Extraction without Movement: is Malagasy a Perfect Language?[†]
Edward L. Keenan, UCLA 2005

abstract I propose "minimal" solutions to two basic problems in Malagasy syntax: the Voice Problem, and the "Extraction" Problem. As is well known, both are prominent in the syntactic analysis of W. Austronesian languages[1] generally. I use the Bare Grammar format of Keenan & Stabler (2003) and contrast it with more mainstream ones: Guilfoyle, Hung and Travis [GHT 1992], Paul (1999), Pearson (2001, 2005).

Voice: Malagasy presents morphologically distinct verbs built from the same root which assign different grammatical cases to DPs with given theta roles, yielding Ss that are theta equivalent, and, with appropriate choice of DPs, logically equivalent, much like active and agented passive Ss in English. The problem is to derive and interpret such Ss so as to yield these judgments of semantic equivalence as theorems. Our solution, which is purely structural, invoking no notion of 'subject', 'topic', 'pivot', 'trigger', etc., is simply an explicit syntactic and semantic interpretation of voice affixes. Deriving Ss built from verbs in different voices involves no A movement of DPs. Voice morphology determines the distinctive syntactic and semantic properties of nuclear clauses. This supports that "Variation of language is essentially morphological ..." (Chomsky 1995:7).

"*Extractions*" as in relative clauses (RCs), are limited to a single structurally definable DP per clause, the DP whose theta role co-varies with the voice morphology. The problem is to derive this fact without taking it as an axiom (a stipulation). But, I claim, this "problem" is a non-problem, the "fact" a non-fact. RCs in Malagasy involve no extraction or variable binding, they are built just from the constituents needed to derive and interpret nuclear clauses. Again our solution is Minimalist: "any structure formed by the computation ... is constituted of elements already present in the lexical items...; no new objects are added in the course of computation ..." (Chomsky 1995:228). We show that Malagasy's rich voice morphology obviates the need for extraction. In English by contrast RCs normally lack sufficient overt structure to support a correct compositional interpretation. Wh-movement and binding is a minimal logical repair which enables Compositionality to be satisfied. The trigger for taking the Movement option is the need to semantically interpret strings that are not interpreted in nuclear Ss. Indeed they are not even constituents of nuclear Ss. We also illustrate the sense in which Malagasy conforms to classical extraction constraints despite lacking extraction.

1. Background Verbs denote relations whose arguments bear *theta* relations such as *Agent*, *Theme*, *Goal*,... to them1. For example in *John chased Bill*, *Bill* bears the Theme relation to *chase*, meaning that the object *Bill* denotes is affected by the chasing action. When A is an argument of a predicate P we write $\theta(A)$ for the set of theta roles that A bears to P. The Theta Criterion (satisfied but not adopted here) implies that $\theta(A)$ is a unit set. And, where *nuclear* Ss are ones built from a single lexical predicate, we define:

Def   Nuclear Ss P and Q are *theta equivalent* iff there is a bijection π from the theta DPs of P to those of Q satisfying $\theta(A) = \theta(\pi A)$, all theta DPs A of P.

(1a,b) are theta equivalent: the bijection π mapping *John* to *Sue* and *the door* to *the window* preserves theta role, Agent and Theme respectively:

(1)   a.   John opened the door   b. The window was opened by Sue

Similarly the nuclear Ss in (2a,b) in Malagasy are theta equivalent, in fact logically equivalent (The identity map between theta DPs is the required bijection).

(2)   a.   [n+aN+enjika (nanenjika) ny jiolahy] Rabe
                past+AF+chase   the thief   Rabe
                Rabe chased the thief

        b.   [n+enjika+ina+Rabe (nenjehin-dRabe)]   ny jiolahy
                past+chase+TF+Rabe.gen   the thief
                Rabe chased the thief (*The thief was chased by Rabe*)

notation Verbs are given in morphemic decomposition followed by their standard orthography1. Square brackets mark the major constituent, for which there is massive evidence (not repeated here) and total agreement among generative grammarians (Keenan 1972, 1976, 1995; [GHT]; Paul 1999; Rackowski & Travis [R&T] 2000; Pearson 2001, 2005). Theory neutrally we call it a P1, *one place predicate phrase*. The theta role of its DP sister is determined by the verb (root+voice affix, *aN-*, *-ina*, etc). If that role is highest on the Theta Hierarchy in (3) of those the verb predicates of its arguments then the P1, its verb and voice affix will be called *Actor Focus* (AF), (adapting Schachter 1976), otherwise they are called τ Focus, τ the theta role of that argument.

(3)   Agent > Experiencer > Theme/Goal > Instrument/Benefactive/Locative > Other

So the P1 in (2a) is AF, that in (2b) TF (Theme Focus). And unsurprisingly many linguists (Rahajarizafy 1960; Keenan op cit; GHT, and Paul 1999) have treated the AF/TF distinction in Malagasy as an Active/Passive one. In (1a) and (2a) for example the Theme forms a constituent with the verb to the exclusion of the Agent, whereas in (1b) and (2b), the Agent is within the predicate constituent to the exclusion of the Theme. Sometimes we translate TF Ss as passives to remind the reader of this alternation. Still, *-ina* suffixed TF verbs behave most like English active transitive verbs2.

A second, important, difference between (2a) and (2b) lies in case assignment to the DPs. These differences are most clearly seen in the pronoun system, which we present below using traditional names as an aid to the reader (though *case 1*, *case 2* and *case 3* would be preferable). A DPα has the same case, α, as its pronominal replacement. See K&P (Keenan & Polinsky 1998).

Case	1sg	2sg	3sg	1pl.excl	1pl.incl	2pl	3pl
nom	aho	ianao	izy	izahay	isika	ianareo	izy2
acc	ahy	anao	azy	anay	antsika	anareo	azy
gen	-ko	-nao	-ny	-nay	-tsika	-nareo	-ny

nominative is the case of the argument of all P1s, including predicate adjectives, PPs and Ns. It functions as a default (Pearson 2005) and is only distinctively marked in the pronouns. It is never selected by Vs, Ns or As in forming P1s, NPs or APs. Only two Preps select **nom** (*noho* 'compared to', because of"), and *afa-tsy* 'except', lit: "free-not".

accusative is the case of the complement of transitive AF verbs. It marks names, some kin terms and, optionally, demonstratives with *an-*. It is selected by many Vs, some Ns (8), a few Ps (a) and a few As (b). DP$_{acc}$s also serve as predicate possessives: *Ahy io* "That is mine", *an-dRabe io* 'That is Rabe's'

genitive is the case of Possessors, internal Agents of Vs and As (Phillips 2000) and most objects of Prepositions. DP$_{gen}$s, pronominal and full, are morphologically bound to their heads (the latter by a process we call *n-bonding*). Pronouns may replace full DPs preserving grammaticality and except monosyllabic *-ko* and *-ny* may carry stress (phonemic). Here are (2a,b) with the DPs replaced by the appropriate pronouns.

(4) a. nanenjika azy izy
         past+AF+chase 3sg.acc 3nom
         He/She chased him/her

     b. n+enjika+ina+-ny (nenjehiny) izy
         past+chase+TF+3gen   3nom
         He/She was chased by him/her

In a text count based on two newspaper articles and selections from three novels, in Malagasy and in English, Keenan (1995) found case distributed as in (5)2. Genitive is the most widely occurring case in Malagasy, accusative the most restricted. Just the opposite pattern obtained for English. Nominative was intermediate in both languages.

(5)	number of DPs	**nom**	**acc**	**gen**
Malagasy	1,237	33.6%	23.0%	43.4%
English	1,345	38.9%	47.1%	14.1%

The Predication Parameter Ls vary according as the dominant expression of the Predicate-Argument relation is N+Possessor or V+Object. Malagasy has a nominal setting, (N+Poss), English a verbal one (V+Obj).

2. The Voice Problem Primary voice affixes combine with roots to build verbs with diverse subcategorizations. The verbs build Ss with diverse phrase structures. We show that the subcategorization of verbs depends *both* on the affix and the root. Some affixes combine with different roots to yield verbs of different valencies. And different affixes may combine with the same root to yield verbs with different subcategorizations. We illustrate with *enjika* 'chase' and then generalize.

We follow *Bare Grammar* (BG) in treating expressions as (string, category) pairs

(s,C), sometimes noted [Cs].   A grammar G is a list LexG of expressions and a set RuleG of functions mapping sequences of expressions to expressions. The language L(G) *generated* by G is the closure of LexG under RuleG, that is, the set of expressions derivable from the lexicon by applying the rules finitely many times. Grammatical categories are language specific, coding the role of classes of expressions in the definition of the rules. A function in RuleG whose values are built by affixing, internal vowel change or reduplication is called *morphological*, otherwise it is *syntactic*. Morphological functions are of necessity language particular, as they mention language specific morphemes. But we expect (but do not try to justify here) that purely syntactic functions are *uniformly definable* across languages, that is, universal2.

2.1 aN-   *Enjika* is a root with two associated theta roles, Agent and Theme. We note its category as RT{AG,TH}. In (2a) *aN-* prefixes to *enjika* to form an AF transitive verb2, one which combines first with a DPacc to form a P1, which in turn combines with a DPnom to form an S. *aN-* prefixation is effected by a morphological function AN mapping roots to verbs, as in (6). We list leftmost in the subcategorization the category, DPacc, the verb combines with first. So DPnom is the category of the argument of the derived P1.

$$AN$$
(6)   enjika: RT{AG,TH}   ===>   -anenjika: S/[DPacc, DPnom]

Syntax *Merge* concatenates a string s of category X/[$C_1\alpha_1$,...,$C_n\alpha_n$] with a t of category $C_1\alpha_1$ to form s⌢t of category X/[$C_2\alpha_2$,...,$C_n\alpha_n$], satisfying a category (C1) and case (α_1) requirement of X/[$C_1\alpha_1$,...,$C_n\alpha_n$]. So Merge concatenates *-anenjika* above with a DPacc to yield a S/DPnom (that is, a P1). Then it concatenates that P1 with a DPnom to form an S. The derivation of (2a) is (7). (We insert tense marking early and ignore it).

(7)   Spst
                              Spst/DPnom

                  Spst/[DPacc,DPnom]
                                                                DPacc   DPnom
                  Tpst   S/[DPacc,DPnom]

                              RT{AG,TH}

            n   aN + enjika   ny jiolahy   Rabe

The use of 'S/...' above is not redundant. Transitive nouns like 'fear' and 'compassion' may take two case marked DP arguments to form a DP, as in (8) from a romance novel.

(8)   DP

   Det   NP

              NP/DPacc

     NP/[DPgen, DPacc]   DPgen   DPacc

     ny   alahelo   n-dRazay   an-drai+ny
     the compassion   gen-Razay   acc-father+her
     Razay's compassion for her father
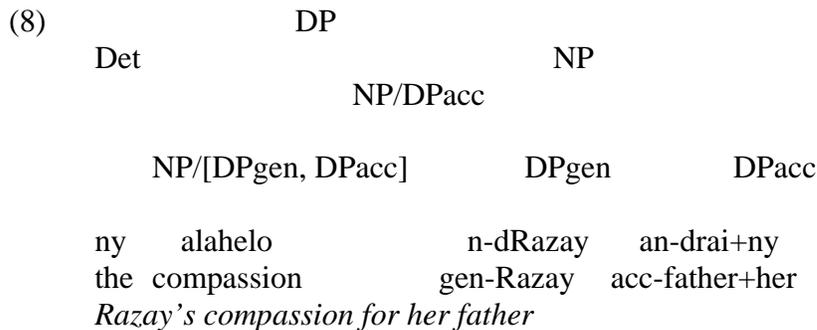
We use traditional trees, as in (7) and (8), to merely record the derivational history of the expressions – the successive application of functions to their arguments.

Semantics We interpret the root *enjika* as a binary relation ENJIKA, with theta roles Theme and Agent. The interpretation of the verb AN(enjika) is given by (9), which says that the pair (x,y) is in the ENJIKA (CHASE) relation, y is its THEME and x its AGENT.

(We write $[\![d]\!]$ for the semantic interpretation of an expression d in an arbitrary model).

(9) $[\![$AN(enjika, RT{AG,TH})$]\!]$(y)(x) = True iff

       (x,y) $\in$ ENJIKA $\wedge$ THEME(y,ENJIKA) $\wedge$ AGENT(x,ENJIKA)

We often omit the category coordinate of an expression, writing just *enjika* above instead of <enjika, RT{AG,TH}>. And from (9) we see that it is the verb+affix, *-anenjika,* not just the root or the affix alone which determines the match of argument and theta role.

    Now, to define the function AN we must define its domain, the set of expressions it applies to, and state its value at each element in that domain. We stipulate its domain to be a listed set of roots, as in Abinal & Malzac (1963). The roots r usually have category RT{AG,TH} and are mapped to AN(r) of category S/[DPacc,DPnom]. Further, almost all roots r of category RT{AG,TH,GOAL} are mapped to AN(r) of category S/[DPacc,DPacc,DPnom]2. Some present tense examples of AN applying to roots of category RT{AG,TH} are: AN(*didy*) = *mandidy* 'cuts', AN(*vaky*) = *mamaky* 'reads', AN(*vono*) = *mamono* 'kills'. In the second case we have: AN(*roso*) = *mandroso* 'serves', AN(*tolotra*) = *manolotra* 'offers', AN(*ome*) = *manome* 'gives'.

But AN also applies2 to a few roots of category RT{AG} to yield a P1 (S/DPnom): AN(*leha*) = *mandeha* 'goes', AN(*dihy*) = *mandihy* 'dances', AN(*lainga*) = *mandainga* 'tell lies', AN(*lohalika*) = *mandohalika* 'kneels'. And AN applies to some roots of category RT{TH}: AN(*hetaheta*) = *mangetaheta* 'is thirsty', AN(*hatsiaka*) = *mangatsiaka* 'is cold', AN(*firatra+firatra*) = *mamirapiratra* 'scintillates'. (10) is the kind of generalization that aids acquisition by providing default interpretations for novel verbs.

(10) *aN-* is an Actor Focus prefix.

To complete defining AN we must say what phonological value it assigns to the string coordinate of the root. We often just write *aN+root*, but as in other Austronesian languages (*m*)*aN-* prefixation is phonologically complex: vowel initial roots just prefix *an-*. If the root begins with a voiceless consonant it (usually) drops; a voiced continuant usually mutates to the "closest"prenasalized non-continuant: AN(roso) = *mandroso*, but sometimes drops: AN(*vono*) = *mamono*. See K&P and Paul (1996), etc.

2.2 i- A second morphological function building verbs from roots is I, which prefixes *i-* to the string coordinate of a root. It has many roots of category RT{AG,TH} in its domain, and like AN, builds AF transitive verbs of category S/[DPacc, DPnom]: I(*kapoka*) = *mikapoka* 'beats', I(*vidy*) = *mividy* 'buys', I(*orina*) = *miorina* 'builds', I(*varotra*) = *mivarotra* 'sells', I(*laza*) = *milaza* 'says', I(*tady*) = *mitady* 'seeks', I(*fidy*) = *mifidy* 'chooses', I(*saotra*) = *misaotra* 'thanks', I(*karakara*) = *mikarakara* 'takes care of', I(*taona*) = *mitaona* 'carries', I(*jery*) = *mijery* 'watches', I(*antso*) = *miantso* 'calls'.

But more often I- builds intransitive verbs, S/DPnoms. It applies to some roots of category RT{AG}: *asa* ==> *miasa* 'works', *teny* ==> *miteny* 'speaks', *tomany* ==> *mitomany* 'cries'. It also applies to some two place roots in the domain of AN: *enjika* ==> *mienjika* 'flees', *manenjika* 'chases'; *sasa* ==> *misasa* 'washes (intr), *manasa* 'washes (tr)'; *hidy* ==> *mihidy* 'is locked', *manidy* 'locks (tr)'. I- is the most productive verb forming function. In distinction to AN it enables direct recovery of the root. Hilda Koopman (pc) notes nonce *mi-* verbs like *miparticiper* 'participates' built on French. And we have another language internal generalization which aids in learning novel verbs:

(11) When *i-* and *aN-* apply to the same root the *i*-verb usually has lesser valence.

A BG style grammar requires extensive listing of just what roots are in the domains of the morphological functions on which the core grammar of Malagasy rides. Much of this work has been done: Abinal and Malzac 1987 [1888], Rajaonarimanana 1995 and Rabenilaina 1993. Below are some further affixes important in our analysis to follow2.

2.3 -ina is suffixed to roots such as *enjika* by a voice function INA forming the transitive TF verb *enjehina* in (2b), one that combines first with a DPgen, interpreted as Agent, and then with a DPnom, interpreted as Theme. (12) shows the syntactic action of this suffix, (13) derives a nuclear S built from the INA-verb, and (14) is its semantic interpretation.

$$\text{INA}$$
(12) enjika: RT{AG,TH}   ==>   enjehina: S[DPgen, DPnom]

(13)   Spst

                    Spst[DPnom]

        Spst[DPgen,DPnom]

                                      DPgen   DPnom
              S[DPgen,DPnom]

    Tpst
        RT{AG,TH}

    n   enjika   +   ina   Rabe   ny jiolahy
    nenjehin-dRabe ny jiolahy (= 2b)

(14) ⟦INA(enjika)⟧(u)(v) = True iff (u,v) ∈ ENJIKA ∧ THEME(v,ENJIKA) ∧
                                    AGENT(u,ENJIKA)

So *enjehina* is a TF verb, the kind INA typically builds. Notice that (15a) follows from (9) and (14), and in general (15b) holds when *root* is in the domain of both AN and INA. The derived verbs are logical converses, from which the two propositions below follow.

(15) a. ⟦AN(enjika)⟧(y)(x) = ⟦INA(enjika)⟧(x)(y) and

     b. ⟦AN(root)⟧(y)(x) = ⟦INA(root)⟧(x)(y) and

Proposition 1 Nuclear Ss built from AN(r) and INA(r) are theta equivalent: the P1 argument of each bears the same theta role as the internal argument of the other2.

Proposition 2 Nuclear Ss built from AN(r) and INA(r) are logically equivalent (true in the same conditions) when their Agents are the same individual denoting DPs and their Themes are the same individual denoting DPs.

Let us emphasize that in building Ss like (7) and (13) with AF and TF predicates respectively no DP movement was invoked in either case.

    Now, unlike AN or I, INA takes two types of verbs in its domain. First there are 6 AF *aN-* verbs which suffix *-ina* directly to form a TF verb: *hataka* ==> *mangataka* (AF) ==> *angatahina* (TF) 'asks'; *halatra* ==> *mangalatra* (AF) ==> *angalarina* (TF) 'steals'; *voatra* ==> *mamboatra* (AF) ==> *amboarina* (TF) 'arranges'. See RR (1971; 102).

(16) n+[[aN+halatra]+ina]+Rabe (nangalarin-dRabe) ny omby
      past[[AF+steal]+TF+Rabe.gen the cow
      The cow was stolen by Rabe

Second, quite frequent, INA suffixes -*ina* to causatives, themselves formed by prefixing *amp-* to an AF verb forming an AF verb:

(17) a. n+amp+aN+taov (nampanao) farafara azy ireo Rabe
   Past+Cause+AF+make bed 3.acc dem.pl Rabe
   Rabe made them make beds

 b. n+[amp+aN+taov]+ina+Rabe (ampanaovin-dRabe) farafara izy ireo
  Past+[[Cause+AF+make]+TF]+Rabe bed 3.nom dem.pl
  They were made by Rabe to make beds

 c. [m+amp+aN+sasa (mampanasa) lamba azy] Rasoa
   pres+Cause+AF+wash clothes him Rasoa
   Rasoa is making him wash clothes

 d. [∅+amp+aN+sasa+ina+Rasoa (ampanasain-dRasoa) lamba] izy
   pres+[[Cause+AF+wash]+TF]+Rasoa.gen clothes he
  He is being made by Rasoa to wash clothes

amp- only prefixes to AF verbs, never roots or verbs in other voices. In the causative of a transitive verb either accusative can be the argument of the INA derived P1. AF verbs built from the functions I, IF (reciprocal), AHA (potential, cause) are not in Dom(INA).

2.3 a- is primary affix, prefixed by a function A which applies only to roots. The argument of a P1 it builds is an "Intermediary": the Theme with roots of transmission, the Instrument otherwise. In the first case -*Vna* externalizes the Goal. See Paul (1999).

(18) a. [n+aN+tosika (nanosika) ny fiara] Rabe   (AF)
   past+AF+push the car Rabe
   Rabe pushed the car

 b. [n+a+tosika+Rabe (natosi-dRabe)] ny fiara (TF)
   past+TF+push+Rabe.gen the car
   Rabe pushed the car

tosika is not in the domain of INA and so does not accept a -*Vna* suffix[2]. However ditransitive roots with both Goal and Theme roles, (19), often have both *a-* and -*Vna* forms. The *a-* verb assigns Theme to the P1 argument and the -*Vna* one assigns it Goal.

(19) a. n+aN+tolotra (nanolotra) boky azy aho
   past+AF+offer   book 3acc 1nomI
   I offered a book to him

 b. no+tolotra+(a)na+ko (notolorako) boky izy
   past+offer+GF+1sg.gen   book 3nom
   He was offered a book by me

c.   n+a+tolotra+ko (natolotro) azy ny boky
     past+TF+offer+1sg.gen   3acc the book
     The book was offered to him by me

Here are the values of AN, A, and INA at ditransitive roots of the form (r, RT{AG,TH,GOAL}):

(20) a. AN(r, RT{AG,TH,GOAL})(z)(y)(x) = True iff

$$\langle z,y,x \rangle \in P \ \& \ THEME(z,P) \ \& \ GOAL(y,P) \ \& \ AG(x,P)$$

   b. A(r, RT{AG,TH,GOAL})(x)(y)(z) = True iff

$$\langle z,y,x \rangle \in P \ \& \ AG(x,P) \ \& \ GOAL(y,P) \ \& \ TH(z,P)$$

   c. INA(r, RT{AG,TH,GOAL})(x)(z)(y) = True iff

$$\langle z,y,x \rangle \in P \ \& \ AG(x,P) \ \& \ TH(z,P) \ \& \ GOAL(y,P)$$

2.4 Secondary verbal affixes build (tenseless) verbs from (tenseless) verbs, combining thus with items which already have (at least) one verbal affix. Of the four major such affixes, *if-* 'Reciprocal', *amp-* 'Causative' and *aha-* 'Cause/Potential' are semantically specific and derive AF verbs from AF verbs (see K&R 2004; Phillips 2000).

(21)   m+if+aN+enjika (mifanenjika) Rabe sy Rakoto
       pres+Rec+AF+chase Rabe and Rakoto
       Rabe and Rakoto are chasing each other

2.5.1 -ana is suffixed by a function ANA to all AF verbs built from AN, I, AMP, AHA, IF and a few other AF building functions not mentioned here. No root is in the domain of ANA. The verbs it forms are non-AF ones, called *Circumstantial Focus* (CF) since the argument of the P1s it forms usually denotes some circumstance of the action or state expressed by the verb[2]. Compare the TF (22b) and the CF (22c). In the latter the Means constituent is the nominative P1 argument and the Agent is the genitive complement, as with TF verbs.

(22) a.   [n+aN+enjika (nanenjika) ny jiolahy t+amin'ny fiara] Rabe
          past+AF+chase   the thief past+with'the car Rabe
          Rabe chased the thief by means of the car

   b.   [n+enjika+ina+Rabe (nenjehin-dRabe) t+amin'ny fiara] ny jiolahy
        past+chase+TF+Rabe.gen   past+with'the car   the thief
        The thief was chased by Rabe by means of the car
        (*Rabe chased the thief by means of the car*)

   c.   [n+[aN+enjika]+ana+Rabe (nanenjehan-dRabe)   ny jiolahy] ny fiara
        past+[[AF+chase]+CF]+Rabe.gen   the thief   the car

The car was used by Rabe to chase the thief

Phonologically, suffixation (as with *-ina* or *-ana*) shifts stress to the right and may trigger an epenthetic consonant[2]. From *enjika* we derive CF verbs such as those on the right in (23) (present tense, primary stress marked):

(23)   miénjika   runs, flees   ienjéhana   circumstance of fleeing
       manénjika   chases   anenjéhana   circ. of chasing
       mifanénjika   chase each other   ifanenjéhana   circ. of chasing each other
       mampanénjika   makes chase   ampanenjéhana   circ. of causing to chase

-ana generally adds an argument for which its AF verb is not subcategorized (Paul 1999, Keenan 2000). It bears some oblique relation (Benefactive, Instrument, Location, Time, Manner, Price, Material,...) to the AF verb it is built from, but we don't know which one. (24) is an illustrative application of ANA to a typical transitive AF verb.

ANA
(24) (verb$_{AF}$: S/[DP$_{acc}$, DP$_{nom}$]) ==> (verb+(C)ana: S/[DP$_{gen}$, DP$_{acc}$, $_{Dpnom}$])

Semantically, oblique DPs and adverbials are predicate modifiers – functions F mapping predicate denotations P to predicate denotations F(P). Normally they are *restricting*: F(P) $\leq$ P, meaning that for all entities x, if F(P)(x) = True then P(x) = True. *for Max* is restricting in (25a), whence (25a) entails (25b). Replacing *for Max* by *with a screwdriver* preserves the entailment since it is also restricting.

(25) a.   Sue opened the box for Max   (**for Max**)(**open** (**the box**))(**Sue**)
     b.   Sue opened the box   (**open**(**the box**))(**Sue**)

So Prepositions denote functions **p** mapping entities to restricting modifiers and CF verbs quantify over such functions, as in (26), the circumstantial of a transitive verb.

(26)   ⟦ANA(verb$_{AF}$)⟧(x)(y)(z) = True iff ∃**p** **p**(z)(⟦verb$_{AF}$⟧(y))(x) = True

The semantic interpretation of the CF verb uses the interpretation of the AF verb it takes as argument. Since **p**(x) is restricting, a nuclear circumstantial S like (27a) entails the corresponding AF one (27b) which lacks an oblique DP corresponding to the P1 argument of the circumstantial. (27a) is derived in (28) and interpreted in (29).

(27) a.   ∅+aN+taov+ana+Rabe (anaovan-dRabe) ny trano Rasoa
          pres+aN+make+ana+Rabe.gen the house Rasoa
          Rabe is making the house for / because of Rasoa

     b.   m+aN+taov (manao) ny trano Rabe
          pres+aN+make the house Rabe
          Rabe is making a house

(28)

$$S_{pres}$$

$$S_{pres}/[DP_{nom}]$$

$$S_{pres}/[DP_{acc}, DP_{nom}]$$

$$S_{pres}/[DP_{gen}, DP_{acc}, DP_{nom}]$$

$$S/[DP_{gen}, DP_{acc}, DP_{nom}]$$

$$S/[DP_{acc}, DP_{nom}]$$

		$RT_{\{AG,TH\}}$		DP_{gen}	DP_{acc}	DP_{nom}
⊘	aN	taov	ana	Rabe	ny trano	Rasoa

(29)	⊘	aN	taov	ana	Rabe	ny trano	Rasoa
	pres	AN	TAOV	ANA	r	t	s

AN(TAOV)

ANA(AN(TAOV)

(ANA(AN(TAOV)))(r)

(ANA(AN(TAOV)))(r)(t)

(ANA(AN(TAOV)))(r)(t)(s)

The last line is True iff for some **p**, **p**(s)AN(TAOV)(t))(r) = True. And since **p**(s) is restricting, the last line entails AN(TAOV)(t)(r), the interpretation of *Rabe is making the house*. Thus our semantics for the circumstantial S entails that Rabe is Agent of *make* and the house is Theme. Consider now the predictions of theta equivalence between a circumstantial S and a corresponding AF one with an overt oblique.

(30) a. ⊘+aN+taov+ana+Rabe (anaovan-dRabe) ny trano Rasoa   (CF)
pres+aN+make+ana+Rabe.gen the house Rasoa
Rabe is making the house for / because of Rasoa

   b. m+aN+taov (manao) ny trano ho an-dRasoa Rabe   (AF)
pres+aN+make the house for Rasoa Rabe
Rabe is making the house for Rasoa

Our semantic analysis shows that the genitive complement *Rabe* in (30a), has the same theta role, Agent, as the P1 argument in (30b). Also the accusative *ny trano* 'the house' has the same theta role, Theme, in both Ss. But while *Rasoa* is a Benefactive in (30b), forced by the choice of Prep, *ho an-* 'for', *Rasoa* is not assigned a theta role on our analysis in (30a). Rather our semantics merely quantifies over the set in which Preps denote, different theta roles for Rasoa correspond to different choices of possible Prep denotations. Most choices – Instrument, Location, Manner – are pragmatically implausible in (30a). But *Rasoa* could still be a Benefactee or a Cause. Thus our analysis predicts that (30b) entails (30a) but not conversely.

On our analysis the P1 argument of a CF S is *vague*, not ambiguous, with respect to theta role. We just know that the type of relation it bears to the predicate is the type expressible by Prepositions. GHT/Paul view it as derived by movement from structures in which it is governed by different prepositions, whence the CF S is semantically ambiguous. But in fact it is merely vague, not ambiguous, like (31a) in English.

(31) a. the bridge where Martin saw an angel
     b. the bridge above/under/beside which Martin saw an angel

(31a) simply doesn't specify the precise location of the angel but it is not felt as ambiguous according as the angel was above, under or beside the bridge, just as *There's a student in your office* is not ambiguous according as the student is Greek or not. Circumstantial Ss in Malagasy are typically non-specific, not ambiguous, though proponents of UTAH are committed to the ambiguity option.,

3. "Extractions" Relative Clauses (RCs) in Malagasy (and W. Austronesian) are formed simply by concatenating a noun with a P1 in any voice: $\mathbf{N} = \mathbf{N} + \mathbf{P1}$[2.]

(32) a. zaza [P1 nanenjika ny jiolahy]   b. zaza [P1 nenjehin-dRabe]
        child (who) [chased the thief]   child [chased-by-Rabe]

P1s denote *properties*, functions from individuals to truth values. The property denoted by *nanenjika ny jiolahy* maps an individual x to True iff x chased the thief. That denoted by *nenjehin-dRabe* maps x to True iff Rabe chased x. In RCs P1s function as property denoting expressions. In (32a) the children referred to are those that *nanenjika ny jiolahy* 'chased the thief' is True of. In (32b) it is those *nenjehin-dRabe* 'was chased by Rabe' maps to True. Formally RCs are interpreted in (33a). (33b) is an example.

(33) a. ⟦Noun + P1⟧ = ⟦Noun⟧ ∩ {x| ⟦P1⟧(x) = True}

     b. ⟦zaza nenjehin-dRabe⟧ = ⟦zaza⟧ ∩ {x|⟦nenjehin-dRabe⟧(x) = True}
                               = CHILD ∩ {x|(CHASED BY RABE)(x) = True}

Thus we compositionally interpret (32a,b)2. No movement is needed, we just concatenate Ns and P1s, both used independently in nuclear clauses. An N is not always

present: demonstrative Dets, the definite article *ny* and the previous mention article *ilay* combine directly with P1s to yield DPs: *ny nenjehin-dRabe* = "the (ones) chased by Rabe", *io nenjehin-dRabe io* 'that one chased by Rabe that', etc.

Note that our analysis captures judgments of selection restrictions. *#Nanenjika azy ny trano* 'The house chased him' is bizarre in Malagasy just as is its English translation, since houses can't autonomously move and so can't chase things. And we interpret the RC *trano nanenjika azy* 'house that chased him' as the set of houses x such that *nanenjika azy x* 'x chased him' is true. So anything in the denotation of the RC must be a house that did some chasing, violating the selectional restrictions.

Now linguists have ethnocentrically modeled Malagasy RC formation and interpretation on the English pattern, yielding structures like (34) in which the post-N expression contains a full S and its P1 has an argument *wh* that moves pre-clausally and is co-indexed with a trace in its original position:

(34)   zaza [whi [S[P1 nanenjika ny jiolahy] [ti]]].
          child who   chased   the thief

(34) is compositionally interpretable but carries superfluous, hence by Full Interpretation, objectionable, structure: an embedded S, an inaudible P1 argument and a variable binding operator. It obscures the fact that to form and interpret RCs, (33a), the Malagasy child just uses Ns and P1s already learned in interpreting nuclear Ss. What would motivate a child to posit and interpret additional, unpronounced, expressions? English, I claim, provides an answer to this question, Malagasy does not.

To see this let us, refreshingly, reverse the ethnocentric bias and ask: Why doesn't English form RCs in the learning theoretically pleasing way that Malagasy does? A partial answer is given by comparing subject and object RCs:

(35) a. (each) child that chased John   b. (each) child John chased

In (35a) *chased John* is a P1 and the RC is interpretable as in Malagasy. But in (35b) *John chased* has no category and is not assigned a meaning in the interpretation of *John chased the child*. By Compositionality2 only constituents are interpreted.

(36)   **The *Wh*-Movement Trigger**

Being forced to interpret strings in RCs which are not independently meaningful triggers recourse to empty categories and variable binding operators: they are a minimal logical extension of what actually occurs that yields a correct compositional semantic interpretation.

So to interpret *John chased* in (35b) as a P1 we add (an imperfection!) silent structure, a trace *ti*, forming an S, and then a silent operator (more imperfection!) *whi* (or λi) to bind

the *ti* so that the whole expression is interpreted as a P1, and not an S. (36) is our account of a "last resort" way voice impoverished languages manage to relativize[2].

3.1 But couldn't English form RCs in the Malagasy way, expressing (35b) as *child that was chased by Rabe* in which the post *that* material is a P1? And in this case it can, but in general the passive capability of English cannot match Malagasy's productivity. Here are seven types of P1s in Malagasy which lack natural passive paraphrases in English.

[1] P1s with oblique arguments, especially P1s built from intransitive verbs (mostly not passivizable in English) as in (37) or ones with Manner arguments, (38).

(37) a. ny fahazotoana n+[[i+asa]+ana]+Rabe (niasan-dRabe)
　　　 the zeal　　　past+[[AF+work]+CF]+Rabe.gen
　　　the zeal with which Rabe worked
　　 the zeal (which) was worked with by Rabe

　　 b. ny andro　n+aha+teraka+ana+Rabe (nahaterahan-dRabe)
　　　 The day on which [past+[[AF+born]+CF]]+Rabe.gen
　　　 the day on which Rabe was born
　　 the day(that) was born(ed) on by Rabe

(38)　 ny hafalian-dehibe n+i+arahaba+ana+nay (niarahabanay) anao
　　　 the happiness-great　past [[AF+greet]+CF]+1.pl.excl 2sg.acc
　　　 the great happiness with which we greeted you
　　 the great happiness (that) were greeted you by us

The P1 in (39a) may take multiple obliques, such as Benefactives and Instrumentals. The identical CF verbs in (39b,c) take arguments corresponding to these obliques. In (39d) what follows *vehivavy* 'woman' independently denotes a property, as in (39b,c), but not so for what follows 'woman' in the English translation of (39d). Thus English has difficulties forming passive P1s whose external arguments would be oblique.

(39) a. [m+aN+taov (manao) trano　ho an-dRasoa amin'ny birikinay] Rabe
　　　　 pres+AF+do　　　　house for acc-Rasoa with'the bricks+our Rabe
　　　 Rabe is making a house for Rasoa with our bricks

　　 b. [∅+anaovan-dRabe trano] Rasoa
　　　　 pres+[AF+do]+CF+Rabe.gen house Rasoa
　　　 Rabe is making a house for/because of Rasoa

　　 c. [∅+anaovan-dRabe trano] ny birikinay
　　　　 pres+[aN+do]+CF+Rabe.gen house the bricks+our
　　　 Rabe is making a house with our bricks

　　 d. ny vehivavy [∅+anaovan-dRabe trano]
　　　　 the woman　　makes-Rabe house

the woman Rabe is making a house for / because of
*?the woman who is being made by Rabe a house for/because of

[2] An object control verb with the P1 argument the controllee

(40) a. [m+aN+ampy an-dRabe h+aN+petraka (hametraka) ny kodiarana] Rasoa
          pres+AF+help acc-Rabe fut+AF+place the wheel Rasoa
          Rasoa is helping Rabe change the tire

     b. [ampian-dRasoa h+a+petraka+Rabe (hapetra-dRabe)] ny kodiarana
          help+TF-Rasoa.gen fut+TF+place+Rabe.gen the wheel
          Rasoa is helping Rabe change the tire

     c. ny kodiarana ampian-dRasoa hapetra-dRabe
          the wheel which Rasoa is helping Rabe change
          * the wheel which is being helped by Rasoa to be changed by Rabe

A literal translation of (40b) is "The wheel is helped by Rasoa to be changed by Rabe",
which is nonsense. TF verbs are better translated as active than passive.

[3] P1 = Vmotion+V form complex TF P1s whose argument again is theta related to V.
 (41)   a. [niakatra hianatra teny vahiny tany Antananarivo] i Soa
          past+AF+go+up fut+AF+study lg foreign past+loc Antananarivo art Soa
          Soa went up to study foreign languages in Antananarivo

     b. ny teny vahiny [niakaran'i Soa hianarana tany Antananarivo]
          the lg foreign past+[AF+go-up]+CF+art Soa.gen fut+[AF+study]+CF there Ant.
          the foreign languages Soa went up to Antananarivo to study
          *the foreign languages that were gone up by Soa to Antananarivo to be studied

[4] Subject control verbs with the P1 argument the Agent

(42) a. [n+i+kasa h+aN+vaky (nikasa hamaky) io boky io] Rabe
          past+AF+intend fut+AF+read that book that Rabe
          Rabe intended to read that book

     b. [no+kasa+ina-dRabe (nokasain-dRabe) ho vaky+ina (vakina)] io boky io
          past+intend+TF+Rabe.gen fut read+TF that book that
          Rabe intended to read that book
          *Rabe intended that that book get read

     c. ny boky nokasain-dRabe ho vakina
          the book that Rabe intended to read
          * the book that was intended by Rabe to be read (wrong meaning)
          *?the book that was intended to be read by Rabe (intended by whom?)

Both the AF and TF forms of *intend* select a future/irrealis complement verb. The English translations in (41c) fail to clearly bind the Agents of *intend* and *read*.

[5] Raising to Object, hosted by over 50 verbs (Paul and Rabaovololona 1998).

(43) a.   manantena aho fa nanaovan-dRabe trano Rasoa
            pres+AF+hope 1.sg.nom that past+[AF+do]+CF+Rabe.gen house Rasoa
            I hope that Rabe made a house for Rasoa[2]

     b.   [manantena an-dRasoa ho nanaovan-dRabe trano] aho   (RTO)
            pres+AF+hope acc-Rasoa as past+do+CF-Rabe.gen house I
            I hope Rabe made a house for Rasoa

     c.   [∅+antena+ina+ko (antenai-ko) ho nanaovan-dRabe trano] Rasoa
            pres+hope+TF+1.sg.gen as made+CF-by Rabe.gen house Rasoa
            I hope Rabe made a house for Rasoa

     d.   ny vehivavy [antenaiko ho nanaovan-dRabe trano]
            the woman hoped+1sg.gen as built+CF-Rabe.gen house
            the woman I hoped Rabe built a house for
          ??the woman who is hoped by me to have been built a house for by Rabe

Complex predicates like those in (43c,d) are the major wayof relativizing arguments of complement clause verbs. The "raised" object, *an-dRasoa* in (42b) for example, must satisfy the theta and selectional requirements of the P1 following *ho*. Virtually all verbs that take complement clause objects permit Raising to Object, as in (43b).

[6] Raising from DPs (Keenan & Ralalaoherivony, 2000)

(44) a.   Hadino+ko ny anarana+ny (anarany)
            forget(TF root)+1sg.gen the name+3.gen
            I forgot his name

     b.   Hadino+ko anarana izy   (Raising from DP)
            forget(TF root)+1sg.gen name 3nom
            He had his name forgotten by me

     c.   ny mpianatra hadinoko anarana
            the student forget(TF root)+1sg.gen name
            The student whose name I forget
          *?the student who is name-forgotten by me

(45)   a. Fohy ny andro n+i+velo+mana+Rabe (niveloman-dRabe)
            Short the days past+[AF+live]+CF+Rabe.gen
            The days during which Rabe lived were short/few

b. Fohy andro nivelomana Rabe   (Raising from DP)
      Short day   lived+CF   Rabe.nom
      Rabe had a short life
  c. ny olona fohy andro nivelomana
      the people short day lived+CF
      the people who had a short life

[7] P1s with internal Agents binding anaphors:

(46) a. [m+aN+vono (mamono) tena ho an'ny taranatsika] isika
         pres+AF+kill self for'the descendant+1pl.gen.incl   1pl.nom.incl
         We kill ourselves for our offspring

  b. [∅+amono+ana+tsika (amonoantsika) tena] ny taranatsika
      pres+[AF+kill]+CF+1pl.gen.incl. self the descendant+1pl.gen.incl
      We kill ourselves for our offspring

  c. ny taranatsika amonoantsika tena
      the descendants that [AF+kill]+CF+1pl.gen.incl self
      the descendants that we kill ourselves for
        the descendants that are killed for by us ourselves

(47) a. [n+aN+ampy (nanampy) azy h+i+tsara ny fanadinana] isika
         past+AF+help 3acc fut+AF+judge the exams we.incl
         We helped them grade the exams

  b   [n+if+an+ampy h+i+tsara ny fanadinana] isika
      past+rec+AF+help fut+AF+judge the exams we.incl
      We helped each other grade the exams

  c. [nifanampian+tsika hitsarana] ny fanadinana
      past+[rec+[AF+help]+CF]+1.gen.pl.incl fut+[[AF+judge]+CF] the exams
      We helped each other grade the exams

  d. ny fanadinana [nifanampian-tsika hitsarana]
      the exams past+[rec+[AF+help]+CF]+1.gen.pl.incl fut+[[AF+judge]+CF]
      the exams we helped each other grade
        the exams that were reciprocally helped by us to be graded

So, (46) and (47), genitive Agent Phrases in Malagasy can antecede the reflexive *tena* 'self' and the reciprocal *-if-*. By-Phrases in English are not as powerful.

3.2 "Extraction" constraints in Malagasy [1] – [7] illustrate Malagasy's rich means of forming complex P1s. All crucially use the voice system. [2] – [6] involve additional P1 building structure. Concatenating an N with these P1s forms RCs on its argument. No movement is used and semantic interpretation is strictly compositional. So the class of relativizable arguments in Malagasy is determined by the class of P1s it can form.

Now assume for sake of argument that [1] – [7] are the only ways of building complex P1s in Malagasy[2], and that RCs are the only candidates for a movement analysis. Then the data which support classical extraction constraints hold in Malagasy. No Vacuous Quantification is satisfied in (48) since what follows the first N is not a P1. Analogously for the Coordinate Structure Constraint, (49b), and Subjacency, (50b).

(48) *ny trano nataon-dRabe ny farafara
     the house built-by-Rabe the bed
     The house that Rabe made a bed

(49)   a. milalao ao an-trano ny zazalahy sy ny zazavavy
       play   there in-house the boys and the girls
       The boys and the girls are playing in the house

       b. *ny zazavavy milalao ao an-trano ny zazalahy sy
         the girls (that) play there in-house the boys and

(50)   a. nandositra ny olona nangalatra ny omby
       fled   the person stole (AF) the cow
       the person who stole the cow fled
       b. *hitako ny omby nandositra ny olona nangalatra
   I+see the cow (that) fled the person (who) stole

So Malagasy satisfies these constraints even though they are not "active"in its grammar (say as constraints on rules or representations). If we eliminate them from UG Malagasy would gain no new expressions, English would gain many. So we treat universal constraints as properties shared by all highly valued grammars for natural languages. Typically they are merely entailed by the independently defined generative mechanisms of any given language but are not usually a functional part of a grammar of a language as they pick the lowest common denominator cross linguistically. Even English likely satisfies a stronger form of Subjacency than Danish or Japanese.

4. "Standard" approaches to Austronesian clause structure are twofold: Advancement Analyses, and V2-Topic Analyses. Among the former we treat the Generalized Passive analysis of GHT (1992) and Paul (1999). There, similar to BJR (1999), the P1 argument in nuclear clauses is base generated empty and DPs move into that position to get nominative case, the movement motivated by the absence of case assignment in its initial, theta, position. Only slightly different from our perspective are Payne (1982), De Guzman (1988), Gerdts (1988), and Aldrich (2002), in which the languages are treated as ergative and the advancement rules generalize Anti-Passive.

The V2-Topic approach is presented in Richards (2000) and Pearson (2001, 2005), most thoroughly in the latter, discussed here. The P1 argument is base generated caseless and assigned case (and theta role) by co-indexing with a case marked empty operator.

Otherwise it behaves like fronted topics in Icelandic or German. We illustrate these two approaches in more detail and then compare them with our BG analysis.

4.1 Generalized Passives GHT was the first attempt to characterize the voice system of Austronesian in GB terms. Some of the initial proposals have since been modified (Paul 1999, R& T 2000, Travis 2005). We consider the original proposal for its simplicity and clarity, focusing on those aspects that survive current reformulation. (51) is GHT's derivation of a nuclear TF S. UTAH (Baker 1988) is assumed throughout, a stipulation with far reaching consequences.

(51)   IP
               I'   SPEC
                                              Theme
  INFL   VP

                 SPEC   V'
                 Agent
                      V   NP

  -Vna   Rabe   enjika t_i ny jiolahy$_i$
  [Enjehin$_k$-[dRabe t_k t_i]] ny jiolahy$_i$
  The thief was chased by Rabe
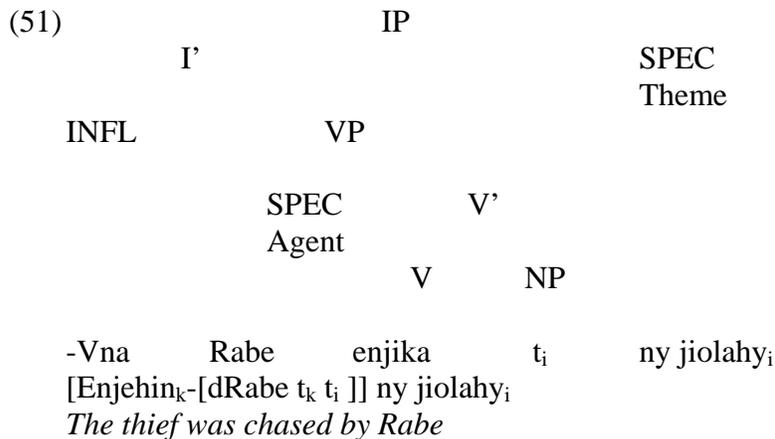
-Vna occurs in INFL and assigns case (by stipulation) to the Agent in SPEC VP (despite not forming a constituent with it). The Theme of V lacks case (by stipulation) and moves to SPEC IP, where case is assigned (a stipulation). (So the Theme moves across the Agent, leaving a trace unbound, a minimality violation). V then moves to INFL to support the suffix *-Vna* (again a stipulation, it could be supported by moving around the constituent to its right, *Rabe*, as with Latin *-que* 'and'). Then the Agent attaches to it in the same complicated way that possessors n-bond to nouns (see Paul 1996a), even though, again, the two are not a constituent. (This remains true in R&T 2000 and Travis 2005 in which the INFL head would occur within a sister to the VP containing the Agent)[2].

The derivation of the corresponding AF S differs from (51) in that *aN-* occurs on V and assigns case to its Theme. Nothing assigns case to the Agent in SPEC VP, so it moves to SPEC IP to get case. V moves to INFL. Now both AF imperatives and present tense indicatives present an AF-specific *m-* in INFL, which Pearson (2005) takes as the distinctive AF morphology. Hosting *m-* could motivate V movement, but why does *m-* not assign case to the Agent as *-Vna* in (51) does? GHT derive CF Ss assuming the simultaneous presence of *-Vna* in INFL and *aN-* on V thereby assigning case to all subcategorized arguments. So an oblique DP raises to get case in SPEC IP, its Prep perhaps absorbed into the verb.

Paul (1999)'s structures are richer than GHT's but use the same X' format and

assume UTAH. Arguments of the verb move to get case as in GHT. The *a-* forms are explicitly treated, with *a-* assigned a position above the arguments that *-Vna* promotes to the P1 argument position. But CF Ss differ from AF and TF ones in that ascension to SPEC IP is not motivated by the Case Filter, just the EPP (which is just the stipulation that the P1 argument has to be filled).

4.2 V2-Topic Like GHT/Paul, Pearson's (2005) pre-movement structures satisfy UTAH and case is determined in different structural positions associated with voice morphology. The P1 argument is base generated in SPEC-TopP (Topic Phrase), by stipulation co-indexed with a case marked empty operator in the SPEC of the next lower projection, WhP, which, by stipulation, licenses *wh*-operators. The empty topic operator is moved there from a case position. The movement is A' movement, to a non-case position, not A movement as in GHT/Paul, so Pearson avoids the minimality violations incurred by GHT/Paul using long A movement. On the other hand GHT/Paul have an in principle motivation for their movement – satisfying the Case Filter. Pearson must postulate an *ab initio* requirement that TopP exists and its SPEC filled. (52) is his structure for the TF *Nenjehin-dRabe Rasoa* 'Rasoa, Rabe chased her:

(52)   TopP

                Top'   DP
                                                 Rasoa$_i$
        Top   WhP

                      Wh'   DP
    Op$_i$

            Wh   TP

                        T   EP
                        n-
                              DP   E'

                                    E   vP
                                   -ny
                                  (weak)

                                        DP   v'
                                        Rabe
                                              v   AspP

                                                DP   AspP'
                                                t$_i$

                                                      Asp   VP
                                                      -Vn
                                                    (strong)

```
                                        DP        V'
                                        t_i

                                        V        enjika
        Nenjehin-dRabe Rasoa                     'chase'
```
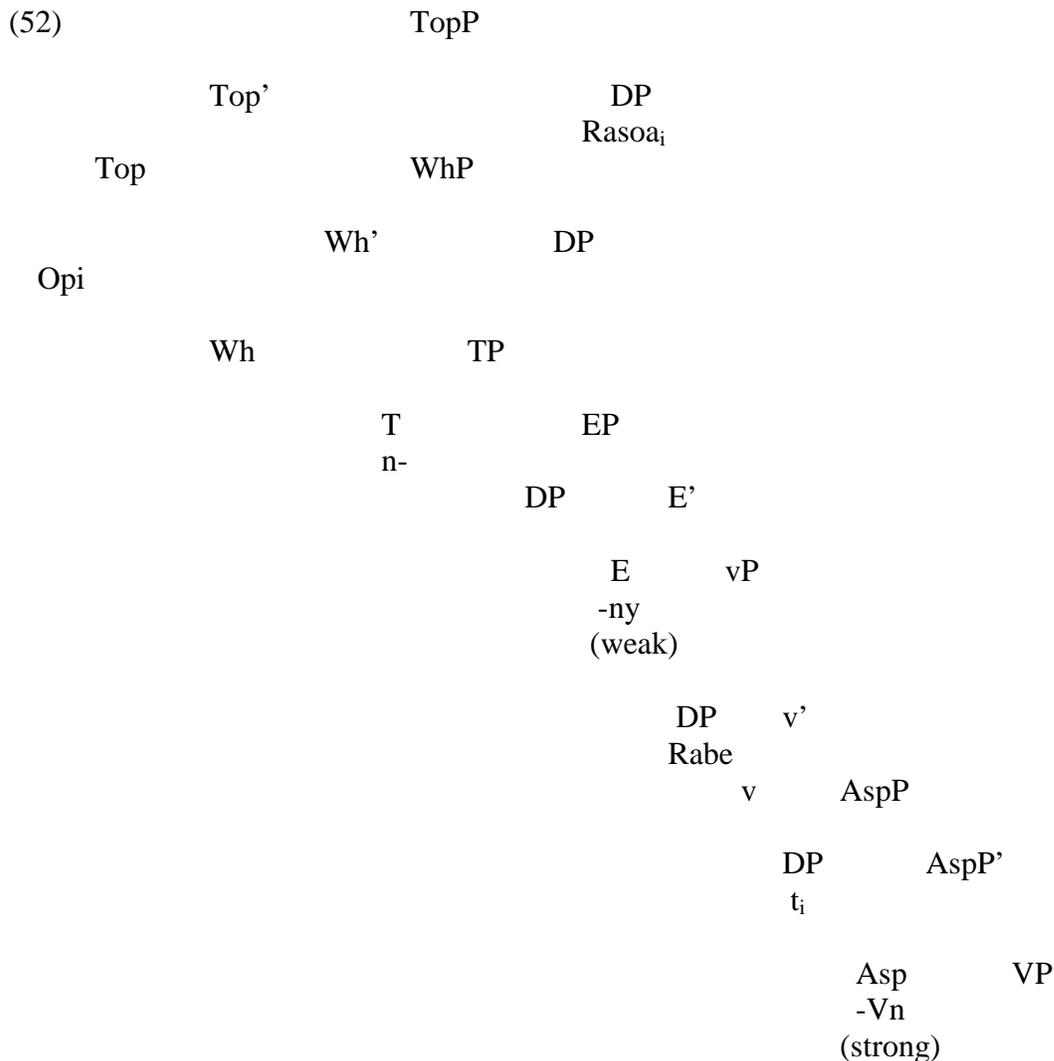
In (52) *Rasoa* in SPEC-TopP is coindexed with a null operator in SPEC-WhP. That operator originates in SPEC VP where it is assigned a theta role (Theme) there by V but not case. Since the Asp head immediately preceding it has a strong (by stipulation) case feature, *-Vn (= -Vna)*[2], it moves across that node into SPEC-AspP, acquiring accusative case. Then it moves directly to SPEC-WhP. *Rabe* is generated in SPEC-vP where it is assigned the Agent role by little v. The E head to its left has a weak (by stipulation) case assigner *-ny*, so *Rabe* remains in situ, where it is assigned case.

The verb is formed from the root *enjika* which then by head movement goes first to Asp, hosting the *-Vn(a)* suffix, then the verb+*-Vn(a)* moves to v (without effect), then to E, hosting the suffixal linker *-ny*, then the whole complex, *-enjehiny,* moves to T where it hosts the tense prefix *n–*.

In contrast the AF version of (52), *nanenjika an-dRasoa Rabe* 'Rabe, he chased Rasoa', presents *aN-* in Asp, stipulated weak, whence *Rasoa* in SPEC-VP does not move and is assigned (accusative) case in situ and the Theme theta role. In E however we now find the AF morpheme *m-*, stipulated as strong. The SPEC of vP is filled with an empty operator which moves to SPEC-EP, acquiring nominative case, and then moves directly to SPEC-WhP. It is coindexed with *Rabe*, generated in SPEC-TopP.

4.3 a comparison Broadly first, the BG approach is bottom up and constructive, deriving semantically interpreted categorized strings. We build *verbs* by functions that affix voice morphemes to roots. The verbs+Merge determine the case of its arguments and the structure of nuclear clauses, which in turn supports a compositional interpretation, including theta role assignment. UTAH is not satisfied, nothing moves in the syntax or binds variables in the semantics. In a slogan, *the syntax of Malagasy nuclear clauses rides on its verbal morphology*.

In contrast GHT/P/P (GHT, Paul and Pearson) assume a clause level tree structure initially and derive others from it with lexical items inserted late. Voice morphology is scattered, not localized on roots, and many stipulations concerning case assignment and its absence are needed to force DP movement, with additional ones needed to force verb movement. UTAH is respected. No compositional semantic interpretation is attempted. Below we consider some more specific comparisons proceeding roughly from general design feature differences to differences in predications concerning Malagasy structure.

4.3.1 Semantic interpretation Of the approaches considered only BG offers a compositional semantics. This has two specific consequences which favor the BG approach: First, it gives us an explicit account of the theta equivalence of nuclear Ss in different voices built from the same root, as well as their logical equivalence when their corresponding DPs are the same individual denoting ones, as in (2a,b). It also gives us an

ccount of the entailment relations between Ss like (30a,b) built from a CF verb or an AF, TF, or GF one with an overt oblique, say a Benefactive as in (30b).

Let us emphasize again that a compositional semantic interpretation is the only way we have of accounting for how speakers understand novel utterances: they know what the lexical items mean and how things built in that way take their meaning as a function of the meaning of the expressions they are built from. Moreover we all focus initially on the same class of voice related nuclear Ss since we recognize that they differ in form and meaning in a regular way. So the GHT/Paul/Pearson proposals are incomplete.

Second, faced with multiple Ss built from a given root in different voices, a theory without a compositional semantics leads one to adopt UTAH, as GHT/P/P do, whereas an explicit semantics leads one to ignore it, as we do. UTAH effectively stipulates that a given theta role can only be assigned in one structural configuration. Now if you lack a systematic means of assigning theta roles you may still claim to account for the *sameness* of assignment in two Ss on the grounds that they are derived from isomorphic syntactic sources where theta role assignment takes place.

But the fact that different configurations may assign the same theta role, as in (9) and (14), is no more surprising than that 7+5 and 9+3 both equal 12. If two people tell you "I've just summed two positive integers to get 12" would you assume they summed the same two? Surely not. Addition is not one to one; neither is theta role assignment.

Accepting UTAH has very significant consequences for the syntactic analysis of Ss built from a given root with different voice affixes. For fixed choices of DPs with AF, TF, and GF variants we see that they have different relative orders in audible structure, so some reordering must be involved. On the BG approach no DPs are moved in building AF, TF, GF and CF variants.

Lastly a compositional semantic interpretation favors minimalist syntax: Each addition to the syntactic structure has to be interpreted (**FI**), leading to a practical "economy" principle: *If you can't say what it means don't use it*. For example Pearson makes crucial syntactic use of a WhP constituent in nuclear Ss. So he should be obliged to tell us its semantic contribution to the meaning of a nuclear declarative it occurs in. But *wh*-expressions are presuppositional in a way that nuclear Ss are not. Similarly how do the empty Wh- and Top heads contribute to the meaning of the whole?

4.3.2 Learnability Children are clearly motivated to learn to use and interpret expressions they actually hear. And on our representations only audible expressions, both functional and content ones, are semantically interpreted. Invoking inaudible objects makes it harder to understand how children learn the language. What triggers the assumption of null elements and operators, and what guides the child in their interpretation? We offered a trigger for invoking variable binding operators in the interpretation of Relative Clauses in English, but Malagasy lacks comparable motivation.
4.3.3 Sui generis? Both Richards/Pearson and GHT/Paul try to understand Malagasy syntax by assimilating it to operations in more familiar languages: Passives for

GHT/Paul, and Germanic V2-Topics for Richards/Pearson. And in each case there *are* substantive similarities. The case and relative hierarchical positions of the DPs in Malagasy nuclear clauses varies with the verbal morphology, in common with actives and their passives in English. And, as in V2-Topic languages, the P1 argument in Malagasy (any voice) is required to be surface-definite, a common property of Topics. More deeply, one cannot relativize a non-Topic from a topicalized clause in German, in analogy to the impossibility of relativizing a DP other than the P1 argument in Malagasy.

But the Malagasy voice system also fails to correspond to Passives in many non-trivial ways, by now well known (Pearson 2005, footnote 4 here, Kroeger 1993). I'll just note first, in distinction to Passives in English, the non-AF -*Vna*, *a*-, and -*ana* verbs in Malagasy present imperative forms (as do the AF verbs). In fact these are the normal forms used when there is a choice. In a recent acquisition study (see Hyams, Manorohanta, and Ntelitheos 2004) covering three children aged 18 - 30 months over an 8 month period, K&M (2004) report that all 26 of the imperatives in the data were non-AF forms. Second, Agent Phrases in non-AF verbs can be controlled. (40b) and (41b) are one type of example, a second comes from coordinate verbs, as the following found examples show (Keenan 1995):

(53)   araka izao [[ahitana-∅i sy ahafantara-nareoi] ahy] izao
       according this [[[see+CF and know+CF]+2.gen.pl] 1.sg.acc] this
       according to that which you see and (you) understand of me

(54)   ataoko fa [voa+jery-∅i sy voa+dinikaoi] tsara ireo sary teo
       I+think that [[TF+see and TF+examine+2sg.gen] good those pictures there]
       I think that those pictures there have been seen (by you) and examined well by you

K&M-01 show that Agent phrases of TF verbs are present or controlled in about 2/3 of their occurrences. So in distinction to English passive verbs, TF verbs derived from transitive roots seem not to be intransitive. And third, the other side of this coin, Agents of verbs in all voices can control other missing Agents, as the same examples show.

That the P1 argument in Malagasy is structurally different from a fronted Topic in German is supported by the fact that in Germanic fronted Topics retain the case they had in the pre-topicalized structure, there being no analogue of a caseless DP there:

(55)   a. Ich habe das Buch dem Hans gegeben
          I.nom have the.acc book the.dat Hans given
          I gave the book to Hans

       b. Dem Hans habe ich das Buch gegeben
          the.dat Hans have I.nom the.acc book given
          To Hans I gave the book

Secondly, V2-Topics with different theta roles do not correspond to verbs with different morphology. Had we fronted *das Buch* instead of *dem Hans* from (55a) the shape of the

auxiliary and past participle would remain unchanged. Further, at least in German, Topic Fronting as in (55b) is largely a root clause phenomenon (Daniel Büring, pc). Thus from the topicalized (56a) it is not natural to form a polar question, (56b), though a Left Dislocated version with resumptive pronoun is possible, (56c)

(56)   a. Das Buch haben die Studenten nicht gelesen
            the book have the students not read
            That book the students haven't read

       b. *Das Buch haben die Studenten nicht gelesen?
            That book, have the students not read (it)?

       c. Das Buch, haben die Studenten das/es night gelesen?
            That book, haven't the students read it?

In contrast questions formed from AF, TF, GF or CF Ss in Malagasy are fully natural.

(57)   a. Tsy novakin'ny mpianatra io boky io
            Not read+TF'the students that book that
            That book wasn't read by the students

       b. Tsy novakin'ny mpianatra ve io boky io?
            Not read+TF'the students ? that book that
            Wasn't that book read by the students?

Similarly Topic Fronted Ss are ungrammatical in many subordinate contexts in German, (58b), but the corresponding Malagasy Ss are fully natural, (58c).

(58)   a. ...weil die Studenten das Buch nicht gelesen haben
            because the students that book not read have
            because the students didn't read that book

       b. ... *weil das Buch haben die Studenten nicht gelesen
            because the book have the students not read

       c. ... satria tsy novakin'ny mpianatra io boky io
            ... because not read+TF'the students that book that
            ... because that book wasn't read by the students

Fourth Malagasy has two topicalizing operations independent of voice morphology. This is unsurprising on the BG approach since Topicalization and P1 formation are independent. It is less expected on the V2-Topic one (though Pearson, his footnote 16, accepts it). Why topicalize from an already topicalized S (TopP)? But the P1 argument does, occasionally, occur fronted, usually with a contrastive meaning in main clauses:

(59)   – Inona no vaovao?
            What Foc new = What's up?

– Tsy misy,   ny mpianatra milalao, ny mpampianatra mifampiresaka
      Nothing,   the students are playing, the teachers are chatting with each other

And second, overt topicalization by fronting a constituent and inserting the particle *dia* is frequent. If the fronted element is not the P1 argument it is usually resumed with a pronoun, but not when it is the P1 argument. Dia-Topicalization even applies to clauses with a clefted constituent, limited in core cases to the P1 argument:

(60)   Ity radary ity dia ny Rosiana no nanao azy
       this radar this top the Russians Foc past+AF+make 3acc
       As for this radar, it was the Russians who built it
In sum then we favor treating the Malagasy voice system as a structure type in its own right. When we have a satisfactory analysis we can compare the results with voice variation in English and V2-Topicalization in Germanic. There is a kind of Hubble Constant methodological moral here:

UG Accessibility
      UG looks the same no matter what language we view it from

4.3.4 Case assignment is a major domain in which our approach and that of GHT/P/P differ (Both omit the morphophonology of case marking, notably *n-bonding* of genitives). On our approach the case of arguments of predicates is simply selected by the predicate, supporting the following natural generalization:

(61) In nuclear Ss, a predicate P selects the case of a DP iff P theta marks that DP.

In contrast, in GHT/P/P DPs get case in two ways: assignment by voice morphology and *free case*, by which I refer to the assignment of nominative in SPEC-IP in GHT/Paul and the co-indexing of the DP in SPEC-TopP with the empty operator in SPEC-WhP in Pearson. By way of comparison:

   1. The GHT/P/P approach results in too many case assigners with intransitive verbs and not enough with ditransitive ones. In GHT/P/P the *aN-* morphology in (62) assigns accusative case to the right sister of the verb, *azy* 'it'. *Aho* 'I' gets free case.

(62)   m+aN+tolotra (nanolotra) azy ny vahiny aho
       pres+AF+offer   3acc the guests 1sg.nom
       I offer it to the guests

But what assigns case to *ny vahiny* 'the guests'? There is insufficient morphology, so some mechanism other than voice morphology must be available to assign case.

      With intransitives, as in *nitomany aho* 'I cried' or *nandainga aho* 'I lied' *aho* gets free case and there are no other DPs in need of case. So the AF morphology *i-* and *aN-* do not assign case, though they do with the many transitive verbs each forms. So AF *i-*

and *aN-* do not have a uniform interpretation: sometimes assigning case, sometimes not. This variation in interpretation makes them harder to learn.

2. Regarding the TF verbs built by suffixing *-ina* to an *aN-* prefixed AF verb, (16), consistency requires GHT/Paul to say that all the arguments get case in situ so none get free case and the derivation crashes (an EPP failure). In Pearson these verbs cannot be derived as *aN-* and *ina* are in complementary distribution, occurring in the same position. But in fact the Theme DP does get free case, (16). And, in GHT/Paul, if case marked DPs can have their case overridden why can not the Agent? BG has no problem with these six verbs: we simply include them in the domain of INA as we did, along with AF causative verbs prefixed with *amp-*, but no other derived verbs.

3. Malagasy also has six or so verbs built from suppletive roots: AF *mi+tondra,* TF *entina* 'carries', AF *m+aka*, TF *alaina* 'takes', AF *mi+hinana*, TF *hanina* 'eats'.

(63)   a. m+i+tondra ny angady Rabe   b. ∅+entin-dRabe ny angady
            pres+AF+carry the spade Rabe   pres+carry+TF-Rabe the spade
            Rabe is carrying the spade   *Rabe is carrying the spade*

On GHT/P/P *i-* in (63a) assigns case to *ny angady* 'the spade'. But then what prevents deriving the phonologically natural **tondrain-dRabe,* suffixing the root in the expected way with *-ina*? In our case *tondra* is not in the domain of INA, nor *ent-* in the domain

of I. And we stipulate that ⟦INA(<ent-, RT{AG,TH}>)⟧(x)(y) =

I(<tondra,RT{AG,TH}>)>)⟧(y)(x). This ad hoc clause in the definition of ⟦·⟧, the semantic interpretation function, yields the logical equivalence of (63a,b) above. GHT/P/P also need a condition blocking insertion of *-ina* when the root is *tondra* and a second blocking the insertion of *i-* when the root is *ent-*. But this still does not derive the logical equivalence above, so they fail to represent the suppletion relation.

4. There are two types of verbs which assign case in the absence of the morphology GHT/P/P use for that purpose. First, about 20 vowel initial roots form AF verbs by prefixing the AF present tense *m* without intervening morphology: *ila* ==> *mila* 'needs', *ino* ==> *mino* 'believes', *indrana* ==> *mindrana* 'borrows', *anana* ==> *manana* 'possesses'. And these verbs do assign case to their object, as in (64). They also have regular TF forms: *ilaina*, *inoana*, *indramina*, *ananana*.

(64) a. M+ila (mila) azy aho   b. Ila+ina+ko (ilaiko) izy
            pres+need 3acc 1nom   need+TF+1sg.gen 3nom
            I need her   *I need her*

It is unexceptional that the lexical verb rather than voice morphology assigns case to its complement. In languages with rich case marking such as Latin, verbs like *amare* 'to love' and *vincere* 'to conquer' take accusative case objects while *parere* 'to obey', *imperare* 'to order', *invidere* 'to envy', *respondere* 'to answer' and *persuadere* 'to

persuade' and others take dative complements though they do not vary with regard to voice morphology. Choice of case is naturally stipulated given the root and its voice morphology, as in BG, but voice morphology alone does not suffice. Cases like (64) can be treated in BG by listing *ila* in the Lexicon both as a root of category RT{EXP,TH} and as a verb of category S/[DPacc,DPnom] whose semantic interpretation is given by

⟦ila⟧(x)(y) = True iff <y,x> ∈ ILA & TH(x,ILA) & EXP(y,ILA).

Perhaps more convincing here are the many deponent verbs in Latin and Greek which are formed with passive morphology but are transitive taking a nominative and a non-nominative argument. Lat: *comitari* 'to accompany', *tueri* 'to protect', *aggredi* 'to attack', *comperiri* 'to find out'. Gildersleeve & Lodge (1913) list 81 examples. Similarly Modern Greek has many such deponents (Dimitris Ntelitheos, pc; see also Embick 1997): *dehome*, *thimame*, *mahome* (I) accept, remember, fight, as in (65). So the case of the complement cannot simply be read off from the voice morphology.

(65) O Giannis dehtike tin protasi mu
    the Giannis.nom accept+NonAct.pst.3sg the proposal.acc my
    Giannis accepted my proposal

The second case concerns 25 or so roots which occur as TF verbs without affixes: *vita* 'finished', *resy* 'defeated', *hadino* 'forgotten'. They may form transitive AF *aN-*verbs in a regular way: *mamita* 'finishes', *mandresy* 'defeats', *manadino* 'forgets', and, surprisingly, also accept *-ina* TF suffixes: *vitaina* 'finished', *resena* 'defeated', *hadinoina* 'forgotten'. The root form is more stative, or telic, as Paul (1999) observes.

(66) a. Hadino+ko ny hira   b. m+aN+hadino (manadino) ny hira   aho
     forgotten+1sg.gen the song   pres+AF+forget   the song.acc 1sg.nom
     I have forgotten the song   *I forget the song*

Thus case assignment must be partially independent of voice morphology and at least be conditioned by the presence of the root.

  5. Case assignment by voice affix does not extend naturally to other case assigners in Malagasy, namely Adjectives and Prepositions. Arguably there is a very general mechanism, represented by a nasal segment (*na* in GHT, *ny* in Pearson) which constructs genitive case complements on verbs, nouns, adjectives and prepositions, (67).

(67) a. jamba (adj) 'blind'   a'. jamban'ny vola 'blinded by money'
    b. marary (adj) 'sick'   b'. mararin'ny tazomoka 'sick from malaria'
    c. ambony (prep) 'on, above'   c'. ambonin'ny latabatra 'on the table'
    d. akaiky (prep) 'near (to)'   d' akaikin-dRabe 'near Rabe'

But a few adjectives, *antonina* 'sufficient', *feno* 'full', *sahaza* 'suitable' and a few prepositions, *ho* (*an*) 'for', *lavitra* 'far (from), select accusative:

(68) a. Antonina azy io satroka io   b. Feno azy ny siny
       sufficient him that hat that   full 3acc the jug
       That hat suits him   *The jug is full (of) it*

a. Lavitra ahy izy io   b. Nividy akanjo ho anao izy
    Far   1sg.acc 3nom dem.sg   bought clothes for 2sg.acc 3nom
       That is far from me   *He bought clothes for you*

In these cases in BG it is natural to enter them in the lexicon, subcategorized for DPs in the appropriate case. GHT/P/P could do the same, but it would mean that there are means of case assigning independent of verbal morphology.

4.3.5 Cooccurrence restrictions are a last case where I feel that the GHT/P/P approach is inadequate in that the separation of the voice morphology and verbal roots presents as independent, expressions whose cooccurrence is restricted. In Paul for example, roots like *tondro* 'point out' and *fafa* 'sweep' take both *a-* and *-Vna*, building P1s whose arguments have different theta roles. So these affixes are not conditioned variants. Indeed Paul (1999:Ch 3) is an empirically enlightening discussion of *a-* vs *-Vna* verbs[2]. But Paul puts them in different projections where they assign case to different DPs. Yet no root takes both affixes at once, whence additional mechanisms forcing complementary distribution are needed. So tree structures that separate voice affixes from each other and from roots are incomplete – they leave cooccurrence restrictions unacknowledged – and misleading, they imply independence where there is none. In the BG format such facts are unproblematic: the domains of the INA and the A functions are simply disjoint.

And in general to define the voice affixing functions INA, AN, I etc. in BG we must define their domain, that is, say what they affix to. Thus we can state selectional and cooccurrence restrictions between roots and their affixes.

Pearson avoids some of these problems: AF verbs lacking *aN-* or *i-* are not problematic, as the distinctive AF morphology for Pearson is the initial *m-*. Also not problematic is the complementary distribution of *-ina* and *a-*, as they occur in the same position. And of course Pearson avoids the Minimality violations inherent in moving one argument over another to get case, since his movement is A' movement.

But avoidance of Minimality violations does oblige Pearson to stipulate that the DP in SPEC-TopP gets case and a theta role by co-indexing with the empty operator in SPEC-WhP. But if this is sufficient to assign it case why does it not have the appropriate case forms, as it does in German. It's form is fixed (our nominative).

And still, as in GHT/Paul there are several cases in Pearson of cooccurrence dependencies between expressions in nodes that are not sisters and cannot directly select each other. For example the head of AspP has a strong feature iff the next lower SPEC is filled with the empty Topic operator, Op. Similarly the AF initial *m-* occurs in E iff the next lower SPEC-vP is filled with Op. There are weak and strong morphemes in E, *m-*

and *-ny*, and weak and strong ones in Asp, *-Vn* and *aN-*, but only two of the four pairs can cooccur, one weak and one strong. The CF morphology *-ana* is spelled out only when its SPEC is the empty Op. (In this last case Pearson appeals to a generalization of Sportiche's Doubly Filled Comp Filter). So guaranteeing these cooccurrence restrictions entails many stipulations not explicitly built in to the initial tree structure.

And in general I don't see how heads, like E or Asp, assign case or theta role to something they do not form a constituent with, namely the next SPEC down. How do they know there is a SPEC there? The SPEC does not project its properties to the maximal projection, and it doesn't seem to make sense to say that the event head E assigns case to its vP sister, or that Asp assigns case to its VP sister, as Predicate + Argument structures are not the kinds of objects that carry case.

Similarly in Pearson little v assigns the Agent theta role to the DP in SPEC-vP. But the theta role of that DP is decided by the root V, which little *v* does not see. How does *v* know that Agent not Experiencer or Recipient is to be assigned? The role that is needed is the highest ranking one selected by the root in V, several maximal projections down.

4.3.6 Constituent structure Only our approach claims that non-AF verbs form a genitive constituent with their Agents. In GHT/P/P the verb moves from elsewhere to be string adjacent to its Agent, but they are not sisters (nor are they in Travis 2005). Primary evidence for this is that such Agent Phrases pronominalize with the **gen** series of pronouns just as possessors do. For example *trano+ko* 'house+my' and *hadino+ko,* (66), 'forgotten by me'. When the host is vowel final (other than *-na*, *-ka*, *-tra*) a nasal is introduced: *tranon-dRabe* 'house of Rabe' and *hadinon-dRabe* 'forgotten by Rabe'. When the host ends with *-na*, *-ka*, or *-tra* and stress is antepenultimate the ending drops or, if the possessor/Agent begins with a continuant, mutates to the closest non-continuant, as in *soroka+Rabe = soro-dRabe* 'shoulder of Rabe' and *atosika+Rabe = atosi-dRabe 'pushed by Rabe'*, (18). See Paul (1996b) and K&P. The claim for syntactic constituency is supported by coordinations like that in (70), taken from a story in a first grade reader, where an agented circumstantial built from a transitive root conjoins with an agented GF ditransitive, both having category S/[DPacc,DPnom]. In (71) from Keenan (2000) a similar CF verb coordinates with a lexical AF transitive verb[2]:

(70)   Raha mahay manaiky ianao ... If you are able to obey ...
        dia hividianako sy homeko fahasambarana (ianao)
        then fut+buy+CF+1sg.gen and fut+give+GF+1sg.gen happiness.acc (2sg.nom)
        If you will be obediant ... then I will buy for and give you happiness

(71) a. n+aN+tondro+ana+ko (nanondroako) ilay toerana ianao
        past+[AF+indicate]+CF+1sg.gen   that place   2sg.nom
        I pointed out that place to you

    b. nanondroako sy naka ilay toerana ianao
       indicate-by-me and took that place you
       You were shown by me, and took, that place

5. "Extraction" Of the mainstream works considered only the V2-Topic approaches, Pearson and Richards (2000), present an explicit "Extraction" Constraint. Pearson's account is that the null Topic Operator Op, and a Wh-operator used in *wh*-questions and RCs compete for the same position, so a *wh*-word and a DP in SPEC-TopP cannot cooccur, (72b). If they did each would have to be co-indexed with its operator in SPEC-WhP (which Pearson must stipulate to be fatal).

(72) a. Iza no nanenjika azy?   b. *Iza no nanenjika azy Rabe
      Who Foc chased.AF 3acc   Who Foc chased.AF 3acc Rabe.nom
      Who (was it who) chased him?

But it is unclear on this view how to derive Wh-questions. Why don't we derive (72b) with *Rabe* in SPEC-TopP, note that the derivation crashes and simply fail to generate (72a)? Somehow the TopP superstorey must disappear when an interrogative operator is present. Pearson (his footnote 26) remarks that "the *no* phrase ... presumably lacks a TopP projection". But the mechanism here is not clear.

GHT/Paul offer no alternative to the stipulation that only "subjects" (SPEC IP) extract. Given that, (72b) is underivable, as both *Iza?* 'Who?' and *Rabe* would have had to be in SPEC IP. Travis (2005) however does suggest without explicit formulation a different line of explanation. Namely assume that underlying structures for Malagasy are in the English SVOX order and derive the audible VOXS order by fronting the tensed [VOX]. Then a general derivational constraint forbidding movement out of an already moved constituent only allows SPEC-IP to be moved, as it was the only part of the original structure not in the moved part. (It was itself moved though, so the derivational constraint must only forbid moving a proper subconstituent out of an already moved constituent). Aldridge (2002) suggests a similar idea. But neither presents an explicit proposal, so we don't actually know what constituent is moving (probably I' in Travis 2005 since it must include tense but leave SPEC-IP behind) and we don't know where it moves to or what motivates the movement.

But to my mind the least satisfying aspect of this suggestion is that it assumes an obviously incorrect underlying order for Malagasy, DPnom+P1 thereby creating the need to move. It also assumes we get the observable order by moving the P1 left around the DPnom creating an island. Surely it would be more economical to move the DPnom rightward around the P1. More generally why not take VOXS as a universal order for head initial languages (as we have done for Malagasy) and derive SVOX by a more modest "Subject Fronting" rule, which has much independent motivation. The constraint then would allow extraction of subjects (and other arguments of the V) but block extraction properly out of subjects, a desirable result. VSOX could be derived by attracting the "S" to the head P1 (and often to a verb initial position, always a possibility in a VSOX language).

That said, we note two problems with P1 fronting creating an island in Malagasy and

then leave it to proponents of this approach to formulate the movement explicitly and to justify it. First, as is well known, adjuncts (locatives, temporals, instrumentals) can move out of a "fronted" P1, (73a), as well as binding the argument of a CF P1, (73b).

(73) a. Aiza no manasa lamba isan'andro Rasoa?
        Where Foc wash+AF clothes every'day Rasoa
        Where does Rasoa was clothes every day?

    b. Aiza no anasan-dRasoa lamba isan'andro?
        Where Foc wash+CF-Rasoa.gen clothes every'day
        "*Where are clothes washed every day by Rasoa?*"

Second, if the DPnom+P1 order in Ss like (74a) reflects the underlying order then no movement of the P1 is needed and we should be able to extract all arguments. But attempts to do so are disastrous, (74b) and (74c).

(74) a. ny zazavavy manasa lobaka   (ary ny zazalahy manasa pataloha)
        the girls are washing (AF) shirts (and the boys are washing(AF) pants)

    b.1 *Lobaka no ny zazavavy manasa (ary pataloha no ny zazalahy manasa)
            shirt Foc the girls are washing (and pants Foc the boys are washing)
    b.2   Lobaka no sasan'ny zazavavy   (ary pataloha no sasan'ny zazalahy)
        shirt Foc wash+TF'the girls   (and pants Foc wash+TF'the boys)

c.1 *ny lobaka ny zazavavy manasa
        the shirts (which) the girls are washing
c.2   ny lobaka sasan'ny zazavavy
       the shirts wash+TF'the girls
        the shirts which the girls are washing

One might dismiss these examples on the grounds that the DPnom+P1 order is a root clause phenomenon (which rules out (74c.1), less clearly (74b.1)). However there is one subordinate conjunction, *noho* 'because of' which requires a DP complement, yet on occasion takes whole Ss as complements. It "compromises" by forcing a DPnom+P1 order, in distinction to *satria* 'because' which takes the normal predicate first order, (75). Here is an example from a Malagasy romance novel:

(75) (These two groups were families who loved each other greatly)
   ...noho Ramavo manambady ny zanaky ny mpirahalahy ...
   ...because Ramavo.nom married the son of the brothers ...

But again in such a case moving a non-P1 argument is not possible:

a. noho Ramavo manambady ilay zanaka
    because Ramavo married that (aforementioned) son

b. *noho ilay zanaka no Ramavo manambady
    because that son Foc Ramavo married
    because it was that son that Ramavo married

In the absence of a concrete proposal we shall not further consider this option.

    To close, we remind the reader that we need no A movement to derive Ss which are voice variants, and we use none of the apparatus of extraction for relative clauses, we just concatenate an N and a P1. Surely our approach wins on grounds of simplicity and descriptive adequacy. And surely considerations of economy and elegance must push us to query the merit of generating DPs and verbs in wrong positions and then contriving ways to move them. At least our BG approach provides an explicit alternative.

5.0 Speculations on Generalizing Two natural ways of generalizing the approach taken here should be pursued: (1) To what extent does our non-movement analysis of RCs extend to other Malagasy structure types often treated as involving movement and variable binding? And (2), to what extent can voice variants, such as actives and passives, in other languages be treated without A movement as done here? In **5.1 − 5.6** below we summarize a variety of potential movement structures in Malagasy indicating briefly how we might directly generate them using P1s. But in each case not only would the analysis have to be justified, the structure would have to be semantically interpreted, as we did for RCs in showing how selectional restrictions were satisfied without movement. Moreover each structure type we mention comports its own specific idiosyncracies, so in each case a thorough analysis would be as long as the one presented here for RCs.

5.1 Embedded Questions generally have the form [*izay* + (N) + P1], closer in form to RCs rather than to direct questions (below).

Tsy fantatro [izay [P1 nanenjika azy]]   Tsy fantatro [izay [P1 nenjehin-dRabe]]
not known-by-me who chased+AF her   not known-by-me who Rabe chased+TF
I don't know who chased her   *I don't know who Rabe chased*

Tsy fantatro [izay mpianatra [P1 nanenjika azy]]
not known-by-me who student chased+AF her
I don't know which student chased her

5.2 Cleft Ss have the form [DP no P1]

Rabe no [P1 nanenjika azy]   Rasoa no [P1 nenjehin-dRabe]
Rabe Foc chased+AF her.acc   Rasoa Foc chased+TF+Rabe
It was Rabe who chased her   *It was Rasoa who Rabe chased*

As noted above, adjuncts may cleft without adjusting the verb voice to form a P1 after the *no*. The P1 argument may optionally tag along (the Bodyguard Condition).

Any ambanivohitra izy no mipetraka
there countryside   3nom Foc pres+AF+sit
It is in the country that he lives

5.3 Wh-Questions are formed as Clefts with a *wh*-question word.replacing the initial DP (the intersection of the set of constituent question words and the relativizers is empty).

Iza no [P1 nanenjika azy]?   Iza no [P1 nenjehin-dRabe]?
Who Foc chased+AF her.acc   Who Foc chased+TF+Rabe
Who was it who chased her?   *Who was it who Rabe chased?*

Aiza izy no mipetraka?
Where he Foc pres+AF+sit
Where is it that he lives?

5.4 Raising to Object constructions have the form [P1V + DPacc + ho + P1] + DPnom, where the DPacc is interpreted as the argument of the P1

a. [P1 mihevitra an-dRabe ho [P1 nanenjika azy]] Rasoa
     thinks acc-Rabe comp chased+AF her Rasoa
     Rasoa thinks Rabe to have chased her

b. [P1 mihevitra an-dRasoa ho [P1 nenjehin-dRabe]] aho
   pres+AF+think acc-Rasoa comp chased+TF+Rabe 1.sg.nom
     I think Rasoa to have been chased by Rabe

c. [P1 ∅+heveriko ho[P1 nenjehin-dRabe]] Rasoa
     pres+think+TF+2sg.gen comp past+chase+TF+Rabe.gen Rasoa
     Rasoa is thought by me to have been chased by Rabe

5.5 Object Control: [P1V + DPacc + P1] DPnom where, again, the DPacc is the understood argument of the following P1.

[P1 manampy azy hametraka ny kodiarana] Rabe
     help+AF him fut+change the wheel Rabe
     Rabe is helping him change the tire

5.6 Tensed Nominalization: [DPDet + P1 + (Det)]

a. Tsy tian-dRakoto ny [P1 n+a+tao+nao azy]
   not like+TF-Rakoto.gen the past+TF+do+2.sg.gen 3.acc
     Rakoto doesn't like what you did to him

b. araka izao [[ahitana-∅i sy ahafantara-nareoi] ahy] izao   (= (53))
according this [[[see+CF and know+CF]+2.gen.pl] 1.sg.acc] this
according to that which you see and (you) understand of me

5.7 Passives in other languages can be enlighteningly treated as predicate level
derivational operations rather than clause level ones (Keenan 1980, 1985; Timberlake
1982; Keenan & Timberlake 1985). It can offer a semantic solution to a central challenge
in Minimalist accounts of passive (Collins 2004), namely how to account for theta role
assignment and the selectional restrictions of Agent Phrases in passives.

European Passive is naturally represented as a way of deriving P1s from P2s, and
more generally, Pn's from Pn+1's: Pass(Pn+1) = Pn (a more explicit definition is given
shortly). Note that in distinction to Malagasy voice functions which preserve valency
(but see footnote 11), Passive in European reduces it.

If "passive" is a property of predicates then verbs can be classed as +/– passive, and
it is not surprising that, as with other classes of verbs, they may select their auxiliaries
("be" in Romance, "become" in German and Dutch, etc.) and even have their own
agreement paradigms, as in Greek and Latin (77):

(77)   Active present indicative   Passive present indicative

'I love'   amo   amamus   'I am loved'   amor   amamur
             amas   amantis   amaris   amamini
             amat   amant   amatur   amantur

And since P1 ⇒ P0 (Sentence) is one case of Passive we expect to find P1s passivized
with the same morphology as P2s and we do in German, Lithuanian, Turkish and Latin:

(78) a. aliquis currit   b. curritur (ab aliquo)   (Latin)
        someone run(Act)   Run(Pass) (by someone)
        Someone is running   *Running is being done (by someone)*

Now unless we restrict Pass to lexical Pn+1's it will apply to syntactically complex ones,
ones including adverbs and PPs within its scope, which is desirable. Goal locatives for
example, such as *ad astra* 'to the stars' in (79), select their verb, making sense when
combined with verbs of motion but not with stative P1s (*Marcus est agricola ad astra*
'Marcus is a farmer to the stars' is nonesense). Thus in (79) we want the P1 passivized
to be *ire ad astra* 'go to the stars', not just *ire* 'go', with *ad astra* 'to the stars' combining
with the P0 Pass(*ire*):

(79) a.   Sic itur ad astra   =   [P0 [Adv sic] + [P0 Pass([P1 ire ad astra)]]]
          Thus go+Pass to stars
          Thus it is gone (one goes) to the stars

Similarly in English agentive *by*-phrases are PPs which we want in the scope of Passive. To motivate this from English consider that we want instrumental PPs in the scope of Passive. When such PPs combine with P1s to form P1s they predicate a "use" property of its argument:

(80) Dana dug with a shovel = Dana used a shovel to dig
     Sasha sang with a microphone = Sasha used a microphone to sing

But semantic equality badly fails in (81):

(81) The can was opened with a screwdriver ≠ *The can used a screwdriver to open

So if the instrumental PP *with a screwdriver* combined with the P1 Pass(*open*) we obtain incorrect predictions. But we obtain correct ones if PPs combine with P2s to form P2s, and as before, they are interpreted as restricting, mapping each P2 R to a subset of the pairs (x,y) which stand in the relation R:

(82) ((WITH z)(OPEN)) ≤ OPEN.   [i.e. (WITH z)(OPEN)(y)(x) ⊨ OPEN(y)(x)]

This assumption, together with the naive semantic analysis of Passive in (83) yields correct judgments of entailment[2].

(83) For P a P2, $\llbracket$Pass(P)$\rrbracket$(y) = True iff for some x, $\llbracket$P$\rrbracket$(y)(x) = True

Now, writing denotations in upper case, observe that our analysis of (81) yields correct judgments of entailment:

(82) PASS((WITH z)(OPEN))(y) = True iff for some x ((WITH z)(OPEN))(y)(x) = True

The righthand side above entails that for some x, OPEN(y)(x) = True since (WITH z) is restricting. So on this semantics we obtain the entailment in (83).

(83) The can was opened with a screwdriver ⊨ The can was opened

And we do not invite any "use" type predication since the instrumental phrase has not combined with a P1 forming a P1. Now with these preliminaries let us give the PP semantics for agentive *by*-phrases. Like other PPs we treat them as combining with P2s to form P2s. All that is unusual about them is that BY is a logical Preposition:

(84) For all P2 denotations P, (BY(z))(P)(y)(x) = P(y)(z)

(84) says that the argument z of BY functions as the last argument of P and thus bears to P whatever theta relation P requires. Now let us verify that Passive works as it should:

(85) a. John was praised by Bill

b. PASS((BY b)(PRAISE))(j) = True iff

    for some x, (BY b)(PRAISE)(j)(x) = T   Def PASS (83)
    for some x, PRAISE(j)(b) = T   Def BY (84)
    PRAISE(j)(b) = T   Def existential quantification

The first and last lines above guarantee that the object b of BY satisfies the theta and

selection restrictions of PRAISE since ⟦praise john⟧ is predicated of it. So, similar to Malagasy, these semantic properties are handled in the semantic interpretation.

Lastly observe that our treatment of Passive commits us to recursive morphology since Pass applies to proper phrases but the distinctive passive morphology in English just affects verbs. A detailed definition would depend (recursively) on the various ways of building P2s in English. But here is a first statement giving the flavor of the definition. We use a strict morphological function EN which maps verbal roots to their past (passive) participles. For example, EN(hit) = hit, EN(break) = broken, EN(praise) = praised, etc. Then the recursion clauses are given as in (b):

(86) **Def** a. Pass([P2 d]) = be + EN([P2 d]),   where
      b. EN([P2 d] + [PP e])   = EN([P2 d]) + ([PP e])   and
        EN([P2 d]) and/or ([P2 d])   =   EN([P2 d]) and/or EN([P2 d])

In this way the *-en* past participle morphology works its way down the structure of a P2 until it gets to the lexical items where its values are given. Ignoring tense the computation of (87) is given below it.

(87) John was robbed and beaten by an angry linguist

    John Pass([rob and beat] [by an angry linguist])
    John be EN([rob and beat] [by an angry linguist])
    John be EN([rob and beat]) [by an angry linguist]
    John be EN(rob) and EN(beat) [by an angry linguist]
    John be robbed and beaten by an angry linguist

This sketch is hardly problem free, and it doesn't pretend to resolve all the standard issues concerning passives. But it does show how to account for semantic notions such as theta role assignment and selectional restrictions in semantic interpretation.

References

Abinal, R.P. et V. Malzac (1963) Dictionnaire Malgache-Français. Mission
   Catholique de Tananarive. Paris: Editions Maritimes et d'Outre-Mer (reprint
   from 1888 edition).
Aldridge, Edith. 2002. Nominalization and WH-movement in Seediq and Tagalog.
   Language and Linguistics 3.2 Special Issue: Nominalization in Formosan Languages,
   Elizabeth Zeitoun (ed). Academia Sinica. Pp. 393 – 427.

Baker, Mark C. 1988. *Incorporation A Theory of Grammatical Function Changing* University of Chicago Press

[BJR] Baker, Mark, Kyle Johnson and Ian Roberts. 1989. Passive arguments raised. *Linguistic Inquiry* 20:219 – 251.

Chomsky, Noam. 1955. *The Minimalist Program*. MIT Press.

Collins, Chris. 2004. A smuggling approach to passive in English. Ms. Cornell University.

Embick, David. 1997. In H. Harley (ed) *Proceedings of the Penn/MIT Workshop on Aspect, Argument Structure, and Events* MITWPL

Erwin, Sean. 1996. Quantity and moras: An amicable separation. In *The Structure of Malagasy* vol I. M. Pearson and I. Paul (eds). Dept. of Linguistics, UCLA. pp. 2 –31.

De Guzman, V. 1988. Antipassives and causatives in Ilokano: Evidence for an ergative analysis. In *Studies in Austronesian Linguistics*, ed. Richard McGinn, 323 – 345. Athens, OH. Ohio University Center for International Studies.

Gerdts, D. 1988. Antipassives and causatives in Ilokano: Evidence for an ergative analysis. In *Studies in Austronesian Linguistics*, ed. Richard McGinn, 295 – 321. Athens, OH. Ohio University Center for International Studies.

Gildersleeve, B.L. and G. Lodge. 1913. *Gildersleeve's Latin Grammar* Macmillan & Co. Ltd. London.

[GHT] 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* 10:375 – 414.

Huang, Lillian. 2002. Nominalization in Mayrinax Atayal. *Language and Linguistics 3.2* Special Issue: Nominalization in Formosan Languages, Elizabeth Zeitoun (ed). Academia Sinica. pp 197 – 227.

Hyams, Nina, Cecile Manorohanta and Dimitris Ntelitheos. 2004. The Acquisition of the Malagasy Voicing System. Dept. of Linguistics, UCLA.

Keenan, Edward L. 1972. Relative Clause Formation in Malagasy (and some related and some not so related languages). *The Chicago Which Hunt* University of Chicago.

Keenan, Edward L. 1976. Remarkable subjects in Malagasy. In *Subject and Topic* C. Li (ed) Academic Press.

Keenan, Edward L. 1980. Passive phrasal (not sentential or lexical). In *Lexical Grammar*, T. Hoekstra, H. Van der Hulst, and M. Moortgat (eds). Foris.

Keenan, Edward L. 1985. Passive in the world's languages. In *Language Typology and Syntactic Description* Vol 1. T. Shopen (ed). CUP

Keenan, Edward L. 1995. Predicate-argument structures in Malagasy. In *Grammatical Relations* C. Burgess, K. Dziwirek, and D. Gerdts (eds). CSLI. Stanford. pp.171 – 217

Keenan, Edward L. 2000. Morphology is structure: A Malagasy test case. In *Formal Issues in Austronesian Linguistics* I. Paul, V. Phillips, and L. Travis (eds). Kluwer. pp.27 – 49.

[K&R] Keenan, Edward L. and Jean Paulin Razafimamonjy. 2004. Reciprocals in Malagasy. in *Oceanic Linguistics* vol. 43.1:177 – 207.

Keenan, Edward L. and Alan Timberlake. 1985. Predicate formation rules in universal grammar. J. Goldberg et al (eds) *Proc. of WCCFL* 4, Stanford University.

[K&M 01] Keenan, Edward L. and Cecile Manorohanta. 2001. A quantitative study of voice in Malagasy. in *Oceanic Linguistics* vol 40:67 – 85.

[K & M] Keenan, Edward L. and Cecile Manorohanta. 2004. Malagasy Clause Structure and Language Acquisition. *Proceedings of AFLA 11,* ZAS, Berlin. Paul Law (ed). Pp. 178 – 202.

[K&P] Keenan, Edward L. and Miriam Polinsky. 1998. Malagasy Morphology. In *The Handbook of Morphology*. A. Spencer and A. Zwicky (eds) Blackwell. pp. 563 – 624.

Keenan, Edward L. and Baholy Ralalaoherivony. 2000. Raising from NP in Malagasy. *Lingvisticai Investigationes* pp 1 – 45.

[K&S] Keenan, Edward L. and Edward P. Stabler. 2003. *Bare Grammar: Lectures on Linguistic Invariants*. CSLI, Stanford.

Kroeger, Paul. 1988. Verbal focus in Kimaragang. In *Papers in Western Austronesian Linguistics* No. 3:217 – 40. Pacific Linguistics, A-78.

Kroeger, Paul. 1993. *Phrase Structure and Grammatical Relations in Tagalog*. CSLI, Stanford.

Paul, Ileana. 1996a. The active marker and nasals in Malagasy. In Pearson & Paul, *The Structure of Malagasy* Vol I, UCLA Occasional Papers in Linguistics 17. Pp 49 – 57.

Paul, Ileana. 1996b. The Malagasy genitive. In Pearson & Paul, *The Structure of Malagasy* Vol I, UCLA Occasional Papers in Linguistics 17. Pp 76 – 92.

Paul, Ileana. 1999. *Malagasy Clause Structure*. PhD Diss. Dept. of Linguistics, McGill University

Paul, Ileana and Lucie Rabaovololona. 1998. Raising to object in Malagasy. In *The Structure of Malagasy* vol II I. Paul (ed). Dept. of Linguistics, UCLA. pp.50 – 65.

Payne, Thomas. 1982. Role and reference related subject properties and ergativity in Yup'ip Eskimo and Tagalog. *Studies in Language* 6.1:75 – 106.

Pearson, Matthew. 2001. *The Clause Structure of Malagasy: A Minimalist Approach* UCLA Dissertations in Linguistics. No. 21.

Pearson, Matthew. 2005. The Malagasy subject/topic as an A' element. *Natural Language and Linguistic Theory* 23:381 – 457.

Phillips, Vivianne. 2000. The interaction between prefix and root: the case of *Maha-* in Malagasy. In *Formal Issues in Austronesian Linguistics* I. Paul, V.Phillips, L. Travis (eds). Kluwer, Dordrecht. pp. 85 – 105.

Prentice, D.J. 1971. *The Murut Languages of Sabah* Pacific Linguistics Series C – No.18. The Australian National University. Canberra, Australia.

Rabenilaina, Roger-Bruno. 1993. *Le verb malgache – Etude des formes predicatives*. Ambozontany: Fianarantsoa, Madagascar.

Rackowski, Andrea and Lisa Travis. 2000. V-Initial Languages: X or XP movement and adverbial placement. In *The Syntax of Verb Initial Languages*. Andrew Carnie and Eithne Guilfoyle (eds) OUP Pp. 117 – 143

Rackowski, Andrea and Norvin Richards. (nd). Phase edge and extraction: a Tagalog case study. Ms. MIT.

Rajaonarimanana, Narivelo. 1995. *Dictionnaire du malgache contemporain* Karthala, Paris.

Rahajarizafy, Antoine. 1960. *Essai sur la grammaire malgache* Imprimerie Catholique, Antanimena, Tananarive

[RR] Rajemisa-Raolison, Regis. 1971. *Grammaire malgache*. Fianarantsoa.

Richards, Norvin. 2000. Another look at Tagalog subjects. In I. Paul, V. Phillips, and L. Travis (eds). *Formal Issues in Austronesian Linguistics*. Dordrecht: Kluwer

Sabel, Joachim. 2005. Grammatical Functions in Malagasy Syntax and the "Complements only" Restriction. Presented at AFLA XII, UCLA

Schachter, Paul. 1976. The subject in Philippine languages: topic, actor, actor- topic, or none of the above. *Subject and Topic* Charles Li (ed) . Academic Press. pp. 491 – 518.

Schachter, P. and Fe Otanes.1972. *Tagalog Reference Grammar*. Berkeley/Los Angeles, Ca. UC Press.   .

Timberlake, Alan. 1982. The impersonal passive in Lithuanian. *Proc of BLS* 8. UC Berkeley.

Travis, Lisa. 2005. VP-internal structure in a VOS language. Ms. McGill University

Wechsler, Stephen and I.Wayan Arka. 1998. Syntactic Ergativity in Balinese. In *Natural Language and Linguistic Theory* 16:387 – 441.

Endnotes

† "A 'perfect language' should meet the condition of inclusiveness: any structure formed by the computation...is constituted of elements already present in the lexical items selected...(Chomsky 1995:228).

Thanks to audiences at UCLA and MIT, and especially to Dimitris Ntelitheos and Ed Stabler for constructive discussion of the issues raised in this paper. Neither should be held responsible for my apostasies.

1. E.g. Tagalog (Schachter and Otanes 1972, Kroger 1993), Timugon (Prentice 1971), and Kimaragang (Kroeger 1988), Balinese (Wechsler and Arka 1998), Atayal (Huang, 2002).

2. Formally, given a domain E of (possibly abstract) objects, an *n-ary relation* R *over* E is just a set of n-tuples of elements of E. Write $\mathbf{R}_n$ for the set of n-ary relations over E. Then a theta role such as *Theme* is, extensionally, a subset of $E \times \bigcup_{1 \le n} \mathbf{R}_n$.

3. Some major orthography-phonology correspondences are: *o* = /u/; word internal *-i-* = word final *-y* = /i/; *tr* is a voiceless prepalatal affricate, *dr* (or *dR*) its voiced counterpart; *j* = /dz/ with *ts* its voiceless counterpart. nC = ^{n}C, mC = ^{m}C

4. Here are three symmetries between AF and non-AF verbs. They differentiate non-AF ones from English passives.

Imperatives The AF and non-AF verbs discussed all have imperative forms with the addressee phrase missed out. That will be the nominative for AF imperative, (i.a), but the genitive complement of non-AF forms in *a-*, *-Vna* and *–ana*, (i.b,c,d):

(i) a. manolóra (m+aN+tólotra+a) vary azy!   (Stress marked)
      offer+AF pres+aN+offer+imp rice them
      Offer them rice

   b. tolóry (tólotra+y) vary izy
      offer+GF offer+imp rice they
       Offer them rice
  c. atolóry (a+tólotra+y) azy ny vary
      offer TF+offer+imp them the rice
      Offer them the rice

   d. anolóry vary azy ny lovia vaovao
      offer+CF rice them the dishes new
      Use the new dishes to serve them rice

As always the P1 argument, when present, must be definite. Imperative suffixes differ for AF verbs and non-AF ones, being the same for *a-*, *-Vna* and *-ana* verbs.

Selection Both AF and non-AF forms are selected by modal auxiliaries. *mahazo* 'permitted' and *mahay* 'able' select AF verbs, their TF roots *azo* and *hay* select non-AF verbs. Replacing AF by non-AF verbs in (ii) and (iii) yields ungrammaticality.

(ii) a. mahazo manao (m+aN+tao(v)) izany aho
        permitted pres+aN+do that I
        I can do that

   b. Tsy azoko (azo + ko) atao izany
      Not can + 1.sg.gen a+do that
      I can't do that

(iii) a. mahay manoratra (m+aN+soratra) amin'io penina io aho
        pres+can write pres+aN+write with'that pen that I
        I am able to write with that pen

   b. Tsy haiko (hay + ko) anoratana (aN+soratra+ana) io penina io
      not able+1.sg.gen aN+write+ana that pen that
      I can't write with that pen

Control With verbs of desire and intent control is vested in the Agent not the nominative (the two coinciding with AF verbs):

(iv) a. n+i+kasa (ny) hamaky (h+aN+vaky) io boky io aho
        past+i+intend (the) fut+aN+read that book that I
         I intended to read that book

b. nokasaiko (no+kasa+ina+ko) (ny) hovakina (ho+vaky+ina) io boky io

past+intend+Vna+1.sg.gen (the) fut+read+ina that book that

I intended to read that book

5. *Izy* can be augmented to force a plural reading: *izy* ireo '3 dem+pl' = they; *izy mivady* '3 spouse' = they spouses; *izy roa lahy* '3 two man' = they two men.
6. We used pronominal replacement as a test for case. In *He left on Monday*, *Monday* was not counted accusative (in English) or genitive (in Malagasy translation) as pronominal replacement is unnatural **He left on it*. Counting such cases would have increased the proportion of accusatives in English and genitives in Malagasy.

7. Here is a hopefully helpful analogy. Consider the closure of the "lexicon" {0} under the ADD 5 function. This "language" is the set {0, 5, 10, 15, 20, ... }. Now consider the closure of {2} under that *same* function. Its language, disjoint from the first, is {2, 7, 12, 17, 22, ... }. So the lexicons and the languages differ, but the generating functions are the same.

8. Standard grammars of Malagasy (e.g. Rahajarizafy 1960) give the *aN-* and *i-* prefixes as *man* and *mi*. The initial *m* is specific to present tense AF verbs.

9. We ignore here the many constraints determining case and order of the two internal arguments. Our only point is that AN builds the core ditransitive verbs.

10. Note that AN is a single function with a fairly large domain, one including unary, binary and ternary roots. It maps unary roots to P1s, binary roots to P2s and ternary roots to P3s. So its value is conditioned by its argument, just as for example a rounding off function maps m.n to m if $n < 5$ and to m+1 if $n \geq 5$.

11. An important TF prefix we do not consider here is *voa-*. It is a TF prefix applying just to roots. It builds perfective verbs, the Agent is often absent, and the derived verb does not form an imperative (*a-* and *-Vna* verbs do, as do AF verbs). They occur much less frequently than TF verbs in *-Vna* (Keenan and Manorohanta 2001). *Voa-* verbs are the closest to English passives.

i. a. Inona no laza+ina+Rabe (lazain-dRabe)   b. Izany no voa+laza+ny (voalazany)

   what Foc say+TF+Rabe.gen   that Foc TF+laza+3gen

   What is Rabe saying?   That's what he said

   c. Voa+kapoka (voakapoka) ilay alika

   TF   beat   that (aforementioned) dog

   That dog was beaten

12. Note that UTAH does not hold on our analysis. The Agent phrases (Theme phrases) in (2a,b) were never in structurally comparable positions. Nonetheless our analysis entails that the genitive DP in (2b) and the nominative one in (2a) bear the role of Agent in their clauses.

13. The most common TF suffix is *-ina*, but such suffixes sometimes show up as (*a*)*na* and occasionally as *-ena*. These are likely best analyzed as variants conditioned by the choice of root, but then one or two cases of contrast still remain. We do not decide this issue here.

14. There are two cases we do not discuss here in which the P1 argument of an ANA verb may be subcategorized by the original AF verb. (1) The P1 argument may be the Theme understood partitively, a case discussed in Paul (1999). Thus "Please buy (me) some paper" could be rendered *Mba ividiano ny taratasy* with a CF imperative and definite P1 argument *ny taratasy* 'the paper'. Second, occasionally CF forms get pressed into duty when appropriate INA forms lack. Thus reciprocal verbs, with prefix *if-*, are not in the domain of INA, so to talk about 'the many things we taught each other' we would say *ny zavatra betsaka nifampianarantsika* with a CF form of "reciprocally caused to learn"

15. Following Erwin (1996) and Pearson (2001) we treat this "epenthetic" consonant as part of the root. It remains in derived forms until the CV template to which Malagasy conforms is applied. At that point it drops if no suffix provides a vowel to follow it. So from the root *taov-* 'do' we derive indicative *manao* (AF) and *atao* (TF), but imperative *manaova* and *ataovy* and CF *anaovana*. AF *aN-* verbs may be ambiguous as the aN- prefix sometimes destroys (or mutates) a root initial consonant. But the TF and CF forms may be distinct as they preserve the root final consonant. Thus *mamono* means 'envelopes' when derived from *fonos-* and 'hits, kills' when derived from *vono*. The *-ina* TF forms are distinct: *fonosina* and *vonoina*, and the CF forms derived from the AF forms are distinct: *amonosana* vs *amonoana*.

16. Occasionally the invariable particle *izay* occurs between the N and the P1.

17. When *izay* occurs (between the N and P1) we interpret it as follows: IZAY(P)(A) = A $\cap \{x | P(x) = \text{True}\}$. A is a N denotation, P a P1 one.

18. In the BG format Compositionality has a particularly simple formulation: writing $[\![d]\!]$ for the interpretation of *d* in an arbitrary model, Compositionality says when an expression t is built by applying a generating function (rule) F to expressions $t_1,...,t_n$ then there is a function

F' taking n-tuples of t_i denotations as arguments satisfying $[\![F(t_1,...,t_n)]\!] = F'([\![t_1]\!], ...,[\![t_n]\!])$.

19. Our claim is preserved under copy theories of movement. If anything like *wh...John chased wh...* is created the learner must interpret expressions other than ones present in nuclear Ss.

20. Note the VSO order here. We assume a rightward CP extraposition rule.

21. There is one other, not mentioned above as it does translate well into English. Namely, as in Tagalog (Rackowski & Richards nd), when the P1 argument is *fa* + S, a "CP", either the entire CP can be regarded as the P1 argument or the P1 argument of the S can be regarded as the overall P1 argument. The two bracketings are illustrated below. The polar question word *ve?* may immediately follow the right bracket in each case. (ii) illustrates Wh-question in each case:

(i) a. [P1 Heverin-dRabe] fa nangalatra ny omby Rakoto

   b. [P1 Heverin-dRabe fa nangalatra ny omby] Rakoto

        Think(TF)-Rabe.gen that stole(AF) the cow Rakoto

        Rabe thinks that Rakoto stole the cow

(ii) a. Inona no heverin-dRabe?

        What Foc think+TF+Rabe.gen

        What does Rabe think?

   b. Iza no heverin-dRabe fa nangalatra ny omby?   (*ho* is preferred to *fa* here)

        Who Foc think+Rabe that stole+AF the cow

        Who does Rabe think stole the cow?

If the matrix verb is AF its nominative DP can be relativized but, as expected, not its accusative CP or anything in that CP. Keenan (1976) suggests that Ss like (i.b) are instances of Subject-to-Subject Raising; Sabel (2005) offers a different, more contemporary extraction based view that would relate extraction asymmetries in W. Austronesian to English.

22. This remains true in R&T 2000 and Travis 2005 which do not explicitly discuss advancements and case assignment, their concerns being adverb order and ellipsis. I assume that voice affixes continue to assign case as in GHT. Travis (2005) does use (to good advantage) constituents like (i) below in which an unnamed 'F' plays the role of INFL in GHT.

(i)   FP

              F   VP

         V   F   Agt   V'

                                              VP

                                        Theme$_{indef}$ V'
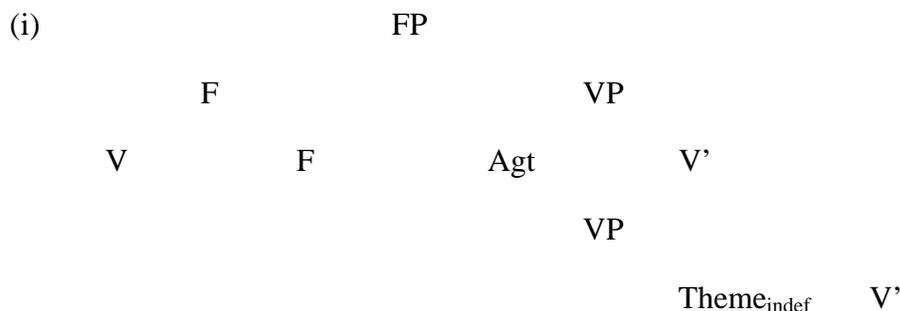
What I don't see here (and elsewhere) is how even the highest F can access something properly inside its sister. F is looking at a tenseless clause – a verb with all its arguments present (or bound from outside). Such objects are not of the right type to be assigned case, so we would reject an analysis in which the "VP" is assigned case and that case feature "trickles down" to an appropriate DP.

23. Pearson writes just -*Vn* where we use -*Vna* for continuity with our previous discussion. Similarly we write -*ana* where Pearson writes -*an*.

24. Since the root moves to become a sister to -*Vna* in GHT they could impose a constraint limiting which roots so move. But this is unenlightening, as the condition is a purely lexical one, not a structural one like blocking movement over a subject, and should be treated as a lexical property of either the root or the affix. Similarly a filter like *(*tosika* + -*ina*) filtering each inadmissible cooccurrence is as ad hoc as it comes.

25. Travis (2005) suggests a significant enrichment of the GHT structures to generate coordinations such as (71) by adding an unnamed constituent noted FP or XP. The new constituent would not make the non-AF verb and its Agent sisters, but it would put them in a constituent which can exclude definite Themes thereby permitting coordinations as in (71). It is not clear that it would permit coordinations as in (70) but the observation that definite and indefinite Themes behave differently seems insightful. For the moment I leave these issues unresolved pending further empirical work and clarification of the motivation and role of the newly added constituent.

26. The general statement is:

For P a P_{n+1}, $[\![Pass(P)]\!](y_n)...(y_1)$ = True iff for some x, $P(y_1)(y_n)...(x)$ = True