vowel initial.

- |      |               |               |                     |               |
|------|---------------|---------------|---------------------|---------------|
| (21) | Root <i>r</i> | <i>ina(r)</i> | <i>man(r)/mi(r)</i> |               |
|      | tápaka        | tapáhina      | manápaka            | ‘cut, decide’ |
|      | hadíno        | hadinóina     | manadíno            | ‘forget’      |
|      | héno          | henóina       | mihéno              | ‘listen to’   |
- (22) a. Tsy mbola tapaky ny olon-dehibe izany raharaha izany  
not yet cut the person-big.gen that work that  
‘That matter isn’t/hasn’t been decided by the leaders yet.’
  - b. Notapahin'ny zaza maditra ity zana-kazo ity  
past+Vna(cut)the child naughty.gen this offspring-tree this  
‘This young seedling was cut off by the naughty child.’
  - c. Tsy mbola nanapa-kevitra (nanapaka hevitra) ny fanjakana...  
not yet past+maN(cut-thought) the government  
‘The government still hasn’t decided.....’

### 2.2.4 Voa

Prefixes *voa* to roots forming *-m* predicates. Its domain overlaps significantly with that of *Vna*: *voatolotra / tolорana* ‘offered’, *voafitaka / fitahina* ‘deceived’, *voavonjy / vonjena* ‘saved’. But *voa*-verbs are telic (Travis 2005) and focus on the final state of an activity, *Vna* on the process. *Voa*-verbs lack imperatives, and like bare roots do not mark past tense. They select a DP<sub>gen</sub> Agent naturally, but it is often absent in discourse. Of the various *-m* verbs the *voa*- ones are most like English passives.

- (23) a. *voa*+kapoka (voakapoka) ilay alika  
       *voa*(beat)                               that (aforementioned) dog  
       ‘That dog was beaten.’  
       b. *voa*+laza+ko (voalazako) izany  
       *voa*(say)+1s.gen                       that  
       ‘I said that.’

Further *-m* affixes are *tafa-* and the non-productive infix *-in/-on-*, as in *tafatolotra* and *tonolotra* ‘offered’. *Tafa*-verbs express action that was either spontaneous on the part of the Agent or unexpected, whence they sometimes have a “manage to” interpretation:

- (24) *tafa*+iditra+ko (tafiditro) ny omby  
       *tafa*(enter)+1s.gen               the cow  
       *I managed to get the cow(s) in (the pen)*

### 2.3 Summary

We have derived and interpreted many affixal functions yielding many theta equivalent Ss. None uses movement. So we know now how an explicit semantics can guarantee theta equivalence when DPs with the same theta role originate in structurally different positions, a standard semantical pattern. We turn now to *ana* suffixation to *+m* verbs, the most complicated case for all approaches.

## 3. *ana* voices (called *circumstantial* or *relative* in the literature)

### 3.1 Syntax

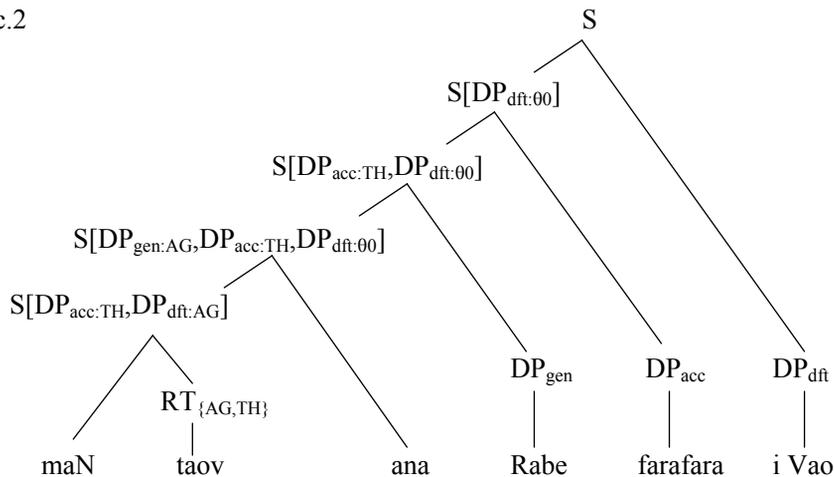
*ana* suffixes *ana* to *+m* verbs of any valency, including ones built from reciprocal *mif-*, causative *mamp-*, and abilitative / indirect causative *maha*. *Ana* increases valency, deriving *-m* Pn+1's from *+m* PnNs. (5d), (25b) and (26b) are illustrative.

- (25) a. [n+*ma*N+enjika (nanenjika) ny jiolahy t+amin'ny fiara] Rabe  
       past+man+chase                       the thief past+with'the car Rabe.dft  
       ‘Rabe chased the thief by means of the car.’  
       b. [n+[*ma*N+enjika]+*ana*+Rabe (nanenjehan-dRabe) ny jiolahy]  
       past+[*ana*[*ma*N(chase)]]+Rabe.gen                       the thief  
       ny fiara  
       [the car].dft  
       ‘Rabe chased the thief in/with/...the car.’

- (26) a. [maN+taov (manao) farafara ho an'i Vao] Rabe  
maN(make) bed for'art Vao Rabe.dft  
 'Rabe is making a bed for Vao.'
- b. [ $\emptyset$ +[[maN+taov]+ana]+Rabe (anaovan-dRabe) farafara] i Vao  
 pres+ana(maN(make))+Rabe.gen bed [art Vao].dft  
 'Rabe is making a bed for / because of Vao.'

*ana* verbs form present, past, future and imperatives as do *-m* verbs generally. The DP<sub>dft</sub> of an *ana* verb usually has an oblique theta role—Benefactive, Instrument, Cause, Location, Time, Manner, ... not selected by the *+m* verb it applies to (Rajemisa-Raolison 1971:112-113, Keenan & Polinsky 1998). Theta roles of the DP<sub>dft</sub> of an *ana* verb are grammatically constrained but not uniquely determined, the precise value depending on context of utterance. In (26a) the Prep *ho an'* forces a Benefactive role for *Soa* but in (26b) the role of DP<sub>dft</sub> is any appropriate oblique. Usually few are appropriate—DPs denoting Instruments, Times, and Benefactees are largely distinct. But in (26b) Cause is possible (Maybe Vao begged Rabe so much he gave in). (27a) illustrates the application of *ana* to a P2. (27b) is the general statement and (27c) illustrates (27a).

- (27) a.  $(v, S[DP_{acc:TH}, DP_{dft:AG}]) \mapsto (v+ana, S[DP_{gen:AG}, DP_{acc:TH}, DP_{dft:00}])$   
 $ana$
- b.  $(v, S[DP_{c1:01}, \dots, DP_{dft:0n}]) \mapsto (v+ana, S[DP_{gen:0n}, DP_{c1:01}, \dots, DP_{cn-1:0n-1}, DP_{dft:00}])$   
 c.1 anaovan-dRabe farafara I Vao (Rabe made a bed for / because of Vao)  
 c.2



Note that on our syntactic analysis the *+m* verb is a constituent of the *ana* derived verb.



is basic: *man-* assigns case to the Theme, so the otherwise caseless Actor raises to Spec, IP to check case. *-ina* assigns case to the Actor, so the Theme raises to Spec, IP. The “two subjects” are Spec, IP and Spec, VP and DP movement is always A-movement.

Pearson in contrast, proposes that the DP sister to P1 behaves like a topic in V2 languages such as Icelandic. He treats a clause as a TopP (Topic Phrase) whose Spec is initially filled with a DP. The complement of the head of TopP is whP whose Spec is initially empty. The complement of its head is TP (Tense Phrase) and the complement of its head is the VP which includes all the arguments of the verb, one of which must be an empty operator. Theta roles and case are assigned *in situ*. The empty operator moves to Spec, whP where it is coindexed with the DP in Spec, TopP passing on its case and theta role. Voice morphology, *man-* and *-ina*, is spelled out according to the position from which the empty operator was moved (an instance of A' movement). So movement determines morphology, rather than morphology determining what can move. But as in Guilfoyle, Hung & Travis (1992) and Paul (1999), theta roles are assigned in fixed positions pre-movement. A DP interpreted as Agent is either in Spec, VP or is coindexed with an empty operator that originated there. Analogously for DP Themes.

The determining difference between these two approaches and ours is that on theirs, theta roles are assigned in pre-movement structure to unique positions (a nearly constant property of Passive from Chomsky 1957 through Mahajan 1995 to Collins 2005). Both assume some form of UTAH (Baker 1988, 1996) and DPs with an Agent (Theme) theta role must move to obtain voice variants in which Agents (Themes) occur in different positions. On our approach by contrast the theta roles of DP arguments are determined by the interpretation of the voice affix. Different affixes assign the same theta roles to syntactically different arguments. So UTAH does not hold and theta equivalence is determined by compositional interpretation. It is still true on our approach that theta roles are assigned as a function of structure.

Recently however, Travis (2006) has suggested an analysis of Malagasy clause structure which differs from the mainstream approaches above in ways rather similar to ours. She analogizes Malagasy voice affixes to verbal clitic pronouns as in Italian, thereby allowing voice variant Ss to be independently generated. The clitic pronoun would bear some sort of binding/agreement relation to the argument of the P1 (our term) the cliticized verb derives. So voice variant clauses would differ morphologically rather than being derived from a common source by movement. Her approach and ours still differ, but are closer to each other than either is to any of the mainstream ones discussed above. And her morphologically different primitive P1s could function as relative clauses, as we advocate (below).

## 5. Relative clauses (RCs) and “subjects only” (Keenan, to appear)

Interpreting RCs in Malagasy (as elsewhere) is usually modeled on operator (*wh* or empty) movement in English, as in (30d) where  $\phi$  is any P0, E an arbitrary domain.

- (30) a. manasa (maN+sasa) lamba amin'io savony io ny tovolohy  
washes (maN(wash)) clothes with'that soap that the young.man  
b. tovolahy (izay) manasa lamba amin'io savony io  
young.man (that) washes clothes with'that soap that  
c. tovolahy Op<sub>i</sub> [[<sub>P1</sub> manasa lamba amin'io savony io] [t<sub>i</sub>]  
young.man who<sub>i</sub> washes clothes with'that soap that t  
d.  $[(Op_i)\phi](B) = B \cap \{b \in E \mid [\phi][t_i/b] = \text{True}\}$  ( $\cap$  is set intersection)

(30d) says that for  $\phi$  a P0 of any complexity,  $[(Op_i)\phi]$  maps a set B—the set of young men in (30c)—to the set of those  $b$  in B that  $\phi$  is true of when  $t_i$  is set to denote  $b$ . So a NP, like *tovolahy* ‘young man’, denotes a set and the expressions in (30b) and (30c) denote sets as well, subsets of that denoted by *tovolahy*. The semantics in (30d) is compositional and extensionally correct. That  $\phi$  be arbitrary is important, for then if the position of  $t_i$  in  $\phi$  violates Subjacency, Coordinate Structure, etc. (*wh<sub>i</sub>*) $\phi$  is still interpretable, ungrammaticality is due to syntactic factors.

In (31a) we cannot relativize the Theme argument from (30a). But its meaning is expressed by the RC in (31a') using the *Vna*-verb. In (31b) we attempt to relativize the Instrument from (30a), but only (31b') with an *ana*-verb is grammatical.

- (31) a. \*ny lamba (izay<sub>i</sub>) manasa t<sub>i</sub> amin'io savony  
the clothes (that) maN(wash) t with'that soap  
io ny tovolahy  
that the young.man  
a'. ny lamba (izay<sub>i</sub>) sasan'ny tovolahy  
the clothes (that) Vna(wash)'the young.man.gen  
amin'io savony io t<sub>i</sub>  
with that soap that t<sub>i</sub>  
b. \*ny savony (izay<sub>i</sub>) manasa lamba (amina) t<sub>i</sub> ny tovolahy  
the soap (that) maN(wash) clothes (with) t<sub>i</sub> the young.man  
b'. ny savony (izay<sub>i</sub>) anasan'ny tovolahy lamba t<sub>i</sub>  
the soap (that) ana(maN(wash))'the young.man.gen clothes t<sub>i</sub>  
'the soap that the young man washed clothes with'

b". \*ny savony (izay<sub>i</sub>) sasan'ny tovolahy lamba (amina) t<sub>i</sub>  
 the soap (that) Vna(wash)'the young.man.gen clothes (with) t<sub>i</sub>

The paradigm in (31) is well supported (Keenan 1972, Pearson 2005) and one says “Only subjects relativize”. But this statement misses a striking generalization:

(32) **RCs in Malagasy are just P1s**, not arbitrary Ss (with a trace)

The acceptable examples in (31) are ones where what follows the head nominal (or *izay* if present) is independently a P1. Thus, informally:

(33) **The Malagasy Relative Clause Generalization: RC = P1**  
 Syntax: Malagasy headed RCs have the form: NP + (*izay*) + P1  
 Semantics:  $[[NP + (izay) + P1]] = [[NP] \cap \{x \mid [P1](x) = \text{True}\}]$

The invariable particle *izay* is usually absent. It plays no role in interpretation, so it does not matter whether it is syntactically present. Formally we derive headed RCs by extending Dom(Merge) to include pairs of an NP and a P1, as in (34):

(34)  $\text{Merge}((s, NP), (t, P1)) = (s \wedge (izay) \wedge t, NP)$

Note that P1s do not combine with NPs to form P0s, but only with DPs—the result of combining a Det with an NP, as in *ny tovolahy* ‘the young man’, *ity tovolahy ity* ‘this young man’, etc.

**RC = P1** is also pleasingly Minimalist. To form and interpret RCs, we just use P1s we have already formed and interpreted. We do not introduce novel elements—variable binding operators or bound variables (traces) in the computation. That is what the voice system does. And we do not need to constrain the positions we can extract from, as we are not extracting. But the content of “Subjects Only” holds.

From an Economy perspective (33) is not surprising. A grammar with seven morphologically distinct forms of *offer* to say *I offered rice to the guests on this dish* will use this apparatus for something. And it does. It codes what we can relativize (question, cleft, ...), functionally comparable to relative pronouns in Latin. Similarly the rich noun class system of Bantu implies the existence of an agreement system (Corbett 1991:105).

## 5.1 Why do English and Malagasy RCs differ?

The formation and interpretation of RCs in Malagasy is simpler than in English: no null operators or traces, no variable binding operators (VBOs). What prompts learners of English to adopt this more complex apparatus? Malagasy gives a clue: for non-subject RCs in English the learner must assign a meaning to strings not independently interpreted in simple Ss. Compare (35a) with the RCs underlined in (35b-c).

- (35) a. John served the rice on the new dishes  
b. the rice John served on the new dishes  
c. the new dishes John served the rice on

The RCs in (35b-c) are not syntactic constituents of (35a) and are not assigned a meaning there. So to interpret (35b-c) the learner must assign meanings to strings that do not independently have one. Invoking a logical constant, a VBO, is a reasonable move. If it is constant there is only one thing (in addition to what is audibly present) we need to learn to interpret the RCs in (35b-c). If not a logical constant its interpretation would vary with how the world is—just like the denotation of *is asleep* or *student* so varies. With no audible cue we would not know which of its possible denotations was intended. Moreover it should determine a P1 (up to isomorphism) as that is the meaning RCs have. A null (or *wh*) operator is a minimal way of fulfilling these functions.

## 5.2 Does RC = P1 capture all RCs in Malagasy?

Since the variable (trace) bound by the operator makes sense in any DP position we might expect that English could form many RCs with no natural translation in Malagasy. In practice linguists suggest many constraints on what can be relativized—Coordinate Structure, Subjacency, Phase Impenetrability, .... Still English can relativize possessors, (36a), and arguments within sentential objects, (36b). Are these RCs expressible as P1s in any natural way in Malagasy?

- (36) a. the [student whose name I forget]  
b. the [thief (who) John thinks that Mary caught]

**5.2.1** Possessors in English are not that easy to relativize. Pied Piping is somewhat learned and we sometimes hear *a student that I can never remember his name*, with a resumptive pronoun, rather than *a student whose name I can never remember*. And semantically we reconstruct the “DP shell”, interpreting (36a) as *the student who<sub>i</sub> I forget*

*t<sub>i</sub>'s name*. Cross linguistically (Keenan & Comrie 1977), some languages relativize possessors directly and some do not. Malagasy does not, (37b). But what it can do is bind possessor positions within the P1, as in (38a-b).

(37) a. [<sub>P1</sub> Hadinoko] [ny anaran'ny mpianatra]  
forget+1s.gen the name'the student.gen  
'I forget the name of the student.'

b. \*ny mpianatra (izay) hadinoko ny anarana  
the student (that) I+forget the name

(38) a. [<sub>P1</sub> Hadinoko anarana] ny mpianatra  
forget+1s.gen name the student  
'The student is name forgotten by me.'

b. ny mpianatra (izay) [<sub>P1</sub> hadinoko anarana]  
the student that forget+1s.gen name  
'the student that is name forgotten by me.'

Here *hadino* is a *-m* root. *Hadinoko anarana* 'to be name forgotten by me' is clearly a P1 in (38a): the question particle *ve* occurs at its right edge, its sister *ny mpianatra* 'the student' is replaceable with a default case pronoun, and it occurs where P1s occur in RCs, (38b). Let us add that they coordinate naturally with unequivocal P1s, as in (39).

(39) a. [Be sandry sady mihetraketraka] Rabe  
big arms and mi(look-for-trouble) Rabe.dft  
'Rabe has big arms and looks for trouble.'  
(Keenan & Ralalaoherivony 2000:5)

b. [Very hevitra ka miherikerika fotsiny] Ranaivo  
lost thought and+so mi(look-around) only Ranaivo.dft  
'Ranaivo is confused and so just looks around.'

c. [Miasa loha sady mifoka loatra] i Soa  
mi(work) head and mi(smoke) too+much [art Soa].dft  
'Soa is preoccupied and smokes too much.'

The bound possessor may even be an Agent phrase of a *-m* verb within the P1: (40b). And if needed, an overt verb of possession can be used, as in (41a-b).

(40) a. Maro [ny raharaha sahaniko (sahana+ina+ko)]  
many the tasks Vna(confront)+1s.gen  
'The tasks I am confronting are numerous.'

- b. [Maro raharaha sahanina]<sub>P1</sub> aho  
 many tasks Vna(confront) 1s.dft  
 ‘I have many things to do.’
- c. mpampianatra maro raharaha sahanina  
 teacher many tasks Vna(confront)  
*teacher(s) with many things to do*
- (41) a. dokotera (izay) [m+anana trano ngeza-be]<sub>P1</sub>  
 doctor that m(0(has)) house really-big  
 ‘doctor who has a really big house’ (≈ ‘doctor whose house is really big’)
- b. trano [m+isy tafo fanitso]<sub>P1</sub>  
 house m(0(be/has)) roof tin  
 ‘house with a tin roof’ (≈ ‘house whose roof is tin’)

**5.2.2** RC formation into complement clauses is derived by a natural recursion, given one new way to form complex P1s. Consider first the natural RC in (42).

- (42) ny [jiolahy (izay) [heverin-dRabe ho nosamborin-dRasoa]<sub>P1</sub>]<sub>NP</sub>  
 the thief (that) Vna(think)+Rabe.gen “as” past+Vna(capture)+Rasoa.gen  
 ‘the [thief (that) Rabe thinks that Rasoa captured]’

Full sentence complements occur clause finally, (43a-b). This is expected when the matrix verb is *-m* as in (43a). We assume here that the *+m* order in (43b) is derived by extraposing the CP in (43b’) satisfying a heavy-to-the-right, light-to-the-left condition.<sup>10</sup>

- (43) a. heverin-dRabe [fa [nisambotra ny jiolahy Rasoa]<sub>P0</sub>]<sub>CP</sub>  
 Vna(think)-Rabe.gen that past+mi(capture) the thief Rasoa  
 ‘Rabe thinks that Rasoa captured the thief.’
- b. mihevitra Rabe [fa [nisambotra ny jiolahy Rasoa]<sub>P0</sub>]<sub>CP</sub>  
mi(think) Rabe.dft that past+mi(capture) the thief Rasoa  
 ‘Rabe thinks that Rasoa captured the thief.’
- b’. \*mihevitra [fa [nisambotra ny jiolahy Rasoa]<sub>P0</sub>]<sub>CP</sub> Rabe  
mi(think) that past+mi(capture) the thief Rasoa Rabe.dft  
 ‘Rabe thinks that Rasoa captured the thief.’

<sup>10</sup> It is actually easier for us to design the category of mi(hevitra) to take the default case DP first, followed by the CP argument. But for reasons we lack the space to discuss, we think the more complicated approach assumed here is correct.

We take the categories of *mihevitra* and *heverina* above to be as in (44).

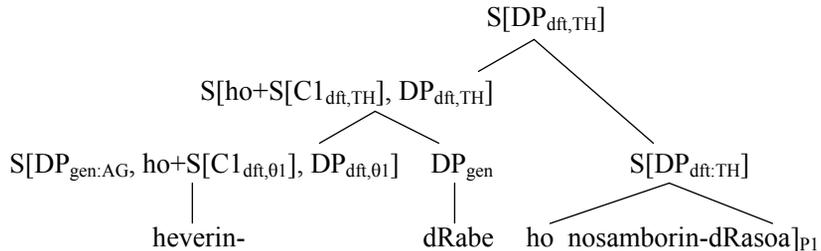
- (44) a.  $\underline{mi}(\text{hevitra}): S[DP_{\text{df}:AG}, CP[\text{fa}]_{\text{acc,TH}}]$   
 b.  $Vna(\text{hevitra}): S[DP_{\text{gen}:AG}, CP[\text{fa}]_{\text{acc,TH}}]$

Consider now (45a-b), which adds these verbs to Lex in a second category which derives the right judgments of paraphrase and theta equivalence.

- (45) a.  $\underline{mi}(\text{hevitra}): S[DP_{\text{acc},\theta 1}, ho+S[C1_{\text{df},\theta 1}], DP_{\text{df},AG}]$   
 b.  $[\underline{mi}(\text{hevitra})](b, ho(p), d) = [\underline{mihevitra}](fa(p(b))(d))$   
 (46) a.  $[mihevitra \text{ ny } jiolahy_{\text{acc}} \text{ ho}([nosamborin-dRasoa]_{P1})]_{P1} \text{ Rabe}$   
 b.  $[heverin-dRabe \text{ ho}([nosamborin-dRasoa]_{P1})]_{P1} \text{ ny } jiolahy$

What is new here is the presence of a *ho*-Phrase consisting of the grammatical constant *ho* followed by a P1. The grammatical status of *ho*-Phrases is treated extensively in Paul & Rabaovololona (1998) who cite over 50 verbs like *mihevitra/heverina*. They leave the precise status of *ho*+P1 open, showing it to differ in crucial respects from Raising-to-Object, ECM and object control structures in English. For our purposes it suffices to say that *ho* is interpreted as a one-to-one function taking the P1 denotation as argument, for then given  $[\text{ho}(p)]$  we can recover the denotation of the P1.<sup>11</sup> In (45a) *mihevitra* ‘thinks’ takes three arguments: a  $DP_{\text{acc}}$ , a *ho*(P1) and a  $DP_{\text{df}}$ . The semantic interpretation of *mihevitra* ‘thinks’ on this categorization guarantees that (46a) is logically equivalent to (43b). Finally, the category of the corresponding *Vna*-verb is:

- (47) a.  $Vna(\text{hevitra}): S[DP_{\text{gen}:AG}, ho+S[C1_{\text{df},\theta 1}], DP_{\text{df},\theta 1}]$   
 b.  $[Vna(\text{hevitra})](d, ho \text{ p}, b) = [Vna(\text{hevitra})](d)(fa(p(b)))$   
 c.



And this P1 is precisely the relative clause in our target structure (42). Thus again

<sup>11</sup> But this does mean that theta role assignment must mention which Pn a given DP bears a given theta relation to (as per footnote 9).

Malagasy maintains logical expressive power by providing a way of forming P1s. (And there are more). Keenan (to appear) discusses other cases of relativization into sentence complements.

## 6. Conclusion

First, we reaffirm the merits of the approach listed at the beginning of this article. And second, we have supported more generally that “core” Malagasy syntax, and implicitly that of Western Austronesian languages more generally, concerns the derivation and interpretation of one place predicates rather than DP movement.

## References

- Aldridge, Edith. 2002. Nominalization and *wh*-movement in Seediq and Tagalog. *Language and Linguistics* 3.2:393-426.
- Andrianierenana, Clément Luc. 1996. Morphological causatives in Malagasy. *The Structure of Malagasy*, Vol. 1, ed. by Matthew Pearson & Ileana Paul, 58-76. UCLA Occasional Papers in Linguistics 17. Los Angeles: Department of Linguistics, University of California.
- Baker, Mark C. 1988. *Incorporation: A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.
- Baker, Mark C. 1996. *The Polysynthesis Parameter*. New York: Oxford University Press.
- Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.
- Chomsky, Noam. 1996. *The Minimalist Program*. Cambridge: MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. *Ken Hale: A Life in Language*, ed. by Michael Kenstowicz, 1-52. Cambridge: MIT Press.
- Collins, Chris. 2005. A smuggling approach to the passive in English. *Syntax* 8.2:81-120.
- Corbett, Greville G. 1991. *Gender*. Cambridge & New York: Cambridge University Press.
- Guilfoyle, Eithne, Henrietta Hung, and Lisa Travis. 1992. Spec of IP and Spec of VP: two subjects in Austronesian languages. *Natural Language and Linguistic Theory* 10.3: 375-414.
- Huang, Lillian M. 2002. Nominalization in Mayrinax Atayal. *Language and Linguistics* 3.2:197-225.
- Hyams, Nina, Dimitris Ntelitheos, and Cecile Manorohanta. 2006. Acquisition of the Malagasy voicing system: implications for the adult grammar. *Natural Language and Linguistic Theory* 24.4:1049-1092.

- Keenan, Edward L. 1972. Relative clause formation in Malagasy (and some related some not so related languages). *The Chicago Which Hunt*, ed. by Paul M. Peranteau, Judith N. Levi & Gloria C. Phares, 169-189. Chicago: Chicago Linguistic Society.
- Keenan, Edward L. 1976. Remarkable subjects in Malagasy. *Subject and Topic*, ed. by Charles N. Li, 247-301. New York: Academic Press.
- Keenan, Edward L. 1995. Predicate-argument structures in Malagasy. *Grammatical Relations: Theoretical Approaches to Empirical Questions*, ed. by Clifford S. Burgess, Katarzyna Dziwirek & Donna B. Gerds, 171-216. Stanford: CSLI.
- Keenan, Edward L. (to appear). Voice and relativization without movement in Malagasy. *Natural Language and Linguistic Theory* 26.
- Keenan, Edward L., and Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8.1:63-99.
- Keenan, Edward L., and Cecile Manorohanta. 2001. A quantitative study of voice in Malagasy. *Oceanic Linguistics* 40.1:67-84.
- Keenan, Edward L., and Maria Polinsky. 1998. Malagasy morphology. *The Handbook of Morphology*, ed. by Andrew Spencer & Arnold M. Zwicky, 563-623. Oxford: Blackwell.
- Keenan, Edward L., and Baholisoa Ralalaoherivony. 2000. Raising from NP in Malagasy. *Linguistic Investigations* 23.1:1-44.
- Keenan, Edward L., and Jean-Paulin Razafimamonjy. 2004. Reciprocals in Malagasy. *Oceanic Linguistics* 43.1:177-207.
- Kroeger, Paul R. 1988. Verbal focus in Kimaragang. *Papers in Western Austronesian Linguistics No.3*, 217-240. Pacific Linguistics A-78. Canberra: The Australian National University.
- Kroeger, Paul R. 1993. *Phrase Structure and Grammatical Relations in Tagalog*. Stanford: CSLI.
- Mahajan, Anoop. 1995. Active passives. *Proceedings of the 13<sup>th</sup> West Coast Conference on Formal Linguistics (WCCFL-13)*, ed. by Raul Aranovich, William Byrne, Susanne Preuss & Martha Senturia, 286-301. Stanford: CSLI.
- Paul, Ileana. 1999. *Malagasy Clause Structure*. Montreal: McGill University dissertation.
- Paul, Ileana. 2001. Concealed pseudo-clefts. *Lingua* 111.10:707-727.
- Paul, Ileana, and Lucie Rabaovololona. 1998. Raising to object in Malagasy. *The Structure of Malagasy*, Vol. 2, ed. by Ileana Paul, 50-65. UCLA Occasional Papers in Linguistics 20. Los Angeles: Department of Linguistics, University of California.
- Pearson, Matthew. 2004. Voice morphology, case and argument structure in Malagasy. *ZAS Papers in Linguistics* 34:229-243.
- Pearson, Matthew. 2005. The Malagasy subject/topic as an A'-element. *Natural Language and Linguistic Theory* 23.2:381-457.

- Phillips, Vivianne. 2000. The interaction between prefix and root: the case of *maha-* in Malagasy. *Formal Issues in Austronesian Linguistics*, ed. by Ileana Paul, Vivianne Phillips & Lisa Travis, 85-104. Dordrecht & Boston: Kluwer.
- Potsdam, Eric. 2004. Wh-questions in Malagasy. *ZAS Papers in Linguistics* 34:244-258.
- Prentice, David J. 1971. *The Murut Languages of Sabah*. Pacific Linguistics C-18. Canberra: The Australian National University.
- Rahajarizafy, Antoine. 1960. *Essai sur la grammaire malgache*. Antanimena-Tananarive: Impr. catholique.
- Rajemisa-Raolison, Regis. 1971. *Grammaire malgache*. Fianarantsoa: Ambozontany Fianarantsoa.
- Randriamasimanana, Charles. 1986. *The Causatives of Malagasy*. Oceanic Linguistics Special Publication 21. Honolulu: University of Hawaii Press.
- Rasolofo, Andoveloniaina. 2006. *Malagasy Transitive Clause Types and their Functions*. Eugene: University of Oregon dissertation.
- Richards, Norvin. 2000. Another look at Tagalog subjects. *Formal Issues in Austronesian Linguistics*, ed. by Ileana Paul, Vivianne Phillips & Lisa Travis, 105-116. Dordrecht & Boston: Kluwer.
- Sabel, Joachim. 2003. Malagasy as an optional multiple *wh*-fronting language. *Multiple wh-Fronting*, ed. by Cedric Boeckx & Kleantes K. Grohmann, 229-254. Amsterdam & Philadelphia: John Benjamins.
- Schachter, Paul. 1976. The subject in Philippine languages: topic, actor, actor-topic, or none of the above. *Subject and Topic*, ed. by Charles N. Li, 491-518. New York: Academic Press.
- Schachter, Paul, and Fe T. Otones. 1972. *Tagalog Reference Grammar*. Berkeley: University of California Press.
- Travis, Lisa. 2000. The L-syntax/S-syntax boundary: evidence from Austronesian. *Formal Issues in Austronesian Linguistics*, ed. by Ileana Paul, Vivianne Phillips & Lisa Travis, 167-194. Dordrecht & Boston: Kluwer.
- Travis, Lisa. 2005. Passives, states and roots and Malagasy. *UCLA Working Papers in Linguistics* 12:395-409.
- Travis, Lisa. 2006. Voice morphology in Malagasy as Clitic Left Dislocation or Through the Looking Glass: Malagasy in Wonderland. *Clause Structure and Adjuncts in Austronesian Languages*, ed. by Hans-Martin Gärtner, Paul Law & Joachim Sabel, 281-318. Berlin & New York: Mouton de Gruyter.
- Wechsler, Stephen, and I Wayan Arka. 1998. Syntactic ergativity in Balinese: an argument structure based theory. *Natural Language and Linguistic Theory* 16.2: 387-441.

[Received 13 November 2007; revised 17 July 2008; accepted 31 August 2008]

Department of Linguistics  
University of California, Los Angeles  
3125 Campbell Hall  
Los Angeles, CA 90095-1543  
USA  
keenan@humnet.ucla.edu

## 馬拉加斯語關係子句之詞組結構性

Edward L. Keenan

美國加州大學洛杉磯分校

通行於馬達加斯加島的馬拉加斯語是一個具有典型語態系統的西部南島語。本文首先說明如何衍生各語態中從動詞發展而成的核心句，以及如何經由語意的組合來理解這些語句。我們認為這些語句是直接經由動詞詞綴投射出來，不需經過論元移位。論旨角色的指派可以清楚地在語意組合的過程中完成。其次，我們將此分析應用到關係子句，主張關係子句的衍生也一樣不需要經由論元移位。和大多數主流的分析相比，我們的分析有下列優勢：

1. 經由清楚的語意組合來解釋關係子句的形成，能（部分地）說明說話者是如何理解新語句。
2. 對馬拉加斯語中的“唯有主語可以移出的限制”提出一個全新的衍生方法，能確實遵守 Chomsky (2001:2) 所提的內包原則，也就是不使用語跡 (traces) 和指標 (indices)。
3. 簡明且易學：衍生核心句只需要約二十個動詞詞綴，並在遵守鄰近原則的屬性查核下即可完成。論旨角色的指派可以很清楚地在詞根加綴所形成的詞類中達成。
4. 像英語這種缺乏語態的語言裡，需要一個認知上的觸發詞 (trigger) 來解釋約束變項運符的移出。這也就是說，英語和馬拉加斯語的差別在於構詞特性的不同 (Chomsky 1996:7)。

關鍵詞：馬拉加斯語，關係子句，語態