

# WORD ORDER TYPOLOGY AND MALAYO-POLYNESIAN LANGUAGES

Charles Randriamasimanana

Massey University

Email: C.H.Rand@massey.ac.nz

## INTRODUCTION

The main purpose of this paper is to provide an overview of the word order typology of the Malayo-Polynesian language family, which covers an immense geographic area of the world, including New Zealand. The description, which will start off from observations made in Greenberg (1966b), Keenan and Comrie (1977) and Hawkins (1983), will utilise Chomsky's X-Bar module of Universal Grammar (1986b) to capture significant generalisations and to outline the importance of languages such as Malagasy, the westernmost branch of Indonesian, and a Polynesian language such as Maori, within the general framework being developed here.

## THE OVERALL PICTURE

In terms of Word Order typology, i.e. according to the relative order of the grammatical subject (S), the verb (V) and the object (O), only the basic orders SVO, VSO and VOS are attested in Malayo-Polynesian; SOV and OVS do not appear in languages of this family. Thus, to illustrate, the relevant word orders in English, Indonesian, Sundanese, Maori, Malagasy and Seediq are as follows:

- (1) Paul chopped the wood.  
S V O
- (2) Paul memotong kayu. [ Bahasa Indonesia ]  
S V O  
'Paul chopped the wood.'
- (3) Paul neukteukan kai. [ Sundanese ]  
S V O  
'Paul chopped the wood.'
- (4) I tâtâhia/tapahia e Paora nga wahie [ Maori ]  
T/A chop/cut-pass erg Paul (fire)wood  
V S O  
'Paul chopped the wood.'

- (5) N-ikapa kitay i Paoly. [Malagasy]  
 past-chop wood art Paul  
 V O S  
 'Paul chopped wood.'
- (6) Mekan wawa ka huling. [Seediq]  
 eat-AF meat part dog  
 V O S  
 'The dog is eating meat.' [A. Holmer, 1996]

The above state of affairs is summarised in the table immediately below.

**Table 1: The Overall Situation**

|                |   |   |         |                                                                                                                                                        |
|----------------|---|---|---------|--------------------------------------------------------------------------------------------------------------------------------------------------------|
| Prototype I:   | S | O | V       | Japanese, Turkish, etc, but not in Malayo-Polynesian                                                                                                   |
| Prototype II:  | S | V | O       | Malay, Indonesian, Javanese, Sundanese, Palauan, Iai, Kaliai-Kove, Arosi, Kapingamarangi                                                               |
|                |   |   | SVO/VSO | Chamorro                                                                                                                                               |
| Prototype III: | V | S | O       | Maori, Ilokano, Cebuano, Bikol, Kalagan, Hiligaynon, Mansaka, Tagalog, West Bukidnon Manobo, Ivatan, Hawaiian, Tahitian, Pukapukan, Samoan, Rennellese |
|                |   |   | VSO/VOS | Rukai, Tsou, Atayal, Tongan                                                                                                                            |
| Prototype IV:  | V | O | S       | Malagasy, Toba Batak                                                                                                                                   |
| Prototype V:   | O | V | S       | Hixkaryana, but not attested in this language family.                                                                                                  |

A very small sample of some thirty-three Malayo-Polynesian languages were used to set up the above chart generally based on published data --as shown in the references provided below-- as well as from elicited sentences provided by native speakers<sup>1</sup> of the relevant languages (notably Indonesian, Ilokano, Sundanese, Maori and Makassar). Prototype refers to the fact that some sentence patterns were unmarked, whereas other patterns were definitely marked in that these other patterns involved, for instance a clefted sentence, i.e. 'It was John who cut the bread', as opposed to the unmarked 'John cut the bread.' In general, I am using the terms

‘prototype’ and ‘marked’ vs ‘unmarked’ as in Comrie (1981); in cases where a language allows more than one basic, unmarked word order, i.e. as in VSO/VOS a slash is used to distinguish between the two possibilities.

### GREENBERG (1966b) AND CORRELATIONS

Greenberg (1966b) assigns natural languages found in the world to twenty-four logically possible types of language, based on the combinations of the four parameters VSO/SVO/SOV, Pr(osition)/Po(s)position, N(oun) G(enitive)/G(enitive) N(oun), N(oun) A(djective)/A(djective) N(oun). These are all unilateral correlations.

In Malayo-Polynesian, the following are attested:

- (a) VSO/Pr/NG/NA
- (b) SVO/Pr/NG/NA.

The following illustrative utterances are from Maori, a VSO language and Sundanese, an SVO language of Indonesia:

- (7) I           tâtâhia/tapahia e Paora nga wahie  
 T/A chop/cut-pass erg Paul (fire)wood  
           V               S           O  
 'Paul chopped the wood.'       [Maori]
- (8) Paul neukteukan kai.  
      S       V       O  
 'Paul chopped the wood.'       [Sundanese]
- (9) I haere a Hone ki Kobe i te taha o Mere  
 T/A go abs John to Kobe prep the side of Mary  
      V S           X       X  
 ma runga motoka  
 by               car  
 'John went to Kobe by car.'   [Maori]
- (10) John indit       ka Kobe kana mobil jeung Mary.  
      John go       to Kobe by car       with Mary  
      S       V       X       X       X  
 'John went to Kobe by car with Mary.' [Sundanese]

- (11) Kua mate te tuahine o Høne  
T/A dead the sister of John  
N G  
'John's sister died.' [Maori]
- (12) Adi John nu awewe maot.  
sister John part dead  
N G  
'John's sister has died.' [Sundanese]
- (13) Kua tuhia e Hone tetahi pukapuka whakamere  
T/A write-pass erg John a book interesting  
N A  
'John has written an interesting book. [Maori]
- (14) John geus nulis buku nu punjul.  
John part write book part interesting  
N A  
'John has written an interesting book.' [Sundanese]

Note that Maori has the word order VSO and that Sundanese has the order SVO, as illustrated in examples (7) and (8); but that both languages have Prepositions as shown in (9) and (10), that the headnoun precedes the genitive or possessor, as in (11) and (12) and that in both languages, the headnoun precedes its adjective, as in (13) and (14). Note that X symbolises an oblique constituent.

Oversimplifying somewhat and leaving aside for the moment S (subject), we can claim the existence of the following tendency (as opposed to an absolute), that in this language family, the following cluster of properties summarized in (c) are attested, whereas those in (d) appear not to be attested:

- (c) VO, Pr, NG, NA  
(d) OV, Po, GN, AN

### **HAWKINS (1983) AND CORRELATIONS**

The following data from some thirty-three different Malayo-Polynesian languages appear to corroborate Hawkins (1983) first three sets of proposed universals, i.e. the so-called implicational

correlations. In essence, each claim goes something like this: If a given language has property A and if in addition, it also has property B, then it is more than likely to also possess property C. One major advantage provided by this framework --as opposed to mere unilateral correlations as in Greenberg (1966b)-- is that the different properties found in diverse natural languages are now assumed to be inter-related and interdependent.

The three relevant proposed universals in Hawkins (1983):

- (e) VSO ---> (NA --> NG)
- (f) Pr --> (NA --> NG)
- (g) Pr & (VSO V SOV) --> (NA --> NG)

In substance, the statement in (e) says that if a language has the VSO order and if in addition, it also has NA (as opposed to AN), then it is more than likely to also have the third property NG (and not GN). The claim made in (f) simply says that if a language has Pr(ositions) --and not Po(st)positions-- and if in addition, it also has NA --and not AN--, then it is more than likely that it will also have the property NG --and not GN. Last, the statement in (g) says that if a language has Pr(ositions) along with the fact that it also has VSO order and if in addition, it has property NA, then it is more than likely that it will also possess property NG. Such properties are summarised on Table 2, wherever there was some doubt a question mark (?) will appear in the relevant cell.

The data shown on Table 2 clearly suggest that the implicational universals proposed in Hawkins (1983) will have to be considered as mere 'tendencies' and not 'as 'absolutes'. Indeed, whereas the first set of proposed universals, i.e. VSO --> (NA --> NG) seem to hold in the Polynesia triangle: Nearly all of those languages have the VSO order and in addition, they also generally have the NA property as well as the third expected property, NG. However, in the Philippines area, even though many of the relevant languages are VSO and also have the property NA, some of the languages appear to have already shifted to AN; nevertheless, most of the languages analysed seem to also have NG as its third property despite the fact that the alternative GN order is possible in languages such as Bikol, Mansaka and West Bukidnon Manobo.

**Table 2: Typology of Malayo-Polynesian**

|                    |                     |       |      |            |               |
|--------------------|---------------------|-------|------|------------|---------------|
| <b>INDONESIA</b>   |                     |       |      |            |               |
| Indonesian         | SVO                 | NG    | Prep | NA         | NRel          |
| Malay              | SVO                 | NG    | Prep | NA         | NRel          |
| Javanese           | SVO                 | NG    | Prep | NA         | NRel          |
| Makassar           | SVO                 | NG    | Prep | NA         | NRel          |
| Sundanese          | SVO                 | NG    | Prep | NA         | NRel          |
| Toba Batak         | VOS                 | NG    | Prep | NA         | NRel          |
| Malagasy           | VOS                 | NG    | Prep | NA         | NRel          |
| <b>PHILIPPINES</b> |                     |       |      |            |               |
| Ilokano            | SVO                 | NG    | Prep | AN         | NRel          |
| Cebuano            | VSO                 | NG    | Prep | AN         | NRel          |
| Bikol              | VSO                 | NG/GN | Prep | AN         | NRel          |
| Kalagan            | VSO                 | NG    | Prep | NA/<br>AN  | NRel          |
| Hiligaynon         | VSO                 | NG    | Prep | NA/<br>AN  | NRel          |
| Mansaka            | VSO                 | NG/GN | Prep | NA/<br>AN  | NRel          |
| Tagalog            | VSO                 | NG    | Prep | NA         | NRel          |
| W.B. Manobo        | VSO                 | NG/GN | Prep | AN         | NRel          |
| Ivatan             | VSO                 | NG    | Prep | AN         | NRel          |
| <b>TAIWAN</b>      |                     |       |      |            |               |
| Tsou               |                     | NG    |      | AN         | NRel/<br>RelN |
| Atayal             | VOS/<br>SVO/<br>VSO | NG    | Prep | AN/<br>NA  | RelN          |
| Rukai              | VSO/<br>VOS         | NG    | Prep | AN         |               |
| Seediq             | VOS                 | NG    |      | NA         | NRel          |
| <b>MICRONESIA</b>  |                     |       |      |            |               |
| Chamorro           | VSO/<br>SVO         | NG    | Prep | NA         | NRel          |
| Palauan            | SVO                 | NG/GN | Prep | NA         | NRel          |
| <b>MELANESIA</b>   |                     |       |      |            |               |
| Kaliai-Kove        | SVO                 | NG    | Prep | NA         | NRel          |
| Arosi              | SVO                 | NG    | Prep | NA         | NRel          |
| Iai                | SVO                 | NG    | Prep | NA         | NRel          |
| <b>POLYNESIA</b>   |                     |       |      |            |               |
| Hawaiian           | VSO                 | NG    | Prep | NA         | NRel          |
| Maori              | VSO                 | NG    | Prep | NA         | NRel          |
| Pukapukan          | VSO                 | NG    | Prep | NA         | NRel          |
| Samoa              | VSO                 | NG    | Prep | NA/<br>AN? | NRel          |
| Kapingamarangi     | SVO                 | NG    | Prep | ?          | NRel          |
| Rennellese         | VSO                 | NG    | Prep | NA         | NRel          |
| Tongan             | VSO/<br>VOS         | NG    | Prep | NA/<br>AN? | NRel          |
| Tahitian           | VSO                 | NG/GN | Prep | NA         | NRel          |

As far as the second set of proposed universals are concerned, Pr --> (NA --> NG), it seems to hold systematically in the Indonesia area, in the Micronesia area, in the Melanesia area as well as by and large, in the Polynesia area although the situation fluctuates considerably in the Philippines. But even there, the tendency is still apparent in languages such as Kalagan, Hiligaynon and Mansaka: In all three languages, prepositions predominate (Pr), the headnoun generally precedes its adjective (NA) and the headnoun precedes the genitive or possessor (NG).

As far as the last universal goes, Pr & (VSO v SOV) --> (NA --> NG), it only holds partially for Malayo-Polynesian since the SOV order is simply not attested in this language family, as will become evident from a rapid perusal of Table 2. One last factor shown on the above table relates to the relative order of a headnoun and its accompanying relative clause: In all Malayo-Polynesian languages --with only one possible exception (Atayal), the relevant property is NRel (not RelN).

### **CHOMSKY (1986b) AND WORD ORDER PHENOMENA**

Lisa Travis (1984) and Eithne Guilfoyle (1990) illustrate possible applications of Chomsky's Universal Grammar to Malayo-Polynesian languages. The methodology used in this type of framework is very different from the one adopted in statistically-based approaches such as the one used in Greenberg (1966b) and Hawkins (1983). A precursor to the latest methodology can be found in Chung (1978), where the author concentrates on a subfamily of Malayo-Polynesian, namely Polynesian in order to be able to look into relevant semantico-syntactic processes in some depth. As a result, this latest methodology turns out to be more reliant on contemporary linguistic theory than the more empirically-oriented, anthropological type of framework utilised in Greenberg (1966b), for instance. The latter also tends to focus on a more manageable range of natural languages for the purposes of cross-linguistic comparisons.

One interesting set of properties found in Chung (1978) and worth noting is that, for example, major syntactic processes such as those associated with Equi-constructions are restricted to the grammatical subject in Polynesian languages such as Maori and

even in ergative languages such as Samoan. (See, for instance, Chung 1976: 130-131). This means that even in such Polynesian languages it is legitimate to refer to notions such as grammatical subjects and that a reanalysis of a sequence Verb-Agent-Patient into Verb-Subject-Object is not impossible, where Agent is now re-interpreted as a Subject and no longer like an X or Oblique constituent --arising from the application of passive.

As far as word order phenomena are concerned, the unilateral correlations summarised earlier as

- (c) VO, Pr, NG, NA
- (d) OV, Po, GN, AN

as well as the universal implicational tendencies

- (e) VSO ---> (NA --> NG)
- (f) Pr --> (NA --> NG)
- (g) Pr & (VSO v SOV) --> (NA --> NG)

can now be described economically using Chomsky's X-Bar Syntax framework in terms of head and complement :

(h) The constructions represented as VO, Pr, NG, NA can now be simply described as head-initial structures in that in the sequence VO, V is the head of construction and O is the complement; in a structure involving a preposition, Pr(osition) is the head and the accompanying Noun the complement; in an NG sequence, N(oun) is the head and G(enitive) or possessor is the complement; and in the NA structure, N(oun) is the head and A(djective) is the complement. In other words, in (c) above, we have head-initial constructions, whereas in (d), we have head-final structures.

(i) With regard to S(ubject) --which arises in (e) and (g)above --we adopt the formal representation proposed in Chomsky (1986b), i.e. the NP occupying the Specifier position in a tree diagram<sup>13</sup> of the following sort:



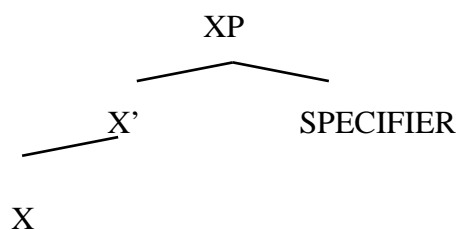


Figure 3

where X can be any major grammatical category {Noun, Verb, Preposition/Postposition, Adjective} projecting itself onto two different levels, i.e. one-bar level (the level of strict subcategorization, indicated by the acute accent) is the first level of projection of the lexical item and where XP, equivalent to two-bar level, is the maximal projection of the same lexical item.

Note that for the purpose of this paper, we need to distinguish between two types of projections: lexical items and functional items. Thus, above the projection<sup>2</sup> for the lexical item corresponding to the lexical category Verb, for example, there will also be another projection for Tense, a functional category. In a language like Malagasy, there will also be another projection above that of the functional category Tense, i.e. AGR(eement) to account for agreement phenomena illustrated in Randriamasimanana (1997). Indeed in Malagasy, whether an overt subject will appear or not depends on the presence or absence of this AGR, as shown in Randriamasimanana (1988b).

### THE NOTION OF ‘BARRIER’ AND PASSIVE

Chomsky (1986b) proposes a definition of government in terms of exclusion rather than domination. This means that in a VOS language like Malagasy the relationship between V and O will be of crucial importance. Thus, if we focus on a verbal predicate involving a transitive verb, we can get the following configuration:

(iii) V ny N

where V stands for the head of construction and ‘ny N’ is its complement. We can analyse the above sequence as:

(iii) V                    ny    N  
 (iv) X    [            Z    Y ]  
                                          barrier

where  $V = X$ , the head stands on its own, whereas 'ny + N', the complement is made up of two different elements, Z as well as Y. Crucially the definite article 'ny' 'the' is a barrier to government of N by V. Simplifying Chomsky (1986b) somewhat: In a sequence where we have two constituents X and Y, if there is an intervening element, say, Z, which is part of Y but not of X, then Z can act as a barrier between X and Y.

One major reason why we are interested in this concept is the following: In the overall grammar for Malayo-Polynesian languages such as Malagasy, there exists some kind of presumably formal mechanism which identifies and forces a given constituent to be selected as grammatical subject within a clause. In the case of Malagasy, the definite article 'ny' 'the' in the configuration described above will not be a barrier if and only if<sup>3</sup> it contains a feature indicating that it is antecedent-governed. Otherwise, it always acts as a barrier and as a consequence, the definite direct object 'ny + N' will have to move into the subject position within the sentence, i.e. passivization will have to take place.

### **SUBJECT SELECTION PRINCIPLE**

Malagasy is a well-behaved and consistent head-initial language where the relevance of the notion of barrier seems rather straightforward. However, there appear to be other VOS languages in the Austronesian family --such as Seediq-- which do not seem to behave in the same manner. This has led linguists such as Arthur Holmer (1996, 1997) to propose a so-called Subject Choice model of voice<sup>4</sup>. Also Joao Costa (1997) put forth his Subject Selection rule based on Optimality Theory and Discourse Analysis. In this article, I will only present illustrative examples based mainly on the presence within the utterance of universal quantifiers such as 'daholo' and 'rehetra', both meaning 'all' in Malagasy and each adjoined to a grammatical subject.

To highlight the crucial importance of something like a Subject Selection principle, note that in Malagasy, for instance, the nominal constituent which is accompanied by a universal quantifier<sup>5</sup> such as 'daholo' or 'rehetra' 'all' has to become a grammatical subject. Thus, to express the following proposition

3. Paul saw all the students

where the constituent 'all the students' occupies a direct object position, at least in languages such as English, such a constituent will have to be promoted to subject position in Malagasy first, as illustrated in (16)c --along lines suggested in Keenan and Comrie (1977) and in essence, this means passivization; otherwise, an irretrievably ungrammatical utterance will ensue, as shown in (16)b. below.

- (16)a. N-ahita           mpianatra        i Paoly.  
 past-see           students         art Paul  
 'Paul saw         (some) students.'
- b. \*N-ahita         ny mpianatra   daholo/rehetra i Paoly.  
 past-see         the students   all               art Paul  
 'Paul saw all the students.'
- c. Hita-n'i Paoly         daholo ny mpianatra (rehetra).  
 pass-see-byart Paul   all     the students   (all)  
 Lit. 'All the students were (all) seen by Paul.'  
 'Paul saw all the students.'

In connection to (16)c., note first of all, as explained in Randriamasimanana (1997) that there is a need to underline the fact that the grammatical subject of a sentence occupies a derived, non-theta position, i.e. not encoding a given semantic role in that it no longer depends on its verb the same way as a direct object would, for instance. With respect to (16)c. and specifically, the presence of two universal quantifiers 'daholo' and 'rehetra' --both meaning 'all'--, 'daholo' obligatorily shows up as the subject internal<sup>6</sup> to the verbal projection, whereas 'rehetra' optionally surfaces on the so-called external<sup>6</sup> subject, a derived, non-theta position as just explained. The explanation for the ungrammaticality of sentence (16)b. would be attributable to the notion of proper government as defined earlier, i.e. that the direct object constituent 'ny mpianatra daholo/rehetra' 'all the students' is not properly governed by its verb 'n-ahita' 'saw'; and the reason for this is simply that the element 'ny' 'the' acts as a barrier to government of 'ny mpianatra rehetra' by 'n-ahita' 'saw': The intervening 'ny' blocks proper government of the definite direct object by its governing verb 'saw'. Linguistic facts like these explain to a certain extent why in a language like Malagasy the passive voice<sup>7</sup> is much more frequent than the active voice, i.e. in the order of 35% to

52% of all utterances in a given text, as opposed to a mere 5 to 15% for languages like English, as noted in Randriamasimanana (1999a).

At this stage, it has to be noted that quite independently of discourse-level phenomena (usually associated with existential presupposition, quantification and the like), as noted in Randriamasimanana (1997), the complementizer ‘ny’ and the definite article or specifier ‘ny’ in a language such as Malagasy also act as a barrier to government. Simplifying somewhat, the above would allow us the following analysis for the relevant sentences of Malagasy:

- (17) M-ikasa      ny      h-andeha      PRO    i Paoly.  
 pres-intend    Comp fut-go      ---      art Paul  
 ‘Paul intends to go.’
- (18) Tsara ny      m-andeha      any      PRO.  
 good spec pres-go      there      -----  
 ‘To go there is good.’

The fact that Complementizer ‘ny’ in (17) and specifier ‘ny’ in (18) act as barriers means that the empty subject PRO is insulated from government from the matrix predicates ‘mikasa’ in (17) and ‘tsara’ in (18). Note crucially that both empty categories occupy a grammatical subject position within each sentence. Finally, note that whereas passive is mandatory in (16)c., with a transitive verb like ‘m-ikasa’ ‘to intend’ as the matrix verb, it is optional in (17), as shown immediately below:

- (17') Ny h-andeha PRO no no-kasa-in'      i Paoly.  
 spec fut-go      --- part pass-intend-by art Paul  
 Lit: 'To go is intended by Paul.'  
 'Paul intended to go.'

### **ASYMMETRY BETWEEN SUBJECT AND OBJECT**

The asymmetry between the treatment of grammatical subject and object will flow from the above observation: The grammatical subject no longer dependent on its verb will not require any special kind of marking when it is moved to the front, for instance; whereas an object still very much dependent on its verb will require some kind of marking when it has to move elsewhere. Now, in a noun-coding type of language like English, the marking that is required is

nominal in nature and, in according with Chomsky (1986b), will leave behind a (nominal) trace in situ, where the moved noun phrase used to be. On the other hand, Malayo-Polynesian languages such as Malagasy tend to be verb-coding and as such, the marking left behind involves the verb itself and not just the accompanying nominal arguments. This typically translates into encoding the initial position of the nominal argument onto an affix of the governing verb, hence the frequent use of passive in such languages.

First, let us substantiate the claim that when a grammatical subject is fronted in a sentence in a Malayo-Polynesian language such as Malagasy, nothing happens to the verb on which it seems to depend. By contrast, when an object gets fronted, then something has to happen to the governing verb. Typically in Malagasy, the subject may be fronted but the verb remains in the active voice, as in:

(20) N-andeha    iza?  
       past-go     Who  
       ‘Who went?’ [ Echo question]

(21) Iza    no    n-andeha --- ?  
       Who   part   past-go ---  
       ‘Who went?’

In (20) we have an echo type of question where the grammatical subject remains in situ and does not move to the front. By contrast, in (21) the grammatical subject ‘iza’ ‘who’ has been fronted. Crucial observation: The verb remains in the active voice. Contrast the above set of data with the following:

(22) N-ahita        an’    iza    ianao?  
       past-see     DO    who    you  
       ‘You saw who?’    [Echo question]

(23) \*(An) iza    no    n-ahita        ---    ianao?  
       (DO) Who   part   past-see        ---    you  
       ‘Who did you see?’

(24) Iza    no    hita-nao        ---    ---  
       Who   part   pass-seen-by you    ---    ---  
       Literally: ‘Who was seen by you?’  
       ‘Who did you see?’

where in (22) we have an echo type question, with the Wh-word left in situ. If, as shown in (23), an attempt is made to move the Wh-word 'iza' 'who' to the front without doing anything to the verb on which it depends, then the sequence becomes irretrievably ungrammatical. On the other hand, if as seen in (24), the object is first promoted to subject --as indicated by the passive voice affix on the verb-- and subsequently moved to the front, the ensuing sequence is perfectly grammatical.

## CONCLUSIONS

We have seen that the following basic word orders are attested in Malayo-Polynesian languages: SVO, VSO and VOS. Researchers in historical Malayo-Polynesian linguistics such as Pawley (1997) suggest that the centre for dispersal of the Austronesian peoples is the island of Taiwan, where precisely a number of so-called aboriginal languages have been found. Judging from data from different sources, many if not most such Austronesian languages found in Taiwan have the VOS word order.

In order to derive the SVO<sup>8</sup> order attested for Western Malayo-Polynesian languages such as Indonesian or Malay, we can envisage fronting of the final element, a move which given the asymmetry between subject and object, does not trigger any special marking process on the verb. On the other hand, in the drift from VOS to VSO there may be an intermediate stage where during passivization VOS becomes something like V-Agent-Patient; subsequently the Agent gets re-analysed as Subject and the Patient is re-interpreted as an Object, finally yielding the VSO word order. Hohepa (1969) proposes such an accusative-to-ergative drift in Polynesian while Coppenrath and Prevost (1975) note that there is in contemporary Tahitian a tendency to passivize even on utterances already in the passive<sup>9</sup> such as (26) below

- |      |            |                 |     |
|------|------------|-----------------|-----|
| (25) | a          | fana-hia        | vau |
|      | perfective | be born-passive | I   |
|      |            | 'I was born.'   |     |
| (26) | a          | fanau           | vau |
|      | perfective | be born         | I   |
|      |            | 'I was born.'   |     |

In conclusion, while it appears that an empirically-based type of approach like the one found in Greenberg (1966b) associated with anthropological linguistics may provide us with a bird's eyeview of the situation as far as word order is concerned in Malayo-Polynesian, a rigorous and systematic approach such as the one utilised in Hawkins (1983) would yield some useful universal tendencies. However, in order to capture the essence of a drift from, let us say, a nominative-accusative system to a so-called ergative-absolutive system, we need an in-depth type of analysis such as is usually proposed in a Chomskyan type of formal approach.

Within such a framework, adopting something like the so-called Subject Selection principle will enable us to understand why Universal Quantification in a VOS language like Malagasy will tend<sup>10</sup> to trigger passive when the object is a definite phrase: The grammatical subject position in this language semantically correlates with definiteness, as shown in Keenan (1976) and is characterised by the presence of the definite article 'ny' 'the' inside the relevant phrase. By contrast, a typical direct object position tends to encode lack of definiteness<sup>11</sup> and absence of the definite article 'ny' 'the'.

It looks as though the functional head of the inflections for tense/aspect in Malagasy may have in it a feature like [+definite], which automatically attracts to it any definite phrase contained in the lexical projection to the exclusion of an indefinite phrase with a feature such as [-definite]. Such semantic properties of the definite phrase correlate with the presence of the definite article 'ny' 'the', which acts like a barrier between the head verb and its putative direct object, hence the necessity to move such a definite phrase into the grammatical subject position --in other words, passivization in a verb-coding language like Malagasy.

In brief, there appears to be a combination of two different types of motivation at work here for the movement of the offending definite phrase, which is probably best analyzed as a projection of the definite article 'ny' 'the', following Abney (1987): (i) first, a push in that if at an earlier or lexical projection stage of a relevant derivation a definite phrase has not been specifically marked as 'antecedent-governed', then it will need to move on up to the relevant functional head, where presumably its case-features will be checked --pretty much along the lines sketched within a minimalist type of framework as described in Radford (1997) (ii) second, a pull in that the [ + definite ] feature contained in the highest functional head (either agreement<sup>12</sup> or tense) will attract to it -- for

instance into the Specifier position shown on Figure 3 above<sup>13</sup>--any definite phrase from within the relevant lexical projection.

Ultimately, when passive becomes more frequent than its active counterpart in the language as a whole, there is a re-interpretation imposed by native speakers on the linguistic system, giving rise to the so-called ergative analysis<sup>14</sup>: From VOS, the language will have shifted to the VSO word order.

### Endnotes

1. All of the Maori utterances used in this paper were graciously provided by Professor Tamati Reedy, School of Maori and Pacific Development Studies, Waikato University, New Zealand, who is a native speaker of Maori. All of the Indonesian and Sundanese sentences were given by Mr. R. Hadjadibrati of Monash University in Melbourne, Australia, who is a native speaker of both languages. All of the Ilokano sentences came from Dr. Wilhelmina Drummond of the Massey University College of Education, who has worked and lived in New Zealand for the past twenty-six years.
2. Given the findings reported in Randriamasimanana (1998) Clausal architecture and movement verbs in Malagasy, where all verbs are to be analysed as comprising a higher V and a lower V, this must be considered a simplification: The higher V combines with a tense-marker, whereas the lower V can only combine with an aspect-marker.
3. This explains why sentences such as the one involving a definite phrase direct object quoted in Randriamasimanana (1997: 493) used in isolation sound like Foreigner Talk:
 

(26a) Mijery ny vehivavy ny lehilahy.  
       see the woman the man  
       ‘The man sees the woman.’  
       Bennett 1986: 25, example 1)
4. In order for Holmer (1997)'s Subject Choice model of voice to be applicable to Malagasy, we need (i) to distinguish between a definite phrase, which is an argument of the verb and a definite phrase, which is only an adjunct, as shown in Randriamasimanana (1997b); (ii) only a definite phrase argument will trigger passive on the verb on which it depends, an adjunct will not. The distinction between an argument and an adjunct is quite sharp in Malagasy, but not in Seediq, another Malayo-Polynesian language described by Holmer (1996, 1997) and with the same VOS order as Malagasy. Apparently in many Austronesian languages found on Taiwan, incorporation seems to have blurred the distinction between an argument and an adjunct.
5. There exist at least three types of quantifiers in Malagasy. The first type, restricted quantifiers, specifically Universal



Quantifiers such as ‘daholo’ ‘all’ and ‘rehetra’ ‘all’, can only show up as subjects: In an internal VP subject position for the first; while the second can only appear (optionally) in the external subject or InflectionProjection position.

The second type, unrestricted quantifiers such as ‘maro’ ‘many, numerous’ and ‘vitsivitsy’ ‘a few’ and other indefinite quantifiers can occupy (but are not restricted to) a non-subject position:

N-ividy            boky maro            i Paoly.  
 past-buy            book(s) numerous    art Paul  
 ‘Paul bought many books.’

In addition, a third type of quantifiers, i.e. Existential Quantifiers like ‘m-isy’ ‘present-exist’ show up inside the projection for the inflections for tense/aspect in Malagasy:

N-isy mpianatra            n-ividy boky maro.  
 past-exist student(s)    past-buy book(s) numerous  
 Literally: ‘There existed some students such that they bought numerous books.’  
 English: ‘Some students bought many books.’

Last but not least, note that the referent of the definite phrase in the external subject position always refers to all the items within the identified set of objects --without any exception. And, of course, there is an existential presupposition accompanying the referent occupying this subject position.

6. The existence of this so-called VP-internal subject as opposed to the external subject position was demonstrated by Guilfoyle, Eithne, Hung, Henrietta & Travis, Lisa. 1992. Spec of IP and Spec of VP: Two subjects in Austronesian languages.
7. See Charles Randriamasimanana (1999a) for a discussion of this specific point. Also see Cecile Manorohanta (1998) for similar results regarding passive in Malagasy.
8. We are dealing here with surface word order and NOT with some putative underlying word order, as it is attested in adult native speaker grammar of the relevant language. So, for example, we are not referring here to the possibility of predominance of SVO word order in Malagasy Child Speech: Indeed, before the emergence of functional categories and projections like tense and aspect in Malagasy, there may be a tendency to have the SVO word order in this variety. From the author’s experience with Baby Talk, for example, Malagasy children may contrast S and O maximally when the verb does not yet merge with any affix for tense and aspect, as in:

Keke ba nay  
 Baby kiss us  
 ‘Baby kiss me.’

Nay ba Keke  
 Us kiss Baby

'I kiss Baby.'

Note the root noun 'ba' for 'kiss', which is the relevant predicate in the above utterances: It has no functional affix of any kind for tense/aspect as of yet. Also note the contrast in word order afforded by the position of S to the left of V and the O to its right.

9. Presumably in (26) the utterance was already in the passive voice, as is quite apparent through the literal translation in English of the predicate 'be born'. However, native speakers of contemporary Tahitian still passivize on 'fanau' 'be born', which is already in the passive voice! We know that they do this since they now use the passive suffix '-hia' on 'fanau'.
10. This is best viewed as a tendency and not an absolute since the environment in which the definite phrase occurs is very important. All throughout this paper, it has been assumed that we are dealing with single utterances in isolation, i.e. that each relevant sentence of Malagasy occurs discourse-initially.
11. Such a proposition is certainly supported by the distribution of indefinite quantifiers in Malagasy, as explained in footnote 5 above.
12. See Randriamasimanana (1998b) for the importance of the distinction between an AGR(ement) projection and one involving tense.
13. Given the Malagasy Child Language data shown under footnote 8 above, it is quite conceivable that the position of the Specifier could be as posited in Kayne (1994), i.e. to the left of the head rather than to its right --contrary to what appears in Figure 3. Nothing that is claimed in the present article crucially hinges on this detail, which will have to await further research. At this stage I am merely describing the situation in Adult Malagasy speech rather superficially for the purpose of the present paper.
14. See Yung-Li Chan (1997) for a treatment of Seediq and Kavalan, two Austronesian languages spoken on Taiwan as split ergative languages.

## References

- Abney, S.P. (1987). *The English noun phrase in its sentential aspect*. PhD dissertation, MIT, Cambridge, Mass.
- Bauer, Winifred (1997). *The Reed Reference Grammar of Maori*. Auckland, N.Z.: Reed Books.
- Bell, Sarah (1976). *Cebuano subjects in two frameworks*. Unpublished PhD dissertation, University of British Columbia, Canada.

- Bunye, Maria and Yap, E.P. (1971). *Cebuano Grammar Notes*. PALI Language Texts: Philippines. Honolulu, Hawaii: University Press of Hawaii.
- Cappell, Arthur (1971). *Arosi Grammar*. Canberra: Linguistic Circle of Canberra.
- Chang, Yung-li (1997). *Voice, Case and Agreement in Seediq and Kavalan*. PhD dissertation. National Tsinghua University, Taiwan.
- Chomsky, Noam (1986b). *Barriers*. Cambridge, Mass.: The MIT Press,
- Chung, Sandra (1994). Word Order and the Subject in Chamorro. Talk given in the Department of Linguistics, July-August 1994, Victoria University of Wellington, Wellington, New Zealand.
- Chung, Sandra (1978). *Case Marking & Grammatical Relations in Polynesian*. Austin, Texas and London, England: University of Texas Press.
- Churchward, Maxwell (1953). *Tongan Grammar*. London: Oxford University Press.
- Clark, Ross (1976). *Aspects of Proto-Polynesian Syntax*. Auckland, N.Z.: Linguistic Society of New Zealand.
- Collins, Grace (1970). *Two views of Kalagan grammar*. Unpublished PhD dissertation, UCB. Microfilm, Ann Arbor, Michigan.
- Comrie, Bernard (1981). *Language Universals and Linguistic Typology*. Chicago: University of Chicago Press.
- Coppenrath, Hubert and Prevost (1974.) *Grammaire approfondie de la langue tahitienne (ancienne et moderne)*. Papeete, Tahiti: Lirairie Pureora,
- Costa, Joao (1997). *Word Order Typology in Optimality Theory*. Manuscript, Universidade de Lisboa & HIL, Leiden University.
- Counts, David (1969). *A Grammar of Kaliyai-Kove*. Honolulu: University of Hawaii Press.
- Dodds, R.W. (1977). *Malay*. Teach Yourself Books. Hodder and Stoughton.
- Elbert, Samuel & M.K. Pukui (1979). *Hawaiian Grammar*. Honolulu, Hawaii: University Press of Hawaii.
- Elkins, Richard (1968). *Manobo-English Dictionary*. Honolulu, Hawaii: University of Hawaii Press.

- Greenberg, J. (1966b). Some universals of grammar with particular reference to the order meaningful elements. In *Universals of Language*, second edition. Cambridge, Mass.: MIT Press: 73-113.
- Guilfoyle, E., H. Hung & Travis L. (1992). Spec of IP and Spec of VP: Two Subjects in Austronesian Languages. *NLLT* 10: 375-414.
- Guilfoyle, Eithne. (1990). *Functional Category Projection and its Applications*. PhD dissertation, Department of Linguistics, McGill University, Montreal, Canada
- Hawkins, John (1983). *Word Order Universals*. New York: Academic Press.
- Hohepa, Patrick (1967). *A Profile Generative Grammar of Maori*. Indiana University Publications in Anthropology and Linguistics, Memoir 20.
- Hohepa, Patrick (1969). The accusative-to-ergative drift in Polynesian languages. *The Journal of the Polynesian Society*, Volume 78, N 3: 295-329. September 1969.
- Holmer, Arthur (1996). *A parametric grammar of Seediq*. Travaux de l'Institut de Linguistique de Lund 30. Lund, Sweden: Lund University Press.
- Holmer, Arthur (1997). A structural analysis of Instrument focus. *Papers for the Eighth International Conference on Austronesian Linguistics (December 28-30, 1997)*, Taipei, Taiwan. Pp 329-346.
- Huang, Lilian (1995). *A Study of Mayrinax Syntax*. Taipei: Taiwan. The Crane Publishing Co. Ltd
- Huang, Lillian (1994). The syntactic structure of Wulai and Mayrinax Atayal: A comparison. Paper read at 7-ICAL, University of Leiden, Leiden, The Netherlands.
- Huang, Lillian (1993). *A Study of Atayal Syntax*. Taipei, Taiwan. The Crane Publishing Co. Ltd.
- Josephs, Lewis (1975). *Palauan Reference Grammar*. PALI Language Texts: Micronesia. Honolulu, Hawaii: University Press of Hawaii.
- Kayne, Richard (1994). *The Antisymmetry of Syntax*. MIT Press, Cambridge, Mass.
- Keenan, Edward and Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8.1: 63-99.

- Keenan, Edward (1976). Towards a universal definition of 'subject'.  
In Charles N. Li, editor, *Subject and Topic*. Academic Press,  
New York.
- Lewis, M.B. (1969). *Sentence Analysis in Modern Malay*. Cambridge:  
Cambridge University Press.
- Li, Paul (1973). *Rukai Structure*. Unpublished PhD dissertation,  
University Microfilms, Ann Arbor, Michigan.
- Lieber, Michael and Dikepa, K.H. (1974). *Kapingamarangi lexicon*. PALI Language Texts: Polynesia. Honolulu, Hawaii:  
University Press of Hawaii.
- Manorohanta, Cecile (1998). A quantitative study of passive in Malagasy. Address to AFLA V, Department of Linguistics,  
Honolulu, Hawaii, March 26-29, 1998.
- Mintz, Malcolm (1971). *Bikol Grammar Notes*. PALI Language Texts: Philippines. Honolulu, Hawaii: University of Hawaii Press.
- Mosel, Ulrike and Hovdhaugen, E. (1992c). *Samoan Reference Grammar*. Oslo, Norway: Scandinavian University Press.
- Nababan, Partabas (1966). *Toba Batak, A Grammatical Description*. Unpublished PhD dissertation. University Microfilms International. Ann Arbor, Michigan.
- Pawley, Andrew (1997). Chasing Rainbows: Implications of the Rapid Dispersal of Austronesian Languages for Subgrouping and Reconstruction. Invited Talk to the 8th International Conference on Austronesian Linguistics, Academia Sinica, Taipei, 28-31 December 1997. Taipei, Taiwan: *Papers for 8-ICAL*, pp 441-482.
- Percival, W.K. (1964). *Toba Batak Grammar*. Unpublished PhD dissertation. Yale University.
- Rabenilaina, Roger-Bruno (1987). *Grammaire-Lexique du Malgache*. Thèse pour le Doctorat d'Etat, Université de Paris VII. Paris, France.
- Radford, Andrew (1997). *Syntactic theory and the structure of English. A minimalist approach*. Cambridge University Press.
- Rajemisa-Raolison, Régis (1995). *Rakibolana Malagasy*. Fianarantsoa, Madagasikara: Ambozontany.
- Randriamasimanana, Charles (1986). *The Causatives of Malagasy*. Honolulu, Hawaii: University of Hawaii Press.

- Randriamasimanana, Charles (1997). *Malagasy and Universal Grammar. Proceedings of the Seventh International Conference on Austronesian Linguistics*, pp 483-496. University of Leiden, Leiden, The Netherlands.
- Randriamasimanana, Charles (1997). Focus in Malagasy. Paper presented at the Eighth International Conference on Austronesian Linguistics, December 28-31, 1997. Taipeh, Taiwan.
- Randriamasimanana, Charles (1998b). Evidence for null subjects in Malagasy. Papers from the fifth annual meeting of the Southeast Asian Linguistics Society 1995. Shobhana L. Chelliah and Willem J. de Reuse, editors. Arizona State University, Tempe, Arizona. Pp 303-315.
- Randriamasimanana, Charles (1999a). The Malayo-Polynesian Origins of Malagasy. In Neile A. Kirk & Paul J. Sidwell, editors, *From Neanderthal to Easter Island. A tribute to, and a celebration of, the work of W. Wilfried Schuhmacher presented on the occasion of his 60th birthday*. Association for the History of Language, 1999: 26-43, Melbourne, Australia. AHL Studies in the Science & History of Language, Volume 2.
- Reid, Lawrence (1966). *An Ivatan Syntax*. PALI. Honolulu, Hawaii: University Press of Hawaii.
- Robson, Stuart (1992). *Javanese Grammar for Students*. Clayton, Victoria, Australia: Centre for Asian Studies, Monash University,
- Schachter, Paul and Fe Otnes (1972). *Tagalog Reference Grammar*. Berkeley, Los Angeles, London: University of California Press.
- Sityar, Emily (1994). Interrogative and Indefinite Verbs in Cebuano. Paper read at 7-ICAL, 23 August 1994. University of Leiden, Leiden The Netherlands.
- Suhandano (1994). *Grammatical Relations in Javanese*. Sub-thesis for the MA degree. Canberra, Australia: The A.N.U.
- Svelmoe, Gordon (1990). *Mansaka Dictionary*. Dallas: Summer Institute of Linguistics.
- Topping, Donald (1973). *Chamorro Reference Grammar*. PALI Language Texts: Micronesia, Honolulu, Hawaii University Press of Hawaii.

- Travis, Lisa (1984). *Parameters and Effects of Word Order Variation*. Unpublished PhD dissertation. MIT, Cambridge, Mass. USA.
- Tryon, Darrell (1968). *Iai Grammar*. Canberra: Australian National University.
- Tung, T'ung-ho (1964). *A descriptive study of the Tsou language*, Taipei, Formosa: Institute of Philology and History, Academia Sinica Special Publication no 48.
- Wolfenden, Elmer (1971). *Hiligaynon Reference Grammar*. PALI Language Textbooks: Philippines. Honolulu, Hawaii: University Press of Hawaii.