A Quantitative study of Voice in Malagasy[†]

Edward L. Keenan   Cecile Manorohanta
UCLA   Université Nord, Antsiranana Madagascar
keenan@humnet.ucla.edu

This paper is a quantitative study of the voice system in Malagasy (W. Austronesian; Madagascar). We show that non-active verbs in Malagasy have a very different distribution in texts than non-active verbs in English, German and Dutch: they occur far more frequently and they typically present Agent phrases. This we claim reflects the very different role of the voicing system in the grammars of W. Austronesian and W. European languages.

   Our presentation is organized as follows: Part I reviews the voice system of Malagasy, classifying the various voice forms into Active vs non-Active; the latter divided in Passive, and Circumstantial. Part II presents the results of our text study and Part III draws some conclusions regarding the nature of the voicing system in Malagasy.

1. The voice system of Malagasy

We first exemplify the Malagasy voice forms. Examples are given in the standard orthography, augmented when helpful by '+' to indicate morpheme boundaries, ´ to mark main stress, and block parentheses to indicate constituency. Pronounced forms are noted in round parentheses next to their morphemic decompositions (when not simply the concatenation of the latter).

(1) a. [n+i+vídy akanjo hoan'ny zaza] i Vao
        past+active+buy clothes for'the child art Vao
        Vao bought clothes for the child

   b. [no+vidy+ina+i Vao (novidín'i Vao) hoan'ny zaza] ny akanjo
      past+buy+pass+art Vao for'the child the clothes
      The clothes were bought by Vao for the child

   c. [n+i+vidy+anana+i Vao (nividiánan'i Vao) akanjo] ny zaza
      past+act+buy+circ+art Vao clothes the child
      The child was bought clothes for by Vao

   There is massive evidence (Keenan 1972, 1995, Pearson 2000) that the bracketed strings in (1a,b,c) are constituents, called here *Predicate Phrases* (PredPh). For example to form the Yes-No questions corresponding to (1a,b,c) it suffices to insert the particle *ve* at the right edge of the

PredPh. Insertion internal to the PredPh is ungrammatical.

Another example: the only relativizable position in (1a,b,c) is that of the NP external to the PredPh (henceforth the *external argument*, EA, a usage we take from Pearson 2000). Thus (2a,b,c) are grammatical; relativizing any other NP, as in (3), is not.

(2) a. ny olona (izay) nividy akanjo hoan'ny zaza
the person (that) bought clothes for the child

b. ny akanjo (izay) novidin'i Vao hoan'ny zaza
the clothes (that) were bought by Vao for the child

c. ny zaza (izay) nividianan'i Vao akanjo
the child (that) were bought+for by Vao clothes

(3) *ny akanjo (izay) nividy hoan'ny zaza i Vao
the clothes that bought for the child Vao

Reading (3) through active *nividy* we understand that it is the clothes that bought something. The rest of (3) lacks a Theme and presents an NP, *Vao*, which lacks a semantic role. Similar ungrammaticality results when attempting to relativize NPs other than the EA from (1b and c). The EA in (1a,b,c) is replaceable with pronominal forms drawn from the following series, which we call *nominative*:

(4) *aho* I; *izahay* we excl; *isika* we incl; *ianao* you sg; *ianareo* you pl; *izy* 3 sg or pl.

This series contrasts with those in (5a), called here *accusative*, and (5b), called *genitive*. Items are given in the same order as in (4).

(5) a. *ahy; anay; antsika; anao, anareo, azy*
b. *-ko, -nay, -(n)tsika, -nao, -nareo, -ny*

The accusative forms replace Patient and Theme arguments internal to the PredPh, as in (1a,c). The genitive forms are used for the Agent phrases in (1b,c) and for possessors in general:

(6) ny trano 'the house';   ny trano+ko 'my house';
ny trano+ny 'his house'   ny trano+nao 'your sg house'
ny tranon'i Vao 'Vao's house'   ny tranon-dRabe 'Rabe's house'

Note the nasals in the last line, and see Paul (1996) for the complicated morphophonemics involved in binding genitives to their hosts,

Definition Verbs which take a genitive argument will be called *non-active*; those which do not, *active*. *Passives* are non-actives which are roots or built by affixing roots (possibly reduplicated). *Circumstantial* verbs are non-active ones built by affixing active ones. □

These three classes of verbs are distinguished by properties in addition to those we chose as definitional. In particular, present tense active verbs (and their imperatives) take an *m-* prefix, whereas present tense non-active verbs (and their imperatives) have no prefix. (All voice forms except roots prefix *n(o)-* to mark past tense; all use *h(o)-* to mark future). Active verbs form imperatives by suffixing *-a*, whereas non-actives suffixe *o* (= /u/) or *y* (= /i/). Suffixation in general induces rightward stress shift in all voices and moods. Finally, the highest argument in the theta grid of an active verb is the external, nominative, one; it is the genitive one with non-active verbs. Keenan (2000) shows that the genitive Agent Phrase in non-actives forms a tight constituent with the verb to the exclusion of its other complements. The EAs of passives correspond to internal arguments of actives; those of circumstantial verbs typically have oblique semantic roles like Benefactive, Instrument, Locative, Temporal, Manner, ...

Below we exhibit the different morphologies in each of the three verb classes. Rabenilaina (1998) and K&P (Keenan & Polinsky 1998) give more detailed exemplification. Rabenilaina (1993), Keenan (1995) and Pearson (2000) discuss the grammatical role of these voices.

1.1 Passives

1.1.1 *root passives*, in which the root itself suffixes genitive Agent phrases directly, as in (7a,b) are morphologically the simplest verb forms. They form imperatives as indicated, (7c).

(7) a. Hadino+ny ny anarako
        forgotten+3gen the name+my
        My name is/was forgotten by him

     b. Heno+ko izy   c. Henoy aho
        listened-to+1sg.gen he   listen+imp I
        He was listened to by me   *Listen to me*!

Root passives include *fantatra* 'known', *resy* 'defeated', *re* 'heard', *tadidy* 'remembered', *simba* 'damaged', *hita* 'seen', *tsinjo* 'perceived from above', *azo* 'received, understood', *vaky* 'broken', *tratra* 'caught' and at least twenty to thirty others[1].

---

[1]Here and later for purposes of immediate recognizability we give roots in their traditional, dictionary entry form. Formally on this view suffixation often induces an epenthetic consonant: root *la* ⇒ *lavina* 'is refused', *manda* 'refuses'. In fact we agree with Erwin (1966) that it is more

1.1.2 *Suffix passives* are the most common type of passive, formed by suffixing the root with -*(V)na*, where *V = i, e* or *a* with *i* the most common and vowel absence rare. The choice of vowel is largely conditioned by the choice of root, but Rahajarizafy (1960) lists a few roots which take more than one suffix with different meanings.

(8) a. no+vono+ina+Rabe (novonoin-dRabe) ilay akoho
        past+kill+pass+Rabe that (aforementioned) chicken
        That chicken was killed by Rabe

     b. ho+soratra+ana+Rasoa (ho soratan-dRasoa) ny taratasy
        fut+write+pass+Rasoa the letter
        The letter will be written by Rasoa

We note that several (but not all) passive roots also have suffix passives (*hadino / hadinoina*; *heno / henoina*; *tsinjo / tsinjovina*,...). Root passives are more stative, suffix forms more dynamic.

1.1.3 *Prefix passives* are of three varieties: *a-*, *voa-* and *tafa-*.

1.1.3.1 *a- passives* ressemble suffix passives in forming past tense with *n(o)-* and forming imperatives. Some roots, like *tao* 'do' only have *a-* passives: *atao* 'is done', *natoako* 'was done by me', etc. But commonly double object roots will have both a suffix passive, whose EA is Goal, and an *a-* prefix passive, whose EA is Theme. Thus from the root *róso* we have

(9) a. Ros*ó*+ana vary ny vahiny   b.   Ros*ó*+y vary izy
        advance+pass rice the guest(s)   advance+imp rice 3nom
        The guests are served rice   *Serve him rice* (lit: *be+served rice he*)

(10) a. A+r*ó*so ny vary   b.   A+ros*ó*+y ny vary
         pass+advance the rice   pass+advance+imp the rice
         The rice is served   *Serve the rice* (*nom*)   (lit: *be+served the rice*)

We note that in general prefixation does not induce stress shift.

1.1.3.2 *voa- passives* by contrast do not form imperatives or take past tense marking with *n(o)-*. Semantically they focus on the result of the action, understood to be completed, rather than the process, in distinction to the suffix passives and the *a-* prefix ones. *voa-* passives in Malagasy

---

accurate to include this consonant as part of the root and account for its absence in terms of a vowel final template.

are those which most ressemble those of English.

(11) a. Voa+laza+ko izany   b. Voa+fitika aho
   pass+say+1sg.gen that   pass+deceive I
   I said that (lit: *That was said by me*) *I was deceived*

1.1.3.3 *tafa- passives* are like *voa-* passives in focussing on the final state not the process denoted by the verb; they do not mark past tense with *n(o)-* and do not form imperatives. In distinction to *voa-* passives they imply that completion of the action was unexpected. Either the action was difficult for the agent, as in (12a), or unintended, even spontaneous, as in (12b), where an Agent phrase is not possible.

(12) a. Tafa+iditra+ko (tafiditro) ny omby
   pass+enter+1sg.gen the cow(s)
   I got the cows in (lit: *The cows were made to enter by me*)

  b. Tafa-tsangana tampoka aho
   tafa-stand suddenly I
   I stood up suddenly (in spite of myself)

1.1.4 *Infix passives* exist as relic structures in official Malagasy: *vaky* 'broken' ⇒ *vinaky* 'is broken'; *tapaka* 'cut' ⇒ *tinapaka* 'is cut'. The few roots that accept such infixation also have suffix passives (*vakina*, *tapahina*) and may even be root passives as well.

1.2 Actives

1.2.1 *root actives*, in distinction to root passives, only number about six and have restricted distributions and interpretations. Almost all are verbs of motion. *tamy* 'about to arrive' and *mby* 'arrived' do not have corresponding non-active forms and do not form *f-* nominals, a very productive derivational process with active verbs. *avy* 'come', *tonga* 'arrive' and *lasa* 'gone' are frequent and often used like auxiliaries, (13). None of these root actives use the distinctive *m-* in the present tense.

(13) Avy n+i+hinana vary (nihinam-bary) izy
   come past+act+eat-rice he
   He just ate

 All other actives are formed by prefixing roots. Following traditional Malagasy grammars (Rahajarizafy 1960, Rajemisa-Raolison 1971) and dictionaries (Abinal and Malzac 1888) we present these prefixes with the distinctively active *m-* (which only surfaces in the present tense).

1.2.2 *mi-* and *man-* prefixes are the most productive primary prefixes (ones that apply directly to roots). The *n-* on the latter may appear just as a prenasalization of the following (mutated) consonant. (See Paul 1966a and K&P for a precise description).

(14)   M+i+petraka eo ambonin'ny latabatra izy
        pres+act+place there on+top'the table she
        She is sitting on the table

   M+aN+petraka (mametraka) boky eo ambonin'ny latabatra izy
   pres+act+place book there on+top'the table she
   She is putting the book on the table

Man- verbs tend to be transitive, *mi-* ones intransitive, though there are exceptions both ways. All ditransitives (except causatives) are *man-* verbs, and if a root takes both *mi-* and *man-*, as in (14), the *man-* verb almost always has more arguments than the *mi-* verb.

1.2.3 *m-* and *ma-* prefixes, in distinction to *mi-* and *man-* apply to closed classes of roots, several of quite common occurrence. In the *m-* case, which we analyze as having a ∅ prefix, we have forms like: *isy* ⇒ *misy* 'exists', *aka* ⇒ *maka* 'takes', and *ino* ⇒ *mino* 'believes'. In the *ma-* case, whose verbs tend to be stative and translate as adjectives in English, we have *rary* ⇒ *marary* 'is sick', *loto* ⇒ *maloto* 'is dirty', *toky* ⇒ *matoky* 'trusts', and *hita* ⇒ *mahita* 'sees'. Usually roots taking *m-* and *ma-* prefixes do not form distinctive passives and the active verbs themselves are usually intransitive.

1.2.4 *maha-*, *mana-* and *mank(a)-* are distinguished from the other primary active prefixes in that they have causative meaning of their own. *maha-* in addition has an abilitative meaning (Phillips 2000) and applies to a surprising variety of expressions in addition to roots (K&P). Some typical instances are *finaritra* 'pleasing' ⇒ *mahafinaritra* 'causes pleasure (transitive), is pleasurable (intransitive)'; *zaka* 'handled, accustomed to' ⇒ *mahazaka* 'can handle'; *tsara* 'good' ⇒ *manatsara* 'makes good'; *rary* ⇒ *mankarary* 'makes sick', *sitrana* 'cured (root passive)' ⇒ *mankasitrana* 'cures'.

1.2.5 *miscellaneous*. There are a few apparent prefixes that occur with just a very few roots that do not (quite) fall into the categories above. For example, from *atsimo* 'South' we have *mianatsimo* 'goes South'. And from locative deictics like *any* 'there, not visible to speaker' we form *mankany* 'goes there', where the apparent prefix *mank(a)-* does not have the causative effect mentioned in the section above.

1.2.6 *Secondary prefixes* apply to already prefixed roots. The main ones are *mamp-*, the major causative affix (See Andrianierenanana 1996 and Randriamasimanana 1986), and *mif-* the reciprocal affix (See Keenan and Razafimamonjy 2000).

(15)   a. m+amp+an+soratra (mampanoratra) taratasy azy aho
            pres+cause+act+write letter him I
            I make him write letters

        b. m+if+an+soratra (mifanoratra) taratasy isika
            pres+rec+act+write letters we (incl)
            We write letters to each other

Verbs with *amp-* form suffix passives and circumstantials (below); ones with *if-* form circumstantials but not passives.

1.2.7 *Tertiary prefixes* combine with tensed active verbs to form further active verbs. There are just two such prefixes: *miha-* 'become' and *miaraka* 'do together', the latter an independent verb meaning to be or go together.

(16)   a.   m+iha+m+an+hatsiaka (mihamangatsiaka) izy
                pres+inch+pres+act+cold he
                He is becoming cold

        b.   m+iaraka+m+i+asa (miara-miasa) izy ireo
                pres+com+pres+act+work they dem+pl
                They work together

1.3 Circumstantial verbs are formed by suffixing *-ana* to active verbs less the distinctive *m-*prefix. The EA typically has an oblique role and in distinction to EAs for active and passive verbs is not subcategorized by the verb. As with passives the Agent phrase is tightly bound as a possessor to the verb. Circumstantial forms are widely used when relativizing, questioning or focusing (clefting) obliques. They have imperative forms, like actives and passives, (17b).

(17)   a. ny zaza (izay) n+i+vidy+iana+Rasoa (nividianan-dRasoa) akanjo
            the child (rel) past+act+buy+circ+Rasoa clothes
            the child for whom Rasoa bought clothes

        b. Mba i+vidy+ana+o (ividiano) akanjo izy
            please act+buy+circ+imp clothes he (nom)
            Please buy him some clothes

        c. ny zavatra if+amp+i+anatra+ana+(n)tsika (ifampianarantsika)
            the things rec+cause+act+study+circ+us.gen(incl)
            the things we teach each other

d. Taiza no n+i+tsangantsangan+ana+nareo (nitsangantsanganareo)?
   Past+where foc past+act+stand+redup+circ+you(gen pl)
   Where did you all take a stroll?

(Verbs lacking passive forms, such as reciprocals, (17c), or zero prefix ones, use the circumstantial when Patients or Themes are extracted).

1.4 Nominalizations are very productive in Malagasy and preserve the voice morphology and subcategorization, though the nominalizing affixes only apply to active and circumstantial verbs, not passive ones.

1.4.1 *Agent nominals* are formed by replacing the active prefix *m-* with *mp-* (pronounced /p/), and thus exist only for active, not passive or circumstantial forms: *teny* ⇒ *miteny* 'speaks' ⇒ *mpiteny* 'speaker'; *leha* ⇒ *mandeha* 'goes' ⇒ *mpandeha* 'voyagers'; *soratra* ⇒ *mifanoratra* 'write to each other' ⇒ *mpifanoratra* 'people who write to each other'; *anatra* ⇒ *mampianatra* 'teaches (lit: causes to study) ⇒ *mpampianatra* 'teacher'.

1.4.2 *f-* nominals are formed by prefixing circumstantial or active verbs (less the *m-* prefix) with *f-*. In the active case *f-* nominals may have an 'instrument of action' interpretation, as in *zaitra* ⇒ *manjaitra* 'sews' ⇒ *fanzaitra* 'needle' or, more productively, a manner of action sense, as in *teny* ⇒ *miteny* 'speaks' ⇒ *fiteny* 'manner of speaking.

The use of *f-* nominalizations with circumstantial verbs is very productive and forms a wide range of abstract nominals:

(18) a.   -an+léha+ana   ⇒   andehánana   ⇒   fandehánana
           act+go+circ   circumstance of going   departure

      b.   -amp-i-ánatra+ana ⇒   ampianárana   ⇒   fampianárana
           cause+act+study+circ   circumstance of teaching   instruction

      c.   -anka+tó+ana   ⇒   ankatoávana   ⇒   fankatoávana
           cause+ accepted+circ   circumstance of accepting ratification

      d.   -aha-márina+ana   ⇒ ahamarínana   ⇒   fahamarínana
           cause-level/true+circ   circumstance of verifying justice

2. The text study

Our study consisted of two novels and some 53 short newspaper articles. The novels[2], NZ and IKM, are romantic/adventure in genre. They are well written, popular, with much dialogue. We did the first few chapters in each, yielding a sample of 23,241 words for NZ, and 20,985 for IKM, for a total sample of 44,226 words. A "word" here is just a segment of text between spaces, so *nifampiananareo* 'were caused to make each other learn by you.pl' is one word (but 8 morphemes). The newspaper articles all appeared in 1995 in the capital Antananarivo. They are reports of current issues: car accidents, break-ins, and current social and political events. We report separately on the two classes of texts, as the results differ somewhat.

2.1 Active vs non-active voice forms

Here first is a summary overview of the frequency of active vs non-active verbs in the two novels:

Table 1: Malagasy novels

verb forms	tokens	percent	types	percent
active	5601	62%	946	50
non-active	3364	38%	945	50
total	8965	100%	1891	100

Thus we have a total of 8,965 occurrences of verbs, of which 5,601 or 62% are active and 3,364 or 38% are non-active. The novels differ little among themselves: In IKM 64% of the verb occurrences are active, in NZ 61%. Contrast these figures with those for English from Svartvik (1966:46), also based on two novels[3].

Table 2a: English (< Svartik 1966)

verb forms	tokens	percent

---

[2]*Ny Zanako* 'My Child' by C. Ratsifandrihamanana, Imprimerie Nationale, 1969 and *Ilay Kintana Mamirapiratra* 'That Shining Star' by Philippe Rajohanesa, Imprimerie Luthérienne, 1963.

[3]*The Echoing Grove* (1958) by R. Lehmann, Penguin Books and and *Eating People is Wrong* (1959) by M. Bradbury, London, sections of 5,000 words from each.

active	1216	94%
passive	79	6%
total	1295	100%

Thus only 6% of the verb occurrences in the English novels Svartvik analyzed were passive. A smaller study by Givon (1979), based on two (American) English novels[4] and limited to main declarative affirmative clauses, yields a comparable figure, 9%:

Table 2b: English (< Givon 1979)

verb forms	tokens	percent
active	177	91%
passive	19	9%
total	196	100%

In contrast in Malagasy novels 38% of the verbs are non-active. So these figures support that in Malagasy literature the non-active voices occur about 5 to 6 six times as frequently as they do in English.

A second measure of frequency is occurrences of non-active verbs per 1,000 words of text. Svartvik (1966:155) computes 8.2 passives per 1,000 words for the two English novels analyzed. Stein (1979:180-3) also analyzed two English novels[5], comprising 140,000 words of text, and found 1284 passive verbs. This computes to 9.1 passives per 1000 words, a figure that substantially agrees with Svartvik's.

Stein also analyzed two German novels[6] and found 1281 passives in 170,000 words of text

---

[4]*Under the Sweetwater Rim* Bantam, by L. L'Amour 1965, and *Animal Farm* by G.Orwell 1945 Penguin Classics.

[5]*A Fairly Honourable Defeat* by I. Murdoch, Harmondworth 1975, 102,000 words in length. And *Pasmore* by D. Storey. Harmondsworth 1976, 38,000 words.

[6]*Das Vorbild* by S. Lenz, 1973. Hamburg, 130,000 words. And *Die verlorene Ehre der Katharina Blum oder: Wie Gewalt entstehen oder wohin sie führen kann* by H. Böll. Köln, 1974. 40,000 words.

(Stein 1979:198-201), which computes to 7.5 passives per 1000 words. So German is comparable to English, with passives occuring slightly less frequently.

But our figures for the two Malagasy novels yield 3364 non-active verb occurences in 44,226 words of text, which computes to 76.0 passives per 1,000 words. Clearly there is an order of magnitude difference between the Malagasy and the English/German cases: non-actives in Malagasy occur about 10 times more frequently than non-actives in English and German. Summarizing:

Table 3
Non-active verbs per 1,000 words of text

Language	non-actives/1,000 words of text
English (Svartvik)	8.2
English (Stein)	9.1
German (Stein)	7.5
Malagasy	76.0

Notwithstanding this order of magnitude difference we are surprised at the high frequency of active verbs in the Malagasy novels. (The figure is significantly lower in the newspaper articles). Comparing numbers of distinct verbs, as opposed to numbers of occurrences (tokens), we see in Table 1 that the number of active and non-active verbs is basically the same. Thus some active verbs occurred with much higher frequency than passive verbs. We note two cases:

First, three root actives *lasa* 'gone', *avy* 'comes/came', and *tonga* 'arrives, arrived' occurred 363 times, over 6% of the active verbs. As noted earlier, these verbs lack distinctive active morphology, have no passives, and function frequently as auxiliaries. But we count all occurrences as active in the interests of obtaining reproducible results.

Second, a very common active verb is the existential *misy* 'exists', a zero prefix verb with no passive (but with a circumstantial form, *isiana*). It occurred, in various tenses, 281 times – 5%

of the active verbs. (19a) is a typical usage, but it also occurs in many fixed expressions in which its verbal force is weak to non-existent, (19b,c).

(19)   a.   Misy zazavavy betsaka ao an-dakilasy
          exist girls many there in-class
          There are many girls in the class

       b.   cafe misy ronono   c.   tsy misy fisaorana
           coffee has milk   not exist nom+thank+circ
           coffee with milk   *You're welcome*

So the degenerate active roots plus the existential verb account for over 11% of the active verbs in Malagasy texts.

     Table 4 shows that in the newspaper articles, non-active verbs and active ones occur with the same frequency, a statistic more in conformity with our expectations than the distribution in the novels.

Table 4
Malagasy newspaper articles

verb forms	tokens	percent	types	percent
active	1620	50%	805	51%
non-active	1614	50%	772	49%
total	3234	100%	1577	100%

     We have no explanation for why the proportion of non-actives is somewhat greater in newspaper articles than in novels, but the increase does further reinforce our observation that non-actives are an order of magnitude more frequent in Malagasy than in either English or German.

2.2 Non-active voice forms Table 5 presents the percentage of non-actives with Agent phrases, broken down into passive and circumstantial groups.

Table 5: Malagasy novels

Voice	tokens	+Agent	% +Agent
passive	1846	1200	65%
circumstantial	1532	814	53%
total	3378	2014	60%

Thus of the 1,832 occurrences of passive verbs, 1200 or 65% presented Agent phrases. By contrast just 53% of the circumstantial verbs presented Agent phrases. Combining the two we see that 60% of non-actives present Agent phrases in Malagasy novels.

These figures contrast again very markedly with those cited for English and Dutch, given in Table 6. Stein's English data are those from the two English novels she studied; those from Dusková and from Svartvik (which go beyond Svartvik's 1966 study), are cited from Stein (1979:126). The figures from Givon (1979:59-60) include ones from an additional novel[7]. Kirsner's counts from Dutch also come from novels. And the Malagasy figures are just from the two novels, to maximize comparability with Stein's, Givon's and Kirsner's figures.

Table 6

Source	#non-active verb tokens	# +Agent	% +Agent
English (Svartvik)	2696	539	20%
English (Stein)	1293	220	17%
English (Dusková)	598	90	15%
English (Givon)	182	30	16%

---

[7]*Venus on the Half Shell* by K. Trout 1974. Dell.

Dutch (Kirsner)	274	39	14%
Malagasy	3378	2014	60%

The figures given for the European languages by the different authors agree remarkably well: from 14% to 20% of non-actives have Agent phrases. So Agent phrases occur in just one in five or six occurrences of non-actives in English and Dutch. But in Malagasy by contrast the figure is 60%. So by this measure non-actives are three to four times as frequent in Malagasy as in English and Dutch.

Other studies support the low percentage of agented forms in English passives. Weiner and Labov (1983:34) report that agented passives make up only 1 to 2 per cent of the total number of passivizable transitive verbs in the spontaneous conversations they studied. They note that Horgan (1978) was obliged to elicit agented passives for lack of sufficient instances in spontaneous speech; Harwood (1959) found no agented passives in over 12,000 utterances of 5-year olds, and Brown (1972) found no instances in 2,100 utterances of parents studied.

Moreover our figure for agented non-actives in Malagasy is in fact artificially low. In several syntactic contexts missing Agent phrases are controlled, and are thus not semantically missing, just as we do not count *read* as subjectless in control contexts like *I intend to read this book*, as *I* binds it. As a first instance, Malagasy active verbs of aspect ("begin",...), intent, and desire may take active verbal complements with controlled empty subjects, (20a); non-active verbs take non-active verbal complements with controlled empty Agent phrases, (20b).

(20) a.   N+i+kasa h+an+vaky (hamaky) io boky io aho
            past+act+intend fut+act+cut that book that I
            I intended to read that book

      b.   No+kasa+ina+ko (nokasaiko) ho+vaky+ina (hovakina) io boky io
            past+intend+pass+1sg.gen fut+cut+pass that book that
            I intended to read that book   lit: *That book was intended by me to be read (by me)*

The Agent in (20b) can be a quantified NP, like *nokasain'ny mpianatra rehetra* 'intended by the students all' preserving the binding relationship.

The natural translation of (20b) into English is active, the literal translation being cumbersome. But (20b) in Malagasy is not at all cumbersome. Asked for a translation of *I intended to read that book* we are at least as likely to get (20b) as (20a). And as we have seen in Malagasy only EAs extract, so if we want to relativize (question, cleft) the book in (20) we must

build on the double passive, (20b), as in (20c,d), (20c',d') being hopelessly ungrammatical. See Law (1995) and Keenan (1995) for more extensive discussion.

(20)   c.   ny boky (izay) nokasaiko hovakina
            the book (that) was intended by me to be read (by me)
            the book I intended to read

       c'. *ny boky (izay) nikasa hamaky aho
            the book (that) intended (act) fut+read (act) I
            the book that I intended to read

       d.   io boky io no nokasaiko hovakina
            that book that Foc intended by me to be read (by me)
            It was that book which I intended to read

       d'. *io boky io no nikasa hamaky aho
            that book that Foc intended(act) fut+read(act) I
            It was that book which I intended to read

Observe now that the non-active verbal complement, *hovakina* 'will be read (by me)' is presented without an overt Agent phrase and was counted as agentless by our program, even though its Agent phrase is controlled. And these constructions are utterly ordinary, despite the cumbersomeness of their English translations. Thus (21b) below is a normal way for a waiter to ask for your order:

(21)   a.   Te-hisotro dite aho
            want-fut+act+drink tea I
            I would like to drink some tea

       b.   Inona no tianao hosotroina?
            what Foc want+pass+2sg.gen fut+drink+pass
            What would you like to drink?

The examples above are paradigm cases. Here is an example from IKM, one which shows that the dependent verb may be circumstantial, not merely passive. Note that it is coordinate, whence this example would (misleadingly) contribute two agentless non-actives to our count.

(22)   ...toerana izay azontsika anaovana afo sy ipetrahana
       ...place where can+by+us(incl) make+circ fire and rest+circ
       ...a place where we can make a fire and rest

Thus we know that our count of agentless non-active verbs is exaggerated, though we do not

know by how much. Here are some further cases in which our program wrongly counts non-active verbs as agentless when in fact their Agent phrase is controlled.

Paul and Ranaivoson (1998) discuss a dozen odd verbs like *maniraka* 'send' in (23a) which, when active, form complex predicates with active complements, controlling their subjects, but when non-active take non-active complements, controlling their missing agent:

(23) a. man+iraka mi+vidy mofo an'ilay vehivavy ianareo
           act+send act+buy acc'that woman you.pl.nom
           You send that woman to buy bread

      b. Inona no anirahanareo an'ilay vehivavy vidina?
           what Foc send+circ+by-you.pl acc'that woman buy+pass
           What do you send that woman to buy?
           lit: *What is sent by you.pl that woman to be bought?*

A third, less common, case of controlled empty Agent phrases are initial conjuncts of coordinate non-actives whose empty Agent phrases are controlled by the overt one in the second conjunct.

(24) a. ...voa+jery sy voa+dinika+nao (voadinikao) tsara ireo sary teo
           ...pass+see and pass+study+2sg.gen good those pictures there
           ...those pictures were seen and studied by you well

      b. ...izao ahitana sy ahafantaranareo ahy izao
           ...that act+seen+circ and cause+known+circ+2pl.gen me that
           ...that which is seen and understood by you about me

(Both these examples are from *Volavola Ranomaso*, a novel by O. Ranaivo, 1957). A final example, partially covered above, is:

(25) a. Tsy azo a+tao izany
           not permit(rt:pass) pass+do that
           That isn't done

      b. Tsy azo anaovana hira f+andihizana ao am-piangonona
           not permit(rt:pass) do+circ song nom+dance+circ there in-church
           One shouldn't make dancing music in church

Thus the very common passive root *azo* takes passive or circumstantial complements and its Agent controls theirs. Note that neither *azo* nor its following complements have overt Agent phrases. One could be added, replacing *azo* with *azoko* 'permitted by me' for example, and

changing the meaning of (25a) to "I couldn't do that" and that of (25b) to "I couldn't make dancing music in church". But without the Agent phrase we understand the Agent to be "One/Anyone", as indicated in the translations. So our counting program finds two agentless non-actives apiece in (25a,b) though in fact none of the four are agentless.

In fact (25b) would contribute two further agentless non-actives according to our counting method. *Fandihizana* is an *f-* nominalization of the circumstantial verb *andihizana*, built from the root *dihy* 'dance'. A*m-piangonana* 'at church' is *aN+fiangonana*, where *aN-* is a locative and *fiangonana* is the circumstantial nominalization built from the root *angona* 'gather'. We have already noted that many abstract nominals are built from circumstantial forms, which we continue to count as circumstantial as they preserve the subcategorization of the verbs:

(26) a. ny antony (izay) anajahan-dRabe azy
          the reason (why) respect+circ+Rabe(gen) him(acc)
          the reason Rabe respects him

      b. ny f+anajahan-dRabe azy
          the nom+respect+circ+Rabe(gen) him(acc)
          Rabe's respect for him

Many nouns in common usage are built from circumstantial verb forms in this way, and are often used without an Agent: *varotra* ⇒ *fivarotana* 'shop'; *afaka* ⇒ *fahafahana* 'freedom'; *voy* ⇒ *fifamoivoizana* 'traffic'; *tatitra* ⇒ *fitaterana* 'transportation'. These contribute significantly to the lower use of Agent phrases in the newspaper articles:

Table 7: Malagasy newspapers

non-active	tokens	per cent	+Agent	%+Agent
passive	903	56%	373	41%
circum.	711	44%	285	40%
total	1614	100%	658	41%

We see that the percentage of agented non-actives here is only 41%, compared with 60% in the novels. Even so the figure is still 2 to 3 times higher than for English/German/Dutch, and, as noted, the difference between the two language groups is even greater, as (1) many overtly agentless non-actives are directly controlled in Malagasy in a way in which they are not in English/German/Dutch, and (2) many words counted as circumstantial are lexicalized nouns

which we expect not to have a possessor (traffic, transportation) or which will not be possessed any more than ordinary common nouns (church, shop).

2.3 Malagasy internal remarks

Lastly we consider briefly the distribution of voice forms internal to Malagasy. We note first that of non-active verbs the passives outnumber the circumstantials in terms of number of occurrences: 2,749 to 2,243, though the difference is not very large. More interesting is the distribution of the different types of passive affixes:

Table 8: Malagasy Novels and Newspapers

passives	tokens	+Agent	% +Agent
suffix	1132	744	66%
root	832	469	56%
prefix	785	360	46%

Unsurprisingly suffix passives outnumber root and prefix passives. More surprising however is the fact that root passives slightly outnumber prefix ones (which include, recall, the *a-*, *voa-* and *tafa-* passives). Equally most suffix and root passives have overt Agent phrases, only among the prefix passives do agentless ones outnumber the agented ones. Again the difference is not massive.

The prominence of root passives is interesting on a more general level: generative treatments of English passives have tended to treat the form of passive verbs as some function (the *-en* function) of active verbs. That is, passive verbs are formed by taking the *-en* (past participle) form of the active verb and using an appropriate auxiliary. But most usually in Malagasy actives and passives are derived directly from a given root by different affixes. E.g. from *vidy* we form active *mividy* 'buys' and passive *vidina* 'is bought'. So typically active and passive verbs in Malagasy are of comparable derivational complexity. But in the common, case of root passives, actives are of greater morphological complexity, being derived directly from the

root passive. Here are a few examples using different active prefixes:

(27)   root passive   active

        hita 'seen'   mahita 'sees'
        re 'heard'   mandre 'hears'
        tapaka 'cut'   manapaka 'cuts'
        azo 'permitted, got'   mahazo 'permits'
        fantatra 'known'   mahafantatra 'knows'

Finally, since we have the data, let us summarize the relative frequencies of the common active forming prefixes. We just consider the primary affixes, and ignore those with very low frequencies.

Table 9: Malagasy Novel and Newspapers

prefix	IKM	NZ	News.	total
mi-	826	1068	543	2437
man-	666	816	348	1830
m+∅-	309	291	147	747
maha-	201	162	82	445
root	276	135	8	419
ma-	166	124	91	381

Clearly the *mi-* and *man-* prefixes are by far the most common, as experience predicts. That *mi-* verbs outnumber *man-* ones is perhaps not surprising: they dominate the intransitive spectrum and also form many transitive verbs (*mividy* 'buys', *mifidy* 'chooses', *mikapoka* 'beats'). *man-* verbs are overwhelmingly (di)transitive (though we have perhaps a dozen intransitives with *man-* (*mandihy* 'dances', *mandeha* 'goes', *mandohalika* 'kneels', *mangetaheta 'is thirsty'*,...).

It is perhaps more surprising that the zero prefix verbs constitute the next most frequent subcategory of active verb. One might have guessed that causative *maha-* verbs and the root actives *lasa* 'gone', *avy* 'comes/came' and *tonga* 'arrives/arrived' were more frequent. Finally, as

noted earlier, the *ma-* prefix verbs are a smallish closed class, a few of whose members are quite common, such as *matory* 'sleeps', *mahita* 'sees' and *marary* 'is sick'.

We have not included active verbs with secondary prefixes (causative *amp-* and reciprocal *if-*), which apply to verbs in the first three frequency categories above. There were in fact 169 *amp-* verbs and 112 *if-* ones, not nearly enough to change the relative order among the first three types even if (as is not the case) *amp-* for example only combined with *man-* verbs. For the same reason we did not include in Table 9 verbs built with the third leval prefixes *miha-* 'become' and *miara-* 'do together'.

3. Conclusions and speculations

The two main conclusions of our quantitative study are:

1. Non-active verbs vie with active ones in terms of frequency of occurrence, and
2. Agent phrases of non-actives are typically present or controlled.

These facts suggest that passives of transitive verbs should be treated in Malagasy as transitve verbs rather than as derived intransitives with Agent phrases expressible as optional adjuncts. Pearson (2000) pursues this approach from a syntactic perspective in an enlightening way. Keenan (1995) found this approach tractable semantically. That is, as long as the Agent phrase in non-actives is treated as an argument and thus accessible from the arity of the verb the semantic interpretation of non-active verbs and their logical synonymy with active ones is compositionally expressible. We might note as well that this perspective is one that Philippinists have always found natural, eschewing the active/passive terminology in favor of a "Focus" system. And in essence Malagasy has a Philippine type clause structure, with less "case marking" morphology on NPs and a more rigid word order and constituent structure.

NOTES

[†] The first author wishes to acknowledge his gratitude to the U.S. Fulbright Commission for supporting this research. Equally thanks are due to Prof. R-B Rabenilaina, head of the DIFP at the University of Antananarivo, whose hospitality greatly facilitated this work.

REFERENCES

ABINAL AND MALZAC (RR.PP). 1888. Dictionnaire malgache-français. Paris: Editions maritimes
   et d'outre-mer.

ANDRIANIERENANANA, C.L. 1996. Morphological causatives in Malagasy. in Pearson & Paul 1966. pp. 58-76

BROWN, R. 1973. *A First Language*. Harvard University Press.

DUSKOVÁ, L. 1972. The Passive voice in Czech and in English. in *Philologica Pragensia* 15:93-118.

ERWIN, S. 1996. Quantity and moras: An amicable separation. in   Pearson & Paul (1996:2 – 31).

GIVON, T. 1979. *On Understanding Grammar*. Academic Press

HOPPER, P. 1983. Ergative, Passive, and Active in Malay Narrative. in *Discourse Perspectives on* Syntax. Flora Klein-Andreu (editor). Academic Press. pp. 67–88.

HARWOOD, F. 1959. Quantitative study of the speech of Australian children. *Language and Society* **2**:236–270.

HORGAN, D. 1978. The development of the full passive. *Journal of Child Language* **5**:65–80.

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

KEENAN, E. L. 1976. Remarkable Subjects in Malagasy. in *Subject and Topic*, C. Li (ed). New York: Academic Press

KEENAN, E.L. 1995. Predicate-Argument Structure in Malagasy. in *Grammatical Relations: Theoretical Approaches to Empirical Questions*. C. Burgess, K. Dziwirek, and D. Gerdts (eds) CSLI. Stanford, Ca. pp. 171–217.

KEENAN, E.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.

KEENAN, E.L. & M. POLINSKY. 1998. Malagasy. in *The Handbook of Morphology*. A. Spencer and A. Zwicky (eds) Blackwell Pub. pp. 563–625.

KEENAN, E.L. & J-P. RAZAFIMAMONJY. 2000. Reciprocals in Malagasy. to appear in *UCLA Occasional Papers in Linguistics*: *Papers in African Linguistics* H. Torrence and P. Munro (eds). Dept. of Linguistics, UCLA

KIRSNER, R. 1983. On the Use of Quantitative Discourse Data to Determine Inferential

Mechanisms in Grammar. in *Discourse Perspectives on Syntax*. Flora Klein-Andreu (editor). Academic Press. pp. 237–257.

LAW, P. 1993. On Grammatical Relations in Malagasy Control Structures. in *Grammatical Relations: Theoretical Approaches to Empirical Questions*. C. Burgess, K. Dziwirek, and D. Gerdts (eds) CSLI. Stanford, Ca.

PAUL, I. 1966. The Malagasy genitive. In Pearson & Paul (eds). pp 76– 92.

PAUL, I. 1966a. The active marker and nasals in Malagasy. In Pearson and Paul (eds) pp. 49–57.

PAUL, I. (ed) 1998. *The Structure of Malagasy* Vol II. UCLA Occasional Papers in Linguistics, 20. Department of Linguistics, UCLA

Paul, I. & J.F. Ranaivoson. 1998. Complex verbal constructions in Malagasy. in Paul (1998), pp. 111 – 125.

PEARSON, M. 1996.Domain phrases and topic arguments in Malagasy existentials. In Pearson & Paul, 1996. pp., 113 – 142.

PEARSON, M. & I. PAUL. (eds) 1996. *The Structure of Malagasy*, vol I. UCLA Occasional Papers in Linguistics, 17. Department of Linguistics, UCLA.

PEARSON, M. 2000. *The Derivation of Malagasy Clause Structure* [working title]. PhD dissertation. UCLA.

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

RABENILAINA, R-B. 1993. *Le Verbe Malgache: Constructions transitives et intransitives*. Centre d'Etudes et de Recherches sur le Malgache, Université d'Antananarivo. (Revised version of *Lexique-grammaire du malgache. Constructions transitives et intransitives*. 1985. Thèse de doctorat d'Etat, Université de Paris 7: D.R.L. et L.A.D.L.

RABENILAINA, R-B. 1998. Voice and Diathesis in Malagasy: An Overview. in Paul 1998.

Rajemisa-Raolison, R. 1971. *Grammaire malgache*. Fianarantsoa: Ambozontany (first edition 1959).

RANDRIAMASIMANANA, C. 1986. *The Causatives of Malagasy*. University of Hawaii Press.

STEIN, G. 1979. *Studies in the Function of the Passive* Gunter Narr Verlag, Tuebingen

SVARTVIK, J. 1966. *On Voice in the English Verb* Mouton & Co.

WEINER, E.J & W. LABOV. 1981. Constraints on the agentless passive. in *J. Linguistics* 19: 29 – 58.

WOUK, F. 1996. Voice in Indonesian discourse and its implications for theories of the development of ergativity. in *Studies in Language* 20:2. pp. 361–410.