# WORD ORDER AND TRANSITIVE CLAUSES IN MALAGASY

## A. Rasolofo
*University of Oregon*
*<arasolof@gladstone.uoregon.edu>*

## 1 Introduction

Malagasy has been considered to have a Verb-Object-Subject (VOS) word order (Keenan 1976, Keenan and Ochs 1979, Randriamasimanana 2000).[1] This paper argues that a distinction between Verb-Agent-Patient (VAP) and Verb-Patient-Agent (VPA) clauses is a better characterization of the language. The scope of this study focuses on the motive behind the choice of these two types of transitive clauses.

This paper starts by addressing issues with considering Malagasy as having a VOS transitive construction. Then, in order to understand the proposed solution, it is necessary to both have a minimum grasp about transitive construction types in the language and present the methodology used for the analysis. The final parts of the paper discuss the results of topicality measurement done on a series of narrative texts.[2]

## 2 The assumption of a single transitive clause form

Considering the language as having a single transitive clause form leads to the assumption that VOS (1a) is the basic word order. From this standpoint, constructions such as (1b) are taken to be passive clauses (Rabenilaina 1991, Keenan and Ochs 1979, among others):

(1)　a. n-an-shuratra　ilai taratashi ilai umbiashi.
　　　　PST-EA:A-write DEF letter　　DEF soothsayer
　　　　'The soothsayer wrote the letter.'

　　　b. n-u-shuratra-ana ilai umbiashi　　ilai taratashi.
　　　　PST-u-write-EA:P DEF soothsayer DEF letter
　　　　'The letter was written by the soothsayer.'
　　　　'The soothsayer wrote the letter.'

Claiming that (1b) is a passive clause does not take into account several facts. First, constructions like (1a) and (1b) are different in terms of word order. For (1a) the arguments' order is patient-agent, hence the word order Verb-Patient-Agent (VPA) while for (1b) the order is agent-patient, hence the word order Verb-Agent-Patient (VAP).

Furthermore, the difference of word order is accompanied by verbal morphology alteration. In (1a), the verb is structured as tense-external argument-root while in (1b) the

---

[1]　Malagasy is an Austronesian language spoken in Madagascar.

[2]　All clauses used in the text counting are from a series of folk tales taken from *Hafatry ny ombiasy iray ho an'ny zanany* by Rainandriamampandry and Ranaivosoa (1900s). There were about 250 clauses in whole, of which 219 were considered for this study.

structure is tense-root-external argument. Permuting the word order without changing the verb morphology results in ungrammatical clauses:

(2)   a. *n-an-shuratra   ilai umbiashi  ilai taratashi.
          PST-EA:A-write DEF soothsayer DEF letter
          'The letter wrote the soothsayer.'

      b. *n-u-shuratra-ana ilai  taratashi  ilai  umbiashi.
          PST-u-write-EA:P DEF letter       DEF soothsayer
          'The letter wrote the soothsayer.'

Third, missionaries and linguists (Cousins 1897, Keenan 1976, Randriamasimanana 2000 among others), who analyzed the language, brought up the fact that, in actual spoken instances, speakers have a preference for constructions such as (1b). If this type of clause was really a passive construction, it might not be expected to have such a high frequency of occurrence.

Finally, depending on the word order, the arguments' constituency is variable. A temporal adverb test is used in this analysis to differentiate constituency. The adverbial phrase *umali* 'yesterday' is used between each pair of words. For both VAP and VPA constructions, if the result produces a grammatical construction, then the words on the left and on the right of the temporal adverb do belong to two separate units. If the result is ungrammatical, then the words on the right of the temporal adverb cannot be separated from the words on the left. Hence, they belong to the same unit. With VPA construction, separating the noun phrases *ilai taratashi* 'the letter' and *ilai umbiashi* 'the soothsayer' with the temporal adverb results in a grammatical clause (3a) while separating the verb *n-an-shuratra* 'wrote' from the noun phrase *ilai taratashi* is not (3b). Therefore, *ilai taratashi* and *ilai umbiashi* belong to different units while *ilai taratashi* forms a unit with the verb phrase:

(3)     a. n-an-shuratra   ilai taratashi umali     ilai umbiashi.
            PST-EA:A-write DEF letter      yesterday DEF soothsayer
            'The soothsayer wrote the letter yesterday.'

        b. *n-an-shuratra    umali     ilai taratashi ilai umbiashi.
            PST-EA:A-write yesterday DEF letter      DEF soothsayer
            'The soothsayer wrote the letter yesterday.'

With VAP construction, inserting the temporal adverb *umali* 'yesterday' between the two noun phrases *ilai umbiashi* 'the soothsayer' and *ilai taratashi* 'the letter' yields a grammatical clause (4a) while putting the adverb between the verb *n-u-shuratra-ana* 'wrote' and *ilai umbiashi* results in an ungrammatical clause (4b). Consequently, *ilai umbiashi* and *ilai taratashi* belong to different units but *ilai umbiashi* forms a unit with the verb phrase:

(4)  a. n-u-shuratra-ana ilai umbiashi    umali      ilai taratashi.
        PST-u-write-EA:P DEF soothsayer yesterday DEF letter
        'The soothsayer wrote the letter yesterday.'

     b. *n-u-shuratra-ana  umali      ilai umbiashi    ilai taratashi.
         PST-u-write-EA:P  yesterday DEF soothsayer DEF letter
         'The soothsayer wrote the letter yesterday.'

Phrases forming a unit with the verb phrase (e.g. *ilai taratashi* in (3b) and *ilai umbiashi* in (4b)) are referred to here as *internal arguments* (IA) while those that can be separated from the verb phrase are referred to as *external arguments* (EA). Readers should note that the terms internal and external argument are therefore used here in a way which does *not* correspond directly to the use of the same terms in transformational grammar/Chomskyean linguistics (where the internal argument of a verb is always its object, and the external argument its subject).

## 3 Transitive Clause Types

Because of the constituency distinction and because of the difference in terms of argument definiteness, Malagasy transitive clauses represent, at least, three different categories (i.e. VAP, VP$_i$A, and VP$_d$A constructions).

External arguments of transitive clauses must be definite. The use of indefinite external arguments results in ungrammatical sentences. Compare examples in (5) with those in (1):

(5)  a. *n-an-shuratra    ilai taratashi ø      umbiashi.
        PST-EA:A-write DEF letter      IDEF soothsayer
        'A soothsayer wrote the letter.'

     b. *n-u-shuratra-ana ilai umbiashi   ø       taratashi.
         PST-u-write-EA:P DEF soothsayer IDEF letter
         'The soothsayer wrote a letter.'

For VPA constructions, the internal argument can be either definite (as in 1a) or indefinite (as in 6):

(6)  n-an-shuratra    ø      taratashi ilai umbiashi.
     PST-EA:A-write IDEF letter      DEF soothsayer
     'The soothsayer wrote a letter.'

Because of the difference of internal argument definiteness, VPA constructions are distinguished as VP$_i$A, for indefinite internal argument, and VP$_d$A, for definite internal argument.

For VAP constructions, this paper will not consider sentences where the internal argument is indefinite, as illustrated in (7):

(7)   ?n-u-shuratra-ana ø      umbiashi    ilai taratashi.
        PST-u-write-EA:P IDEF soothsayer DEF letter
        'A soothsayer wrote the letter.'

Although clauses such as in (7) may be grammatical, I will only consider VAP constructions in which both internal and external arguments are definite, as in (1b).

## 4  Methodology

Coorman (1982)'s and Givón (1983)'s topicality measurement is applied throughout this analysis to establish the difference between Malagasy transitive constructions. *Topicality* is a property of a noun (or noun phrase) when it is recurring through discourse (Givón 1990). According to Thompson (1994), a topical argument is one that carries some weight of importance and predictability within a text. An argument is important when, after its current mention, it appears often within the following discourse, and it is predictable when it is mentioned at least once in a previous nearby clause.

Cooreman and Givón define *anaphoric reference* (Referential Distance in Coormen and Givón's terminology) as the number of clauses between the present occurrence of the referent and its last occurrence in the preceding clauses. When a referent is located in the clause immediately preceding the present occurrence, the value of 1 is assigned. If there is an occurrence in two or three clauses before the present occurrence, the value of 2 or 3 is respectively assigned. If the present occurrence is the first mention of the referent or there are no occurrence of the referent in the three preceding clauses, the value of >3 is assigned. Referents with values between 1 and 3 are said to be topical while those getting values >3 are said to be less or non-topical.

*Cataphoric reference* (Topic Persistence in Cooreman and Givón's terminology) is defined as the number of times a referent is mentioned in the next 10 clauses following its present occurrence. If the referent is not mentioned at all within the next 10 clauses the value of 0 is assigned. If it is mentioned 1, 2, 3, …, or 10 times the value of 1, 2, 3, …, or 10 is assigned. Referents with values greater than two (i.e. >2) are said to be topical while those getting values between 0 and 2 are said to be less or non-topical.

In the series of narrative text used for this study, for each type of transitive clause, both anaphoric and cataphoric references are counted in order to determine the topicality of the arguments.

## 5  VP$_i$A-Type Clause Arguments Topicality

For anaphoric reference, the text count for VP$_i$A constructions shows that agent arguments are commonly present in the previous three clauses while patients are not. In table 1a, the number of agent arguments receiving a value between 1 and 3 is much higher than patients:

**Table 1a:** *anaphoric reference count for VP$_i$A constructions*

| | AGT | | PAT | |
|---|---|---|---|---|
| | N | % | N | % |
| 1 | 12 | 40 | 0 | 0 |
| 2 | 10 | 33.3 | 0 | 0 |
| 3 | 2 | 6.7 | 0 | 0 |
| >3 | 6 | 20 | 30 | 100 |
| | | | | |
| **1-3** | **24** | **80** | **0** | **0** |
| **>3** | **6** | **20** | **30** | **100** |
| | | | | |
| Total: | 30 | 100 | 30 | 100 |

Similarly, for cataphoric reference, agent arguments are more persistent than patients. Agent arguments with a value >2 occur more frequently than patients in the next ten clauses following its present occurrence:

**Table 1b:** *cataphoric reference count for VP$_i$A constructions*

| | AGT | | PAT | |
|---|---|---|---|---|
| | N | % | N | % |
| 0 | 4 | 13.3 | 15 | 50 |
| 1 | 4 | 13.3 | 7 | 23.4 |
| 2 | 5 | 16.7 | 5 | 16.7 |
| 3 | 8 | 26.7 | 1 | 3.3 |
| 4 | 2 | 6.7 | 1 | 3.3 |
| 5 | 6 | 20 | | |
| 6 | 1 | 3.3 | | |
| 7 | | | 1 | 3.3 |
| 8 | | | | |
| 9 | | | | |
| 10 | | | | |
| | | | | |
| **0-2** | **13** | **43.3** | **27** | **90.1** |
| **>2** | **17** | **56.7** | **3** | **9.9** |
| | | | | |
| Total: | 30 | 100 | 30 | 100 |

Results in tables 1a, b tell us that, for VP$_i$A constructions, agents are highly topical while patients are not topical at all.

## 6 VP$_d$A-Type Clause Arguments Topicality

For VP$_d$A constructions, the count for anaphoric reference points to the conclusion that agent and patient arguments are both topical to some extent. Both agents and patients receive a value between 1 and 3. However, according to table 2a, within three clauses pre-

ceding the present occurrence of the referent, agent arguments appear roughly twice more frequently than patients:

**Table 2a:** *anaphoric reference count for VP$_d$A constructions*

|      | AGT |      | PAT |      |
|------|-----|------|-----|------|
|      | N   | %    | N   | %    |
| 1    | 27  | 31.4 | 10  | 11.6 |
| 2    | 11  | 12.8 | 7   | 8.1  |
| 3    | 8   | 9.3  | 5   | 5.9  |
| >3   | 40  | 46.5 | 64  | 74.4 |
|      |     |      |     |      |
| **1-3** | **46** | **53.5** | **22** | **25.5** |
| **>3**  | **40** | **46.5** | **64** | **74.5** |
|      |     |      |     |      |
| Total: | 86 | 100  | 86  | 100  |

For the same type of construction, the count for cataphoric reference correlates with the results for anaphoric reference. In table 2b, both agent and patient arguments receive a value >2 but agents occur more frequently than patients within the next ten clauses following the current occurrence of the referent:

**Table 2b:** *cataphoric reference count for VP$_d$A constructions*

|      | AGT |      | PAT |      |
|------|-----|------|-----|------|
|      | N   | %    | N   | %    |
| 0    | 30  | 34.8 | 43  | 50.0 |
| 1    | 18  | 20.9 | 15  | 17.5 |
| 2    | 11  | 12.8 | 10  | 11.7 |
| 3    | 10  | 11.6 | 6   | 6.9  |
| 4    | 10  | 11.6 | 3   | 3.5  |
| 5    | 3   | 3.5  | 6   | 6.9  |
| 6    | 1   | 1.2  | 3   | 3.5  |
| 7    | 1   | 1.2  |     |      |
| 8    | 1   | 1.2  |     |      |
| 9    | 1   | 1.2  |     |      |
| 10   |     |      |     |      |
|      |     |      |     |      |
| **0-2** | **59** | **68.6** | **68** | **79.1** |
| **>2**  | **27** | **31.4** | **18** | **20.9** |
|      |     |      |     |      |
| Total: | 86 | 100  | 86  | 100  |

For VP$_d$A constructions, although both agent and patient arguments are topical, tables 2a, b show that agents are more topical than patients.

## 7  VAP-Type Clause Arguments Topicality

For VAP constructions, the anaphoric reference count indicates that patient and agent arguments are both topical since they both receive a value between 1 and 3. However, according to the figures in table 3a, patient arguments occur more frequently than agent arguments:

**Table 3a:** *anaphoric reference count for VAP constructions*

|  | AGT | | PAT | |
|---|---|---|---|---|
|  | N | % | N | % |
| 1 | 6 | 5.8 | 37 | 35.9 |
| 2 | 9 | 8.7 | 5 | 4.9 |
| 3 | 7 | 6.8 | 4 | 3.9 |
| >3 | 81 | 78.7 | 57 | 55.3 |
|  |  |  |  |  |
| **1-3** | **22** | **21.4** | **46** | **44.7** |
| **>3** | **81** | **78.6** | **57** | **55.3** |
|  |  |  |  |  |
| Total: | 103 | 100 | 103 | 100 |

The counting for cataphoric reference of VAP constructions shows a slightly different result from the anaphoric reference count. In table 3b, there are more occurrences of agent referent with value >2 within the next ten clauses following the present occurrence of the referent than there are occurrences of patient referents. However, because the difference is very small, I consider them as equal:

**Table 3b:** *cataphoric reference count for VAP constructions*

|  | AGT | | PAT | |
|---|---|---|---|---|
|  | N | % | N | % |
| 0 | 38 | 36.9 | 31 | 30.0 |
| 1 | 15 | 14.5 | 32 | 31.0 |
| 2 | 20 | 19.4 | 11 | 10.7 |
| 3 | 13 | 12.6 | 10 | 9.8 |
| 4 | 5 | 4.9 | 4 | 3.9 |
| 5 | 7 | 6.7 | 7 | 6.8 |
| 6 | 3 | 3.0 | 1 | 1 |
| 7 | 2 | 2.0 | 7 | 6.8 |
| 8 |  |  |  |  |
| 9 |  |  |  |  |
| 10 |  |  |  |  |
|  |  |  |  |  |
| **0-2** | **73** | **70.9** | **74** | **71.9** |
| **>2** | **30** | **29.1** | **29** | **28.1** |
| Total: | 103 | 100 | 103 | 100 |

Although it is assumed that, for cataphoric reference count, the agents are as topical as patients, the results for anaphoric reference maintains the fact that patient referents are more topical than agent referents in VAP constructions.

## 8  Comparisons

The percentage comparison for the text count for VP$_i$A and VP$_d$A constructions tells us that they are different because, for VP$_i$A-type clauses, patient arguments are not topical while for VP$_d$A-type clauses patients may have a certain degree of topicality:

**Table 4:** *VP$_i$A vs. VP$_d$A constructions*

|          | Anaphoric reference (value 1-3) | | Cataphoric reference (value >2) | |
|----------|-----------|-----------|-----------|-----------|
|          | AGT % | PAT % | AGT % | PAT % |
| VP$_i$A  | 80   | 0    | 56.7 | 9.9  |
| VP$_d$A  | 53.5 | 25.5 | 31.4 | 20.9 |

Furthermore, for VPA constructions (i.e. both VP$_i$A and VP$_d$A), agent arguments are more topical than patients. For VAP constructions, with anaphoric referent count, patients are more topical than agents while, according to the cataphoric referent count, they are assumed to be equally topical:

**Table 5:** *VPA  vs. VAP constructions*

|      | Anaphoric reference (value 1-3) | | Cataphoric reference (value >2) | |
|------|-----------|-----------|-----------|-----------|
|      | AGT % | PAT % | AGT % | PAT % |
| VPA  | 133.5 | 25.5 | 88.1 | 30.8 |
| VAP  | 21.4  | 44.7 | 29.1 | 28.1 |

Finally, considering both VP$_d$A and VAP constructions, we can see that these two types of constructions are somewhat a mirror image of each other. For anaphoric referent count, both agent and patient referents are topical. However, with VP$_d$A-type clauses, the agent is more topical than the patient while with VAP-type constructions the patient is more topical than the agent. For cataphoric referent count, for VP$_d$A-type clauses, the agent is more topical than the patient while with VAP-type clauses, again, agent and patient are assumed to be equally topical:

**Table 6:** *VP$_d$A vs. VAP constructions*

| | Anaphoric reference (value 1-3) | | Cataphoric reference (value >2) | |
|---|---|---|---|---|
| | AGT % | PAT % | AGT % | PAT % |
| VP$_d$A | 53.5 | 25.5 | 31.4 | 20.9 |
| VAP | 21.4 | 44.7 | 29.1 | 28.1 |

Arguments' topicality of the three types of transitive constructions is summarized in table 7:

**Table 7:** *degree of topicality of patient and agent referents*

| Construction type | Argument topicality |
|---|---|
| VP$_i$A | agent >> patient |
| VP$_d$A | agent > patient |
| VAP | agent < patient |

For VPA constructions, the agent is always more topical than the patient. However, depending on the type of VPA clause the patient may not be topical at all (i.e. for VP$_i$A) or caries a certain degree of topicality (i.e. for VP$_d$A). For VAP constructions, patient arguments are more topical than agent arguments but these latter retain a certain degree of topicality.

## 9 Discussion

The text count methodology applied to this analysis allowed us to determine the motive behind the different word orders in Malagasy. As argued throughout the paper, the most topical referent is located in the external argument position and both patient and agent can occupy that position. When the patient is not very topical or not topical at all, one of the VPA-type constructions is used, while when the patient has a considerable degree of topicality the language uses VAP constructions. The topicality count illustrated that VAP and VPA constructions are not just different structurally but they have different functions too.

Furthermore, from this standpoint, the issue in Malagasy is more complicated than asking if VOS is the basic word order or if VOS is a more basic word order than VSO in the language. It can be argued that the external argument has the same privileges/status as "subject" (i.e. being a topical argument in the case of the present study). If we say this, the expectation might be that the "subject" is the agent of the clause. However, as discussed throughout the paper, the external argument can in fact also be a patient. Also, like an external agent argument, this external patient argument can be topical. Therefore, considering Malagasy to have solely a VOS structure in its transitive clauses does not adequately capture the observation that the external argument can be either a patient or an agent.

By way of brief conclusion, it can be said that the VOS claim about Malagasy does not account for why there would be different argument orders in transitive clauses in Malagasy. Consequently, considering the relative topicality of arguments instead arguably leads to a better description of the variation in word order attested.

**Notes**

I would like to thank Doris Payne for her comments for this paper.

Abbreviations

AGT:    agent

DEF:    definite

EA:     external argument

EA:A:   agent external argument

EA:P:   patient external argument

IA:     internal argument

IDEF:   indefinite

PAT:    patient

PST:    past tense

VAP:    verb-agent-patient contruction

VPA:    verb-patient-agent construction

VP$_d$A:   verb-patient-agent construction with definite patient

VP$_i$A:   verb-patient-agent construction with indefinite patient

**References**

Coorman, Ann. 1982. "Topicality, ergativity, and transitivity in narrative discourse: evidence from Chamorro." *Studies in Language* 6(3): 343-374.

Cousins, W. E. 1897. *Introduction sommaire a l'étude de la langue malgache*. Tananarive, Madagascar, Imprimerie de la société de missions de Londres.

Givón, T. 1983. *Topic continuity in discourse: an introduction. Topic continuity in discourse: a quantitative cross-language study*. T. Givón (ed.). Philadelphia, PA, John Benjamins Publishing. 3:4-41.

Givón, T. 1990. *Syntax: A functional-typological introduction*. Philadelphia, PA, John Benjamins Publishing Company.

Keenan, Edward L. and Elinor Ochs. 1979. Becoming a competent speaker of Malagasy. *Languages and their speakers*. Timothy Shopen (ed.). 113-158.

Keenan, Edward L. 1976. Remarkable subject in Malagasy. *Subject and topic*. Charles Li (ed.). New York, NY, Academic Press, inc. 247-301.

Rabenilaina, Roger-Bruno. 1991. "Voix et diathèse en malgache." *Linguisticae investigationes* 15(2): 325-335.

Rainandriamampandry and Georges Ranaivosoa. 1900s. *Hafatry ny ombiasy iray ho an'ny zanany*. Antananarivo, Madagascar, Editions Madprint.

Randriamasimanana, Charles. 2000. Word order typology and Malayo-Polynesian languages. *The 12th International Conference of the New Zealand Asian Studies Society*. Charles Randriamasimanana and Rosemary Haddon (eds.). Massey University, Palmerston North, New Zealand. 266-288.

Thompson, Chad. 1994. Passives and inverse constructions. *Voice and inversion*. T. Givón (ed.). Philadelphia, PA, John Benjamins Publisher. 28:47-63.