ZASPiL Nr. 34 – October 2004

**Proceedings of AFLA 11
ZAS, Berlin 2004**

Editor: Paul Law

Table of Contents

Foreword

The papers in this volume were presented at the eleventh meeting of the Austronesian Formal Linguistics Association (AFLA 11), held from April 23-25 at the Zentrum für Allgemeine Sprachwissenschaft, Berlin, Germany. The conference was organized by Hans-Martin Gärtner, Joachim Sabel, and myself, as part of the research project *Clause Structure and Adjuncts in Austronesian Languages*. We gratefully acknowledge the financial support by the German Research Foundation (Deutsche Forschungsgemeinschaft). We would like to thank Wayan Arka, Agibail Cohn, Laura Downing, Silke Hamann, S J Hannahs, Ray Harlow, Nikolaus Himmelmann, Yuchua E. Hsiao, Lillian Huang, Ed Keenan, Glyne Piggott, Charles Randriamasimanana, Joszef Szakos, Barbara Stiebels, Jane Tang, Lisa Travis, Noami Tsukido, Sam Wang, Elizabeth Zeitoun, Kie Ross Zuraw, and Marzena Zygis for reviewing the abstracts. We are thankful to Mechthild Bernhard, Jenny Ehrhardt, Fabienne Fritzsche, Theódóra Torfadóttir and Tue Trinh for their help during the conference. I would like to thank Theódóra for providing essential editorial assistance.

Paul Law
Berlin, October 5, 2004

Errata

p. 135

Belleti -> Belletti

p. 144

de Sousa 2002 -> de Sousa 2001

p. 145

Belleti -> Belletti

p. 195   Nthelios -> Ntelitheos

p. 200, footnote 16 should read:

The problem is more serious on the GHT/Paul view, since movement of the Theme is an instance of A movement but must cross a theta position, which is in general not possible.   For Pearson the movement is an instance of A' movement, which applies unproblematically across a co-argument (*the man he saw*) [Dimitris Ntelitheos pc].

p. 201   Nthelios -> Ntelitheos
            The Acquisition of the Malagasy Voicing System ->
                The Acquisition of the Malagasy Voicing System

p. 189, above example (36)
            (34) -> (36)

p. 189, above example (37)
            to Bill -> to Max

Antipassive and Specificity in Tagalog

Edith Aldridge
SUNY, Stony Brook

1. Introduction

It is common knowledge in the field of Philippine linguistics that an *ang*-marked direct object in a non-actor focus clause must be definite or generic, while a *ng*-marked object in an actor focus clause typically receives a nonspecific interpretation.

(1) a. K-**in**-ain=ko   **ang**   **isda**.
       -Tr.Perf-eat=1sErg   Abs   fish
       'I ate the/*a fish.'

    b. K-**um**-ain=ako   **ng**   **isda**.
       -Intr.Perf-eat=1sAbs   Obl   fish
       'I ate (a) fish.'

However, in contexts like *wh*-questions, the oblique object in an antipassive may be interpreted as specific, as noted by Schachter & Otanes (1972), Maclachlan & Nakamura (1997), Rackowski (2002), and others.

(2)   Sino   ang   k-um-ain   **ng**   **isda**?
       who   Abs   -Intr.Perf-eat   Obl   fish
       'Who ate a/the fish.'

In this paper, I propose to account for the specificity effects seen in (1) and (2) within the analysis of Tagalog syntax put forth by Aldridge (2004). I analyze Tagalog as an ergative language, which accounts straightforwardly for the alternation in interpretation shown in (1). (1a) is a transitive clause, in which the internal argument has absolutive status. (1b), on the other hand, is an antipassive. Cross linguistically, antipassive oblique objects receive a nonspecific interpretation, while absolutives are definite or generic. I show in this paper how the Tagalog facts can be subsumed under a general account of ergativity.

The difference between declarative antipassives like (1b) and *wh*-questions like (2), is accounted for by the structural characteristics of this clause type. Aldridge (2004) argues that *wh*-questions of this type take the form of pseudoclefts. The possibility of a specific reading for the oblique object is due to the fact that this object is contained within the headless relative in matrix subject position, which is the presupposed part of the utterance. The object, then, can also receive a presupposed interpretation.

1

2. Tagalog Ergativity

This paper assumes an ergative analysis of Tagalog, as argued extensively in Aldridge (2004). Non-actor focus clauses are taken to be active and transitive. Locative and benefactive focus markers are treated as applicatives. I take actor focus clauses to be intransitive. Those containing a direct object are considered to be antipassives.

(3) Traditional Analysis   Ergative Analysis
     Non-actor focus (-in-)   Transitive
     Locative focus (-an)   Locative/dative applicative
     Benefactive focus (i-)   Benefactive applicative
     Actor focus (-um-/mag-)   Intransitive

Given this, Tagalog can be seen to display an ergative case-marking pattern. Absolutives are marked with *ang*, while *ng* marks ergative external arguments and non-absolutive themes and patients.

Non-actor Focus (Transitive)
(4) B-in-ili   **ng**   babae   **ang**   isda.
     -Tr.Perf-buy   Erg   woman   Abs   fish
     'The woman bought the fish.'

Actor Focus (Intransitive)
(5) D-um-ating   **ang**   babae.
     -Intr.Perf-arrive   Abs   woman
     'The woman arrived.'

(6) Actor Focus (Antipassive)
     K-um-ain   **ang**   babae   **ng**   isda.
     -Intr.Perf-eat Abs   woman   Obl   fish
     'The woman ate (a) fish.'

3. Information Structure in Antipassives

Several ergative analyses of Tagalog have been proposed in the literature (Payne 1982, De Guzman 1988, Aldridge 2004, and others). It is also fairly uncontroversial among Philippine linguists in general that non-actor focus clauses are active and transitive (Schachter 1976 and 1994, Kroeger 1993, Maclachlan & Nakamura 1997, among others). However, the question of whether Tagalog has an antipassive construction is still controversial. This section presents evidence that Tagalog actor focus clauses with a direct object are antipassives. In particular, I show that the object must receive an indefinite, nonspecific, or otherwise non-presuppositional interpretation, as is observed to be the case for antipassive constructions cross-linguistically (Cooreman 1994, Bittner 1994, Basilico 2003, and others).

Recall first the alternation observed in (1), repeated below as (7). The absolutive object in (7a) is definite, while the antipassive oblique in (7b) is indefinite and nonspecific.

(7) a. K-in-ain=ko   **ang**   **isda.**
        -Tr.Perf-eat=1sErg Abs   fish
        'I ate the/*a fish.'
    b. K-um-ain=ako   **ng**   **isda.**
        -Intr.Perf-eat=1sAbs Obl   fish
        'I ate (a) fish.'

This alternation is commonly found in ergative languages, as illustrated by the South Baffin Eskimo examples in (8). As in (7), the oblique object in the antipassive in (8a) is indefinite, while the absolutive object in (8b) is definite.

(8) a. Joosi   ququq-si-y-up-0   tuttu-mik
        Joosi.Abs   shoot-*si*-Part-Monop-3 caribou-MIK
        'Joosi shot a caribou.'
    b. Joosi-up   quqi-kkaniq-t-a-nga   tuttu
        Joosi-Erg   shoot-again-Part.Polyp-3/3 caribou.Abs
        'Joosi shot the same caribou again.'   (Kalmar 1979:124)

A survey of Tagalog antipassives found in a text confirms the tendency for oblique objects to be nonospecific. There are 65 antipassive clauses in 93-page long novela I examined. In 50 of these, the direct object is indefinite and nonspecific. (9) shows some typical examples of this type.

(9) a. Na-ka-kuha   **ng**   **scholarship**   ang   kaibigan.
        Perf.Pot-Intr-get   Obl   scholarship   Abs   friend
        '(Her) friend was able to get a scholarship.'
    b. Hindi ito   nag-karoon   **ng**   **injury.**
        Neg   3sAbs Perf.Intr-have Obl   injury
        'He didn't have an injury.'
    c. H-um-ugot   **ng**   **hininga**   si   Gilbert.
        -Perf.Intr-drawObl   breath   Abs   Gilbert
        'Gilbert drew a breath.'
    d. Hindi ito   nag-hintay   **ng**   **katugunan.**
        Neg   3sAbs Perf.Intr-wait Obl   answer
        'He didn't wait for an answer.'

In the remaining 15 examples, the object could be understood as definite.

(10)   a. Mag-bu-buslo   **ng**   **bola**   si   Gilbert.
           Intr-Red.Fut-shoot Obl   ball   Abs   Gilbert
           'Gilbert will shoot the ball.'

3

b.	Nagmamadaling	l-um-abas		**ng**	**silid**.
	quickly	-Perf.Intr-leave		Obl	room
	'(She) quickly left the room.'				

c.	Nag-taas	**ng**	**mukha**	si	Trini.
	Perf.Intr-lift	Obl	face	Abs	Trini
	'Trini lifted her face.'				

d.	Saan	ba	ako	puwede-ng	mag-palipas	**ng**	**gabi** rito?
	where	Q	1sAbs	may-Lk	Intr.spend	Obl	night here
	'Where can I spend the night here?'						

However, I suggest that the definiteness is these cases is not discoursally determined. In (10a), for instance, *bola* "ball" could be considered to be definite, but not because it has been introduced in the preceding discourse. Rather this is due to some type of pragmatic inference, due to the fact that the sentence is describing a scene from a basketball game, which always involves a ball (though not a specific ball). This type of definiteness has been as analyzed as bridging by Asher and Lascarides (1998) and need not be assumed to involve specificity. Thererfore, all of the 65 antipassive sentences examined above contained nonspecific objects and therefore pattern with antipassives in other languages.

4. Analysis of Specificity Effects in Declarative Clauses

In this paper, I propose an account of the above specificity effects, based on the ergative analysis of Tagalog syntax proposed by Aldridge (2004). The analysis takes as its theoretical foundation the theory of Multiple Spell-Out as proposed by Chomsky (2000, 2001a, 2001b). The status of vP as a phase and the Phase Impenatrability Condition (Chomsky2001b:5) play crucial roles in this account.

(11)   Phase Impenatrability Condition (PIC)
       The domain of a phase head is not accessible to operations, but only the edge is.

The PIC dictates that movement of VP-internal material must first pass through the edge of vP, i.e. the outer specifier. In the case of object *wh*-movement, for example, v must have an EPP (or occurrence) feature to first draw this DP into its outer specifier. From this position in the edge of vP, the object is accessible to the [wh] feature on C and can undergo further movement to [Spec, C]. Direct movement from within VP to [Spec, C] would violate the PIC.

(12)   What did you [$_{v}$P t$_{what}$ [$_{v'}$ t$_{you}$ [$_{v[EPP]}$ [VP eat t$_{what}$]]]]?

4.1. Ergative Languages
It is assumed for English that EPP features are generated on v when needed. What I propose for ergative languages is that the appearance of EPP features on v is restricted.
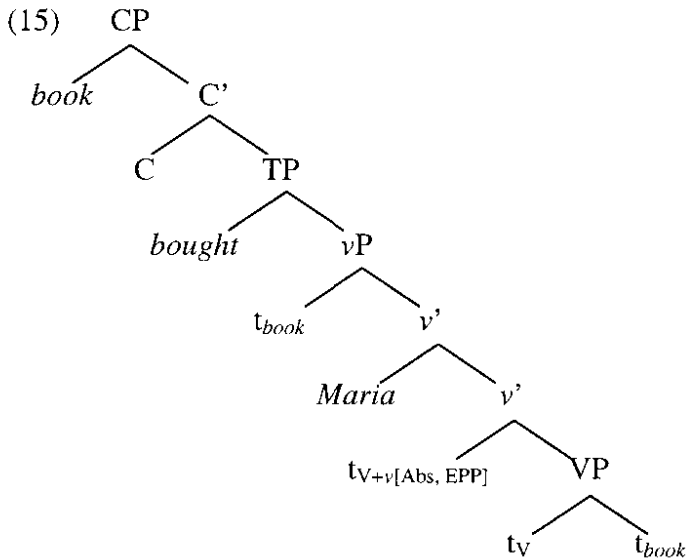
4

(13) a. Transitive *v* checks absolutive case and has an EPP feature, drawing the absolutive DP to its outer specifier.

  b. Intransitive *v* has no EPP feature; the direct object in an antipassive does not raise out of VP.

This accounts for the hallmark characteristic of syntactic ergativity, i.e. only absolutive DPs can undergo A'-movement. (cf S/O Pivot (Dixon 1979, 1994)) For example, a relative clause can be formed on a direct object in a transitive clause, as in (13a), but not on the oblique object in an antipassive.
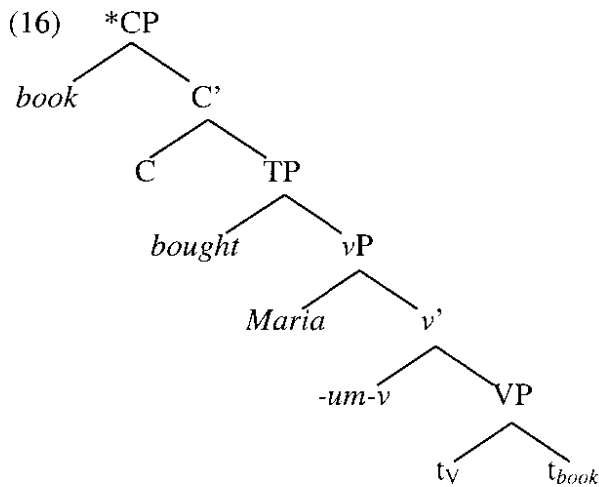
Relativization

(14) a. **libro**-ng b-**in**-ili   ni   Mara?
       book-Lk   -Tr.Perf-buy Erg Maria
       'book which Maria bought'

  b. ***libro**-ng b-**um**-ili   si   Maria?
       book-Lk   -Intr.Perf-buy Abs Maria
       'book which Maria bought'

In (14a), *v* is transitive and therefore has an EPP feature, which attracts the absolutive (the null operator coindexed with the head of the relative clause) to its outer specifier. From there, it can be further attracted to the specifier of CP.

(15) CP
```
         CP
        /  \
     book   C'
           /  \
          C    TP
        bought  \
               /  \
            vP
           /  \
       t_book  v'
             /   \
          Maria    v'
                 /    \
         t_V+v[Abs, EPP]  VP
                        /   \
                      t_V   t_book
```

Since antipassive verbs are intransitive, *v* in (14b) has no EPP feature. Direct extraction of the operator from object position within VP violates the PIC, thereby accounting for the ungrammaticality of (14b).

(16) *CP

```
        CP
       /  \
   book    C'
          /  \
         C    TP
      bought  / \
            vP
           /  \
        Maria  v'
              /  \
          -um-v   VP
                 /  \
               t_V   t_book
```
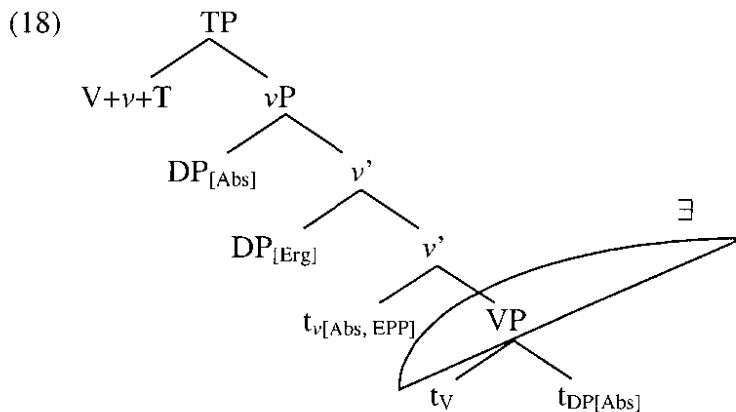
The difference in interpretation between absolutive and oblique direct objects is also accounted for the presence or absence of an EPP feature on v.
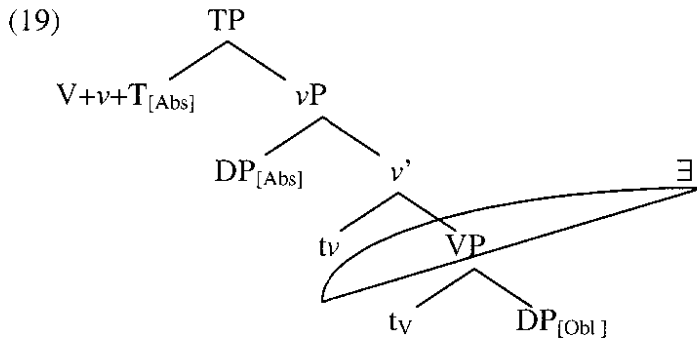
(17) a. B-**in**-ili   ng   babae   ang   isda.
          -Tr.Past-buy   Erg   woman   Abs   fish
          'The woman bought the/*a fish.'

      b. B-**um**-ili   ang   babae   ng   isda.
          -Intr.Perf-buy   Abs   woman   Obl   fish
          'The woman bought a/*the fish.'

According to Diesing's (1992) Mapping Hypothesis, a clause is divided into two parts at LF. Material inside VP is mapped to the nuclear scope, where it undergoes existential closure and receives a non-quantificational, existential interpretation. Material outside VP, on the other hand, is mapped to the restrictive clause and receives a quantificational or presuppositional reading.

In ergative languages, I have proposed that absolutive DPs raise out of VP to the outer specifier of vP, with the result that they will be mapped to the restrictive clause at LF. Therefore, absolutives must receive a presuppositional interpretation.

(18)

```
              TP
             /  \
        V+v+T    vP
                /  \
           DP_[Abs]  v'
                    /  \
               DP_[Erg]  v'
                        /  \
              t_v[Abs, EPP]  VP
                            /  \
                          t_V   t_DP[Abs]
```

∃

The oblique object in an antipassive, on the other hand, remains inside VP, since v is intransitive in antipassives and does not have an EPP feature. Consequently, the object will be mapped to the nuclear scope and undergo existential closure at LF to receive a nonpresuppositional reading.

(19)

$$
\begin{array}{c}
\text{TP} \\
\diagup \diagdown \\
\text{V}+v+\text{T}_{[\text{Abs}]} \qquad v\text{P} \\
\diagup \diagdown \\
\text{DP}_{[\text{Abs}]} \qquad v' \\
\end{array}
$$

V+v+T[Abs]   vP

DP[Abs]   v'   ∃

tv   VP

tv   DP[Obl]

The proposal presented in this section accounts for the asymmetry between declarative transitive and antipassive clauses. I will return to the question of pseudoclefts in section 5.

4.2. Objections?

Though most Austronesian linguistics accept the fact that non-actor focus clauses in Philippine languages are active and transitive and therefore possibly ergative, many continue to reject the ergative analysis, claiming that these languages do not have a true antipassive construction (Shibatani 1988, Kroeger 1993, Schachter 1994, Paul & Travis 2003).

For example, Kroeger (1993) objects to the analysis of Tagalog actor focus as antipassive, because the object can control a gap in a participial clause. He claims that this shows that the object in question is a term, i.e. argument, and not an oblique.

(20)   Nanghuli   **ng**   **magnanakaw**$_i$ ang   polis$_j$
        AV.Perf.catch Gen   thief   Nom   police
            [nang   PRO$_{i/j}$ pumapasok   sa   bangko]
            Adv   AV.Imperf.enter   Dat   bank
        'The police caught a/the thief when entering the bank.'

                                            (Kroeger 1993:47)

Kroeger substantiates his distinction between terms and obliques by showing that the latter can undergo focus movement, while the former cannot.

(21)   a.   Mag-bi-bigay=akong   bulaklak   kay   Maria.
            Intr.Perf-Red-give Obl   flower   to   Maria.
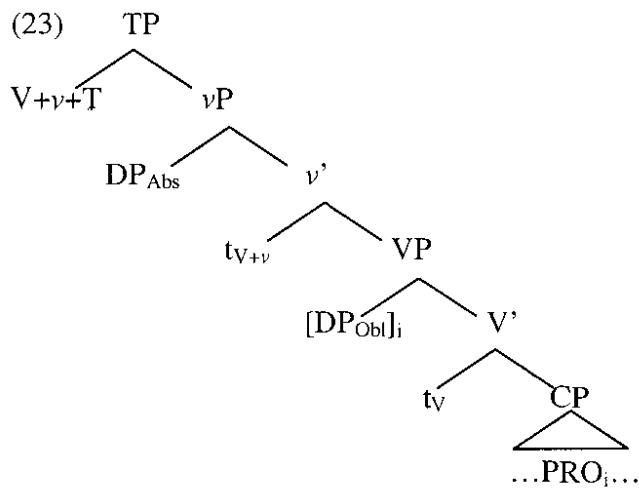   'I will give flowers to Maria.'

b. **Kay**   **Maria**=ako   mag-bi-bigay   ng   bulaklak.
   to   Maria=1sAbs   Intr.Perf-Red-give   Obl   flower
   'I will give flowers *to Maria*.'
c. ***Ng**   **bulaklak**=ako mag-bi-bigay   kay   Maria.
   Obl   flower=1sAbs Intr.Perf-Red-give   to   Maria
   'I will give *flowers* to Maria.'

Kroeger's distinction is actually between DPs and non-DPs, rather than between arguments and adjuncts. In terms of control, Kroeger's analysis breaks down when confronted with examples of the following type. In (22), it is clear that PPs (obliques on Kroeger's analysis) can be controllers.

(22)   a.   Nag-utos   ang   nanay sa   [anak=niya]$_i$-ng
            Intr.Perf-order   Abs   mother Dat   child=3sGen-Lk
               [PRO$_i$ mag-bantay   ng   bahay]
                       Intr-watch   Obl   house
            'The mother ordered her child to watch the house.'
       b.   Na-ka-tingin   si   Maria$_i$ sa   [kaibigan=niya]$_j$
            Perf-Intr-see   Abs   Maria Dat   friend=3sGen
               [nang   PRO$_{i/j}$ p-um-apasok   sa   bangko]
               C   -Intr.Perf-enterto   bank
            'Maria saw her friend entering the bank.'

The difference between PPs and oblique DPs is therefore one of category. Assuming that an EPP feature is realized as a strong [D] feature, the lack of such a feature on *v* is what prevents oblique DPs from fronting. Movement of PPs should not be relevant to the presence or absence of an EPP feature.

As to the question of how it is that oblique objects can be controllers, under my analysis of ergativity, this is not at all surprising. What distinguishes antipassive from transitive clauses is whether *v* checks absolutive case and has an EPP feature. In both cases, the verb still has an internal theta-role to assign, i.e. the object in both clause types is merged in argument position inside VP and is therefore in a position to c-command inside a complement clause.

8

(23)   TP
       /\
  V+v+T   vP
        /\
    DP_Abs   v'
           /\
        t_V+v   VP
              /\
          [DP_Obl]_i   V'
                    /\
                  t_V   CP
                       /\
                      ...PRO_i...

5. Specificity of Embedded Obliques

As we have seen in section 1, oblique objects in antipassives can receive a specific interpretation in *wh*-questions.

(24)   Sino   ang   k-um-ain   ng   isda?
        who   Abs   -Intr.Perf-eat Obl   fish
        'Who ate a/the fish.'

The analysis I proposed in section 4 does not predict this possibility, since intransitive *v* has no EPP feature and the oblique object is not forced to raise out of VP. Therefore, it should be in the nuclear scope at LF.

However, it is generally accepted that, in a great many Austronesian languages, *wh*-questions formed on DPs take the form of pseudoclefts (Aldridge 2002a, 2002b; Paul 2000; and many others). This fact suggests a natural account for (24). The direct object is contained inside the headless relative which provides the presupposition of the clause and is located in the matrix subject position, outside of the domain of existential closure. As part of the presupposition, then, the embedded direct object can also receive a specific interpretation.
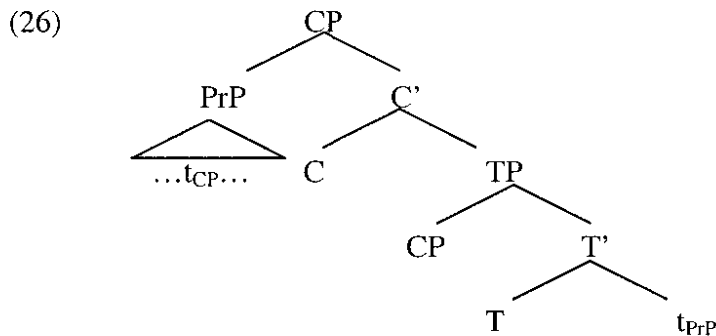
5.1. Wh-questions as Pseudoclefts

In Aldridge (2004), I propose the following structure for Tagalog *wh*-questions formed on pseudoclefts. The *wh*-phrase forms the matrix predicate. The rest of the clause is contained in a headless relative, which is located in the matrix subject position.

(25)   a.   Sino   ang   k-um-ain   ng   isda?
              who   Abs   -Intr.Perf-eat Obl   fish
              'Who ate a/the fish.'

9

b.

```
            CP
           /  \
              C'
             /  \
            C    TP
                /  \
                   T'
                  /  \
                 T    PrP
                     /      \
                   DP        Pr'
                  /  \      /   \
                ang   CP   Pr    NP
                     /  \         |
                   Op   TP       who
                        /\
                   ate fish t_Op
```
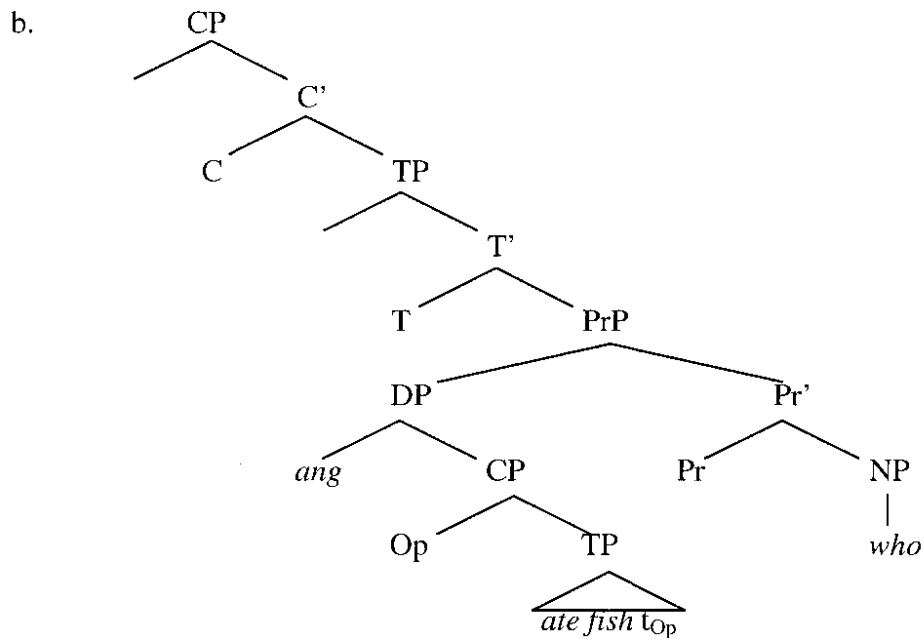
In order to obtain predicate-initial word order, the headless relative raises to matrix subject position, and PrP fronts to its left.

(26)

```
              CP
             /  \
          PrP    C'
          /\    /  \
     ...t_CP... C    TP
                    /  \
                  CP    T'
                       /  \
                      T    t_PrP
```

What is important for the discussion at hand is how the interpretation is derived. The headless relative, located in the [Spec, T] subject position, is mapped to the matrix restrictive clause. The embedded object, which is contained within the subject phrase, may also receive a definite or specific reading by virtue of the fact that it is part of the presupposition of the clause.

## 5.2. Relative Clauses in Argument Position

This account is further supported by examples in which an antipassive appears inside a relative clause in argument position. (27a) shows that an oblique object contained in a relative clause which itself has absolutive status in the matrix clause, may receive a specific or definite interpretation. On the other hand, if the relative clause is itself an oblique in an antipassive matrix clause, as in (27b), then the embedded object must be interpreted as nonspecific.

10

(27) a. B-in-ili=ko
          -Tr.Perf-buy=1sErg

	ang	[pusa-ng	k-um-ain	ng	daga].
	Abs	cat-Lk	-Intr.Perf-eat	Obl	rat

          'I bought the cat which ate a/the rat.'

      b. B-um-ili=ako
          -Intr.Perf-buy=1sAbs

	ng	[pusa-ng	k-um-ain	ng	daga].
	Obl	cat-Lk	-Intr.Perf-eat	Obl	rat

          'I bought a cat which ate a/*the rat.'

This is accounted for in the following way. If the relative clause has absolutive status, it moves to the outer specifier of *v* in the matrix clause. The clause itself is mapped to the presupposition at LF, and material contained inside it may also be presupposed, yielding the possibility of a specific interpretation for the embedded object.

(28)
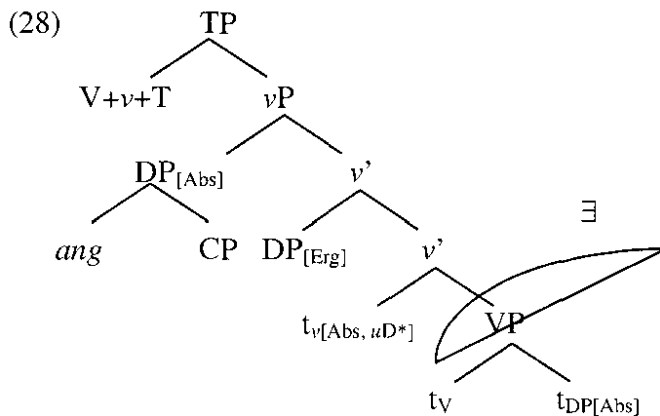
On the other hand, if the relative clause has oblique status in the matrix clause, then it will not move out of VP and will underdo Existential Closure at LF. Material inside the clause will also be interpreted as nonspecific.

(29)   TP

V+*v*+T_{[Abs]}   *v*P

DP_{[Abs]}   *v*'   ∃

t*v*   VP

t_V   DP_{[Obl]}

ng   CP
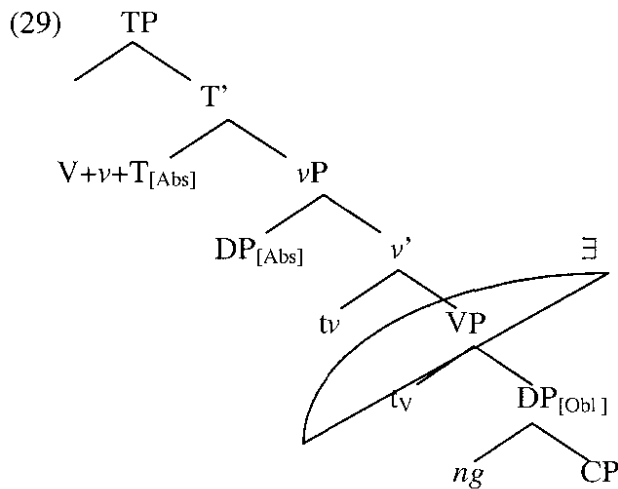
The key point of the analysis here is that the possibility of a specific interpretation for an embedded oblique object is dependent on mapping in the matrix clause. For lack of space, I am unable to introduce the analyses themselves, but there are proposals by Rackowski (2002) and Maclachlan and Nakamura (1997) which tie the specific interpretation of the oblique object to extraction of the external argument. However, the above examples clearly show that the specificity of the embedded oblique object cannot be tied to extraction of the agent. This is because the agent is extracted in the relatives clauses in both (27a) and (27b), but only the object in (27a) can receive a specific interpretation. Aldridge (to appear) presents Rackowski's (2002) analysis and arguments against this approach.

6. Conclusion

This paper has proposed an analysis of two types of specificity effect in Tagalog. As an ergative language, absolutives in Tagalog always receive a definite or generic interpretation, while oblique objects in antipassives are typically nonspecific. This alternation is accounted for by restricting the appearance of an EPP feature on *v* to transitive contexts only. The effect of this is that *v*'s EPP feature attracts the absolutive DP to its outer specifier, where they will be mapped to the restrictive clause at LF and receive a generic or specific interpretation. Intransitive *v*, on the other hand, does not have an EPP feature. Oblique objects in antipassives typically remain inside VP and undergo Existential Closure at LF.

As we have seen, however, oblique objects may be specific when contained in a relative clause. However, the possibility of a specific interpretation for the embedded object is again dependent on the Mapping to LF in the matrix clause. When the containing relative is mapped to the restrictive clause, the embedded object is also allowed a specific interpretation. However, if the containing relative is mapped to the matrix nuclear scope, then the embedded object will also receive an existential, non-presuppositional interpretation.

12

References

Asher, Nicholas. and Alex. Lascarides. 1998. Bridging. Journal of Semantics. 15, 1.83-113.

Aldridge, Edith. 2002a. Nominalization and Wh-movement in Seediq and Tagalog. Language and Linguistics 3, 2.393-426.

Aldridge, Edith. 2002b. Wh-movement in Seediq and Tagalog. Proceedings of AFLA 8: the Eighth Meeting of the Austronesian Formal Linguistics Association, ed. by A. Rackowski & N. Richards, 1-15. MIT Working Papers in Linguistics 44.

Aldridge, Edith. 2004. Ergativity and Word Order in Austronesian Languages. Ph.D. dissertation, Cornell University.

Aldridge, Edith. to appear. Antipassive, Clefting, and Specificity. Pennsylvania Working Papers in Linguistics 11.1: Proceedings of the 28th Annual PLC.

Basilico, David. 2003. The Antipassive: Low Transitivity Objects, Syntactic Intertness, and Strong Phases. Paper presented at the North American Syntax Conference. Montreal: Concordia University.

Bittner, Maria. 1994. Case, Scope, and Binding. Dordrecht: Kluwer Academic Publishers.

Bos◊kovic≅, Z◊eljko. 1997. Pseudoclefts. Studia Linguistica 51, 3.235-277.

Chomsky, Noam. 2000. Minimalist Inquiries. Step by Step: Essays in Minimalist Syntax in Honor of Howard Lasnik, ed. by R. Martin et al., 89-155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001a. Derivation by Phase. Ken Hale: A Life in Language, ed. by M. Kenstowicz, 1-52. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001b. Beyond Explanatory Adequacy. MIT Occasional Papers in Linguistics 20.

Collins, Peter. 1991. Pseudocleft and Cleft Constructions: A Thematic and Informational Interpretation. Linguistics 29.481-519.

Cooreman, Ann. 1994. A Functional Typology of Antipassive. Voice: Form and Function, ed. by B. Fox and P. Hopper, 49-86. Amsterdam: John Benjamins

De Guzman, Videa P. 1988. Ergative Analysis for Philippine Languages: An Analysis. Studies in Austronesian Linguistics, ed. by R. McGinn, 323-345. Athens, Ohio: Ohio University Center for International Studies

Diesing, Molly. 1992. Indefinites. Cambridge, MA: MIT Press.

Dixon, R.M.W. 1979. Ergativity. Language 55.59-138

Dixon, R.M.W. 1994. Ergativity. Cambridge: Cambridge University Press.

Gerdts, Donna B. 1988. Antipassives and Causatives in Ilokana: Evidence for an Ergative Analysis. Studies in Austronesian Linguistics, ed. by R. McGinn, 295-321. Athens, Ohio: Ohio University Center for International Studies.

Kalmar, I. 1979. The Antipassive and Grammatical Relations in Eskimo. Ergativity, ed. by Franz Plank, 117-143. New York: Academic Press.

Knowles, John. 1986. The Cleft Sentence. Lingua 69.295-317.

Kroeger, Paul. 1993. Phrase Structure and Grammatical Relations in Tagalog. Stanford: Center for the Study of Language and Information.

Maclachlan, Anna and Masanori Nakamura. 1997. Case-checking and Specificity in Tagalog. The Linguistic Review 14.307-333.

Paul, Ileana. 2000. Malagasy Clause Structure. Ph.D. dissertation, McGill University.

Paul, Ileana and Lisa Travis. 2003. Ergativity in Austronesian Languages: What it can do, what it can't, but not why. Paper presented at the 10th meeting of the Austronesian Formal Linguistic Association (AFLA), University of Hawaii.

Payne, Thomas E. 1982. Role and Reference Related Subject Properties and Ergativity in Yup'ip Eskimo and Tagalog. Studies in Language 6, 1.75-106.

Prince, Ellen. 1978. A Comparison of Wh-clefts and It-clefts in Discourse. Language 54, 4.883-906.

Rackowski, Andrea. 2002. The Structure of Tagalog: Specificity, Voice, and the Distribution of Arguments. Ph.D. dissertation, MIT.

Schachter, Paul. 1976. The Subject in Philippine Languages: Topic, actor, actor-topic, or none of the above. Subject and Topic, ed. by Charles Li, 491-518. New York: Academic Press.

Schachter, Paul. 1994. The Subject in Tagalog: Still None of the Above. UCLA Occasional Papers in Linguistics 15.

Schachter, P. and Fe T. Otanes. 1972. Tagalog Reference Grammar. Berkeley, CA: University of California Press.

Shibatani, Masayoshi. 1988. Voice in Philippine Languages. Passive and Voice, ed. by M. Shibatani, 85-142. Amsterdam: John Benjamins.

Primary Source

Caliente, Kayla. *Ngayong Mahal na Kita.* Manila: Bookware Publishing Corporation.

Towards a Typology of Austronesian Pronominal Clisis

Loren Billings   &   Daniel Kaufman
National Chi Nan University   Cornell University

This paper reports ongoing work on the following aspects of Austronesian clisis: the relative ordering of multiple clitics, the position of the clitic cluster.

Clitic phenomena have captured a great deal of attention because they constitute a primary testing ground for the various conceptions of modularity in the architecture of Universal Grammar. Specifically, the linearization of clitics in languages such as South Slavic (Rudin et al. 1999) and Chamorro (Chung 2003) requires access to prosodic phrasing. Clitics in languages such as these therefore represent the most clear-cut cases of movement in PF. However, unlike other such cases, morphosemantic features are still required at this level to determine the relative order of multiple clitics within the cluster. Capturing precisely this interplay between prosody and morphosyntax has been one of the major challenges in work on the subject. The variation among Austronesian languages as to cluster-internal ordering is shown to represent different compromises between morphosemantic and prosodic requirements.

An area that has not received much attention at all in the theoretical Austronesianist literature is the morphosyntax of verb-adjacent clisis. We discuss in this regard several languages of Central Sulawesi, focusing on pre- and post-verbal alternation. A paradigm-based Optimality-theoretic analysis is offered to account for the distribution of pre-verbal positioning in these languages. In addition, we offer observations concerning portmanteaux, co-occurrence restrictions, clitic doubling, and non-pronominal clitics.

An area which has been explored less in the literature is the boundary between verb-adjacent clisis and person agreement. This is, in part, a problem of definitions, as there are still residual problems with the difference between affix- and clitic-hood (despite the good progress in this direction). We therefore seek a robust distinction between verb-adjacent clitics and verbal-agreement affixes based on Austronesian data. In this regard, the variation among the languages of Sulawesi, Indonesia, appears to be highly promising.

1. Internal ordering
Like normal syntactic constituents, clitic pronouns can be ordered based on morphosemantic factors. These include morphological case, semantic role, or φ-features. Unlike normal syntax, their order can also be based on prosody.

1.1 Orderings based on morphological case or semantic roles
It is often difficult to determine whether it is morphological case or semantic roles which determines order. This is because in many Austronesian languages, multiple pronominal clitics appear in only a limited number of voice configurations because pronominal (more generally, definite) patients cannot be realized in the GEN case. Generally, in non-Actor voice constructions, the Actor is realized in the GEN. As such, it is impossible to tease case apart from roles. In at least one language that we have come across, however, it is possible to show that the semantic role and not case is the determining factor.

In a number of languages, a GEN (Actor) precedes a NOM (non-Actor) pronoun. The Central Luzon subgroup (Kapampangan, Sambal, Bolinao, and Aita) is categorical in this respect. In Central Philippines, Mamanwa is likewise categorical, but other languages have a mere preference as long as more highly ranked prosodic constraints are satisfied (e.g., Tagalog and Bikol).

Other languages position the NOM (non-Actor) in front of the GEN (Actor). Although we know of no languages that require this ordering, some of them appear to favor this order if other factors (such as prosody or person) are controlled for: Agusan Manobo (Weaver and Weaver 1964), Sarangani Manobo (Dubois 1976: 48–50), and Mayrinax Atayal (Huang 1995: 28–36).

We know of no language where case and crucially not semantic roles determines the order of two *clitic* pronouns. If such a language exists, it would not be a problem, because the pronouns would be ordered using normal syntax.

The only language of this kind we have found in which case and roles can be teased apart is Isbukun Bunun, a language that seems to order the Actor before the other pronoun regardless of case (NOM vs. OBL) or the verb's voice:

(1) a. Masaiv -*ik* -*su* tasa ahil. (Isbukun Bunun)
     give.AV 1SG.NOM 2SG.OBL one book
     'I gave you a book.' [Huang et al. 1999: 186]
   b. Tahuan -*ku* -*as* bunun tu halinga. (Isbukun Bunun)
     teach.NAV 1SG.OBL 2SG.NOM Bunun LNK word
     'I taught you the Bunun language.' [Huang et al. 1999: 187]

Though (1a–b) suggest that grammatical person may be the factor, Huang et al. (1999: 188) add that person (as well as case or number) do not affect ordering.

It is not clear how these Isbukun Bunun data can be handled by prevailing theories, where the syntax tends to have access to a nominal expression's case but not its semantic role as such. Assuming that the syntax precedes the morphological component derivationally, it would be even harder for a morphological theory of clitic ordering to have access to the pronouns' semantic roles. Clearly, this phenomenon deserves further investigation.

1.2 Orderings based on φ-features

Unlike case or semantic role, φ-features do not change to reflect the predicate's voice. The features of this kind relevant to Austronesian are person and number. (The languages we are aware of do not exhibit pronominal genders.)

Quite a number of languages utilize person in ordering clitic pronouns. It is interesting to note in this regard that all three descriptions of Manobo and Atayal languages mentioned above, in which a NOM (non-Actor) precedes a GEN (Actor), point out explicitly an overriding ordering constraint requiring a third-person clitic pronoun to follow a first- or second-person clitic pronoun.

At least two accounts of the Squliq dialect of Atayal argue additionally for person to be used to order first- and second-person pronouns relative to each other. Using different combinations of such pronouns, Rau (1992: 146–147) argues for a first-second order, whereas Huang (1995: 34–35) proposes the opposite order. Liao (2004) shows convincingly that prosody rather than person decides the ordering of such clusters. Indeed, if person is used as a factor, the only conclusive examples in our view are of a third-person clitic pronoun having to follow either a first- or a second-person clitic pronoun.

Nor does number appear to be a clear factor in ordering clitic pronouns. Although Huang (1995: 28–36) proposes number as a factor in ordering clitic pronouns for Mayrinax Atayal, prosody is just as valid an explanation.

To summarize, of all the φ-feature factors, the only one that seems to matter is person (and only third person as opposed to first or second person). This factor is also noteworthy in that it frequently interacts with both NOM-before-GEN and prosody-based systems. Number does not seem to be crucial.

1.3 Orderings based on prosody

In several Austronesian languages, linguists have argued that number of syllables is the factor determining the order of two clitic pronouns. In each of these languages, if prosody is a factor, the order is light before heavy.

As has been widely publicized for Tagalog (e.g., Schachter 1973), monosyllabic clitic pronouns precede disyllabic ones. Other languages—all from Central Philippine—such as Bikol and Cebuano (Billings and Konopasky 2002) and Tagakaulo, Kaagan, Mansaka, and Davawenyo (Lee 2004), use the number of syllables to order clitic pronouns within the cluster. As mentioned above, Liao (2004) makes the same argument about Squliq Atayal, and these ideas could be extrapolated to Huang's (1995) Mayrinax Atayal data as well.

In some of these languages the prosodic constraint is categorical; only if the pronouns have the same number of syllables do morphosemantic factors emerge. Tagalog, Bikol, Mansaka, and Davawenyo are languages of this kind. All the aforementioned accounts of Atayal varieties—Rau (1992), Huang (1995), and Liao (2004)—argue for a morphosematic constraint requiring a third-person clitic pronoun to follow first- or second-person one. Liao further argues that the prosodic factor emerges in clusters of first- and second-person clitic pronouns. Tagakaulo, heavily influenced by Sarangani Manobo, also shows a strong preference to position third-person forms last.

1.4 Generalizations about cluster-internal ordering

To summarize the factors that influence the order of two clitic pronouns in a cluster, several types emerge. In each of these, there is just one possible order.

First are the languages that order a GEN-case Actor before a NOM-case non-Actor: the Central Luzon group and Mamanwa. Isbukun Bunun is similar in ordering an Actor before the other role, regardless of voice/case. We assume as a hypothesis that the crucial ordering factor in the other languages as well is the semantic role and not case. Our main reason for developing this notion is that each of the factors—role, case, person, and even prosody—have a single order; with regard to semantic roles, it is the Actor that precedes any other role.

Next are the languages with the opposite order: NOM-case non-Actor preceding GEN-case Actor. On strictly conceptual grounds, we assume that case is the relevant factor here, with the subject preceding non-subject cases. Our hypothesis predicts that if case is the factor, then NOM is leftmost.

Similarly, it is quite common cross-linguistically for third-person pronouns to follow any others in the cluster; see, for example Rudin et al. (1999) regarding Slavic. As with the preceding two factors, if person is the relevant factor, then a third-person pronoun is predicted to follow any other.

Finally, if prosody is the relevant factor, then our hypothesis predicts that a lighter pronoun precedes a heavier one, not vice versa. Billings and Konopasky (2003) offer a rationale for this ordering; only a monosyllabic clitic

can adjoin prosodically to a preceding word in Tagalog. It remains to be seen how this rationale could be extended to languages (e.g., Tagakaulo) in which trisyllabic clitic pronouns behave differently from disyllabic ones.

So far, we have compiled four necessarily pre-theoretic constraints on cluster-internal ordering. These represent a hypothesis in and of themselves. We leave it to future work to determine if the languages under discussion can continue to be accounted for by alternative rankings of these constraints.

2. External ordering

As the preceding discussion of particles has shown, there are two basic positions for the clitics relative to the rest of the clause. We give only a brief overview of Wackernagel systems in order to concentrate on the verb-adjacent systems, which have attracted less attention in the Austronesianist literature.

2.1 Wackernagel systems

The vast majority of Philippine languages instantiate Wackernagel clisis: following some initial element. Several restrictions on clitic movement are common. Initial topics do not constitute the initial element, in (2a), and initial OBL-case foci generally behave as impermeable constituents for clitics, in (2b):

(2) a. [Buhat (*siya) sa= Maynila =ay] (*siya) mag-la-lakad siya.   (Tagalog)
        from   OBL Manila   TOP   AV-IRR-walk 3SG.NOM
        "He will walk [_Topic_ from Manila]."

   b. [Buhat (*siya) sa= Maynila] siya   mag-la-lakad.   (Tagalog)
        from   OBL Manila   3SG.NOM AV-IRR-walk
        "He will walk [_Focus_ from Manila]."

   c. Buhat siya   sa= Maynila.   (Tagalog)
        from 3SG.NOM OBL Manila
        "She/He is from Manila." [(b–c) from Schachter and Otanes 1972: 190]

Another common feature is that the Case Phrase (KP) represents an island for clitic movement. Exemplifying again with Tagalog, clitics can never precede the _ang_ 'NOM', _ng_ 'GEN' or _sa_ 'OBL' phrases they are associated with.

(3) a. Hindi (*ko)   iyon [_KP_ ang= problema ko].   (Tagalog)
        NEG   1SG.GEN that   NOM problem 1SG.GEN
        'That's not my problem.' (= 'That's not [the problem I have].')

   b. Hindi ko   iyon problema.   (Tagalog)
        NEG   1SG.GEN that problem
        'That's not my problem.' (= 'That's not a problem to me.')

Clitic climbing from a lower clause raise to a higher one, is also rarely attested in Philippine languages; (4a–b) are typical examples. However, modals and serial verbs generally do not block clitic movement, as (5a–b) both show.

(4) a. Hindi ko   (*siya)   s<in>abi [na   taksil siya].   (Tagalog)
        NEG   1SG.GEN 3SG.NOM PV.PRF-say COMP traitor 3SG.NOM
        'I didn't say he's a traitor.'

b. Achi -*na*   aforot-on   [ay   fiakon -*ak*   om-ali]. (E. Bontoc)
NEG -3SG.GEN believe-PV COMP NEG   -1SG.NOM AV-come
'He does not believe that I am not coming.'   [Fukuda 1997: 44]

(5) a. Hindi *ka*   dapat   (*ka)   um-alis   (*ka).   (Tagalog)
NEG 2SG.NOM should   2SG.NOM AV-leave   2SG.NOM
'You shouldn't leave.'
b. Kaman-*ak*   om-oy ay   ma-rqos.   (E. Bontoc)
seems   1SG.NOM AV-go LNK STA-pass
'Seems like I go passing by.'   [Fukuda 1997: 48]

What qualifies as the initial element is subject to minor variation. In a number of languages, the constituent relevant for hosting Wackernagel clitics is the PrWd, as exemplified by the following sentences, from Tagalog.

(6) a. L\<um\>utas   *na ako*   ng= isa -ng   suliranin.   (Tagalog)
AV.PRF-solve PRF 1SG.NOM GEN one   LNK problem
'I already solved one problem.'
b. Hindi *ka*   *ba* l\<um\>utas   ng= bigla -ng   suliranin? (Tagalog)
NEG 2SG.NOM Q AV.PRF-solve GEN sudden LNK problem
'Haven't you ever solved a sudden problem?'

The NOM-case pronominal set may also appear in clause-initial position in cleft-like constructions and topicalizations. However, for the 2SG pronoun (whose clitic version *ka* is the only monosyllabic form in the NOM paradigm), there exists a non-clitic disyllabic form *ikaw*. Compare examples (7a–b):

(7) a. *Ako*   ang= guro.   (Tagalog)
1SG.NOM NOM teacher
'I am the teacher.'
b. *Ikaw*   ang= estudyante.   (Tagalog)
2SG.NOM NOM student
'You are the student.'   (cf. *Ka ang estudyante.)

The 2SG form is thus the only formal diagnostic to distinguish between the nearly identical clitic and nonclitic paradigms of the NOM personal pronoun.

Languages differ with regard to what constituent can serve as the initial element. All the languages of the Philippines that we have encountered treat negation, fronted adverbs, and fronted foci as a potential first element:

(8) a. Hindi   *ka*   mag-lu-luto   ng= itlog.   (Tagalog)
NEG   2SG.NOM AV-IRR-cook GEN egg
'You won't cook (the) eggs.'
b. Ya' *ku*   pilay pa luma' bi   batna'a.   (Mapun)
NEG 1SG.NOM go   DIR house 2SG.GEN now
'I'm not going to your house now.'   [Collins et al. 2001: 591]

Languages differ as to whether complementizers may serve this purpose. As seen from (9a–b), Tagalog does not allow this but Mapun (Sama Bajaw) does:

19

(9) a. Kung hindi *ka*   mag-lu-luto   ng= itlog, ...   (Tagalog)
     if   NEG 2SG.NOM AV-IRR-cook GEN egg
     'If you didn't cook (the) eggs, ...'
   b. Bong *ko*   ya' pa-tagong nangis, ...   (Mapun)
     if   2SG.NOM NEG AV-stop   AV.cry
     'If you don't stop crying, ...'   [Collins et al. 2001: 115]

2.2 Verb-adjacent systems

Moving southwards from the Philippines, we find a major change in clitic placement upon reaching Central Sulawesi. This difference accompanies a distinct kind of verbal paradigm for the Actor voice/Patient voice distinction. Because this represents an important divergence in the linearization of clitics, we will analyze clitic behavior in Kulawi, a representative Kaili-Pamona language, and offer an Optimality-theoretic account of the attested variation. (The case labels in the rest of section 2 are synchronically better analyzed as ERG and ABS. The more neutral labels of GEN and NOM, resp., are used in order to maintain the historical linkage between the paradigms across Austronesian.)

Like most Kaili-Pamona languages (Mead 2002), Kulawi possesses only two diatheses: Actor voice and Patient voice. There are only two moods: realis and irrealis, indicated in the actor voice by *nang-* and *mang-*, respectively. The patient voice, by contrast, is marked overtly only in the realis, by the *i-* prefix (cognate to the *-in-* affix commonly found in the Philippines and Taiwan). The irrealis aspect of the patient voice is marked by the pre-verbal attachment of a pronominal Actor or by *ra-* (3PL.GEN) if the Actor is suppressed. The relevant paradigms can be seen in (10).

(10) Kulawi verbal paradigms (partial)   [Adriani and Esser (1939)]

	φ-features	Realis	Irrealis
Actor voice (intransitive)	1SG	*nang*-STEM-*a*	*mang*-STEM-*a*
	2SG	*nang*-STEM-*ko*	*mang*-STEM-*ko*
	3SG	*nang*-STEM-*i*	*mang*-STEM-*i*
	1PL.IN	*nang*-STEM-*ta*	*mang*-STEM-*ta*
	1PL.EX	*nang*-STEM-*kami*	*mang*-STEM-*kami*
	2PL	*nang*-STEM-*komi*	*mang*-STEM-*komi*
	3PL	*nang*-STEM-*ra*	*mang*-STEM-*ra*
Patient voice (transitive)	1SG	*i*-STEM-*ku*	*ku*-STEM
	2SG	*i*-STEM-*mu*	*mu/nu*-STEM
	3SG	*i*-STEM-*na*	*na*-STEM
	1PL.IN	*i*-STEM-*ta*	*ta*-STEM
	1PL.EX	*i*-STEM-*kami*	*ki*-STEM
	2PL	*i*-STEM-*mi*	*mi*-STEM
	3PL	*i*-STEM-*ra*	*ra*-STEM

In the Patient-voice realis, we see that the *i-* prefix marks the mood while a pronominal Actor is expressed as a verb-adjacent GEN enclitic. The corresponding irrealis, on the other hand, is signalled solely through the proclitic attachment of the pronominal. What is generally found in the Kaili-Pamona subgroup is that the NOM pronominals are expressed as either free

pronouns or bound forms. (Free pronouns may be ordered more or less freely within the sentence, whereas the bound sets appear to be verb-adjacent.)

Unlike the Wackernagel positioning observed above in the Philippine languages, the GEN pronominals with verbs prefixed by (n)i- do not encliticize to negation or any preverbal material in Kulawi and related languages. (In the Da'a examples, it's not clear whether ni should be analyzed as a PV prefix.)

(11)   Position of the GEN pronominals: invariably after the verb
a.   Moma i-   epe -ku   ka= rata -mu.   (Kulawi)
     NEG   PV.REAL hear 1SG.GEN NOM come 2SG.GEN
     'I haven't heard about your arrival.'   [Adriani and Esser 1939: 30]
b.   Da'a ninjani -ku.   (Da'a)
     NEG know   1SG.GEN
     'I don't know.'   [Abas and Friberg 1989: 21]
c.   Wei-wei ninjani -ku   -mo.   (Da'a)
     now   know 1SG.GEN already
     'Now I already know (it).'   [Abas and Friberg 1989: 21]

By contrast, NOM clitics may encliticize to preverbal elements such as negation, in (12a), and adverbials, (12b), as well as certain post-verbal elements, (12c):

(12)   Position of the NOM pronominals: Wackernagel clitics
a.   Moma -'a   t<um>ai.   (Kulawi)
     NEG   1SG.NOM AV-come.here
     'I didn't come here.'
b.   Iwengi   -'a   no-   towa kau i=   bone.   (Kulawi)
     yesterday 1SG.NOM AV.REAL chop tree OBL field
     'Yesterday, I chopped down a tree in the field.'
c.   Ba i-   tudu tina   -mu   -da   -ko?   (Kulawi)
     Q   PV.REAL send mother-2s.GEN -EMPH -2SG.NOM
     'Are you sent by your mother?'   [all from Adriani and Esser 1939: 30]

Although a full treatment of NOM clitics cannot be given here, it appears that postverbal positions such as the one in (12c) may be a result of the syntactic integrity of the GEN argument and the verb. Such sensitivity to syntactic constituency is expected if the NOM forms are phrasal clitics (Billings 2002).

2.3 Formalizing the transition from Wackernagel to verb-adjacent clisis
As is evident from this paper so far, languages of Taiwan and the Philippines tend toward the Wackernagel type of clisis, whereas in Sulawesi we begin to see verb-adjacent systems. Because the verb is usually initial or right after an initial element (such as NEG), the Wackernagel position often coincides with one side or the other of the verb, making it difficult for linguists to determine with certainty whether the language has a Wackernagel or a verb-adjacent clisis. To complicate matters, in some languages, once the crucial environment is found, speakers will often accept more than one clitic position, as in (13):

(13)   Kailan (ka)   sa= Maynila (ka)   p<um>unta?   (Tagalog)
     when   2SG.NOM OBL Manila   2SG.NOM AV.PRF-go
     'When did you go to Manila?'   [NB: only one ka clitic or the other]

21

Billings and Konopasky (2003) take the position that Tagalog clitic pronouns are verb-adjacent. Anderson (to appear) assumes a Wackernagel analysis.

Returning to the verb-adjacent GEN set in Kulawi and related languages, there has been some debate among Austronesianists as to the genesis of so-called conjugated verbs (van den Berg 1996, Mead 2002, Wolff 1996, Zobel 2002). Wolff offers the following reanalysis scenario (for Malay):

Stages: I. NEG=pron.GEN PV.verb *Wackernagel enclitic*

      II. NEG pron.GEN=PV.verb *Wackernagel proclitic ~ verb-adjacent clitic*

      III.   pron.GEN=PV.verb *verbal prefix*

Though agreeing overall that pronouns shift through stages such as these, we see several problems with such an analysis. We discuss each in turn.

Our first concern is with part of Wolff's characterization of stage II. While we agree that there exists a diachronic stage in which pronouns are verb-adjacent, we dispute the existence of pronouns that are positioned as Wackernagel clitics, following the first element, while affiliating prosodically to the following word. Billings (2002: 59–65), correcting Klavans (1985) and others, argues that Wackernagel clitics can be prosodically affiliated with only the preceding PrWd. Thus, we correct stage II to just "verb-adjacent clitic".

A related problem arises also with stage II. In a clause with multiple clitics preceded by a preverbal non-clitic element, a verb-adjacent pronoun can, under certain conditions, be separated from the verb. In a number of Philippine languages, in structures such as (14), although clitic pronouns are required to follow the verb, adverbial clitics can—under specified conditions that need not concern us here—follow the first PrWd of an initial phrasal constituent.

(14)   [Bukas   *ba* ng= gabi] 'y   <u>sa-sayaw</u> *sila*   ...? (Tagalog)
          tomorrow Q GEN night TOP   IRR-dance 3PL.NOM
          'Will they dance ... tomorrow night.' [Schachter and Otanes 1972: 429]

Billings and Konopasky (2003) argue that this is due to an intrinsic difference in the positioning of the two clitic types: the pronouns being verb-adjacent and the adverbials being Wackernagel clitics. If the Wackernagel position happens to be immediately pre-verbal, then it is possible for a (verb-adjacent) pronoun to precede one or more (Wackernagel) adverbial clitics; (6b) is repeated here:

(15)   Hindi *ka*   *ba* l<um>utas   ng= bigla -ng suliranin? (Tagalog)
          NEG 2SG.NOM Q AV.PRF-solve GEN sudden LNK problem
          'Haven't you ever solved a sudden problem?'

Despite being a verb-adjacent clitic, *ka* (2SG.NOM), as a monosyllabic pronoun, is still prosodically suffixal: [PrWd [PrWd Hindi] *ka*] [PrWd l<um>utas] ... As such, the Wackernagel adverbial clitic *ba* (Q) is placed after the first (matrix) PrWd, resulting in the attested form: [PrWd [PrWd Hindi] *ka*]-*ba* [PrWd l<um>utas] ... Similarly, Rudin et al. (1999: 562–66) discuss a non-pronominal Wackernagel clitic of sorts in Bulgarian that can appear right after a clitic pronoun, which itself ordinarily must be adjacent to the (following) verb. The generalization in Bulgarian is that only other clitics—even Wackernagel ones—can separate a verb-adjacent clitic from its verb. Indeed, Wackernagel clitics can appear

between pronominal clitics in both languages (Billings and Konopasky 2003: 9, Rudin et al. 1999: 564). This concern turns out to be just a red herring. It shows, however, the way these types of clitics can interact in the same clause.

Our final concern with Wolff's stages is that there is no clear connection between the so-called conjugation in (10) and other morphosyntactic features of the languages in question. We address this problem by combining Wolff's general proposal with a suggestion made by Himmelmann (1996): that the morphological asymmetry between the irrealis and realis triggers rebracketing. Specifically, the fact that the patient-voice realis is marked by a prefix while the irrealis is morphologically unmarked creates a mood slot of sorts which may be filled in the irrealis by a pronominal clitic. This can then be linked to the fact that all languages in question have externalized the Proto-Austronesian perfective infix *-in-* to a prefix (*ni-* or *i-*) and employ the bare form of the stem for the irrealis/non-past. This situation, in combination with the presence of Wackernagel clitics, can be seen as a primary factor in the move towards verb-adjacent clisis. Furthermore, following recent work in Optimality Theory on paradigmatic and systemic markedness (McCarthy 2002/2003, Itô and Mester to appear, and Kaufman 2003) we can take these motivations to be present in the synchronic grammar:

(16)   McCarthy's Optimal Pardigms (OP) approach to markedness
    a.   A candidate consists of an entire inflectional paradigm.
    b.   Markedness and input-output constraints evaluate all members of the candidate paradigm cumulatively. The violation-marks incurred by each paradigm member are added to those incurred by all the others.
    c.   The stem (shared lexeme) in each paradigm member is in a correspondence relation $\Re_{OP}$ with the stem in every other paradigm member. (That is, for every candidate paradigm P there is a relation $\Re_{OP}$ on P×P.) There is no distinctive base. Rather, every member of a paradigm is a base of sorts with respect to every other member.
    d.   On $\Re_{OP}$ there is a set of output-output faithfulness constraints.

The paradigm which is most harmonic according to both the OP constraints and the regular faithfulness and markedness constraints wins out. Thus, it is not necessary to stipulate a base form/attractor to which all forms in a paradigm are compared. Attractors are epiphenomenal in that the member that can influence other members to satisfy a given OP constraint in the most harmonic way possible will naturally do so. The following constraints are employed:

ANCHOR (STEM, L, PRWD, L) OP: Violated when a stem's left edge coincides with a PrWd's left edge in one paradigm member but not in another.
REALIZE-MORPH: A morpheme in the input must have some phonological exponent in the output (Kurisu 2002).
DEP: Every segment in the output has a corresponding segment in the input.

Given a grammar that contains a REAL prefix but no IRR affix, if ranked sufficiently high, then ANCHOR-OP is satisfied through another means. In Kulawi, this is done by placing the pronominals on the left edge, producing a symmetrically prefixed inflectional paradigm: <REAL-STEM, pron-STEM>.

(17) Tableau (Kulawi)

	Realis input: STEM *i* - PV, REAL	Irrealis input: STEM	ANCHOR-OP	DEP	REALIZE MORPH
a.	PrWd[*i*-STEM-pron]	PrWd[STEM-pron]	*		
b.	PrWd[*i*-STEM-pron]	PrWd[*i*-STEM-pron]		*	
c.☞	PrWd[*i*-STEM-pron]	PrWd[pron-STEM]			
d.	PrWd[STEM-pron]	PrWd[STEM-pron]			*
e.	PrWd[STEM-pron]	PrWd[pron-STEM]	*		*

The ANCHOR-OP constraint is satisfied either by uniform prefixing, as in candidates (b) and (c), or its uniform absence across an inflectional paradigm, as in (d). As such, candidate (a) is ruled out because the left edge of the stem is aligned with the left edge of the PrWd in one member of the paradigm but not in the other. Candidate (b) is ruled out because the realis prefix is employed without corresponding realis semantics in the input. By contrast, candidate (d) is ruled out because the realis semantics in the input does not correspond to a realis marker in the output. Thus, the winning candidate is one which employs an element whose presence is independently licensed by the input but whose position may be determined by the prosodic constraint ANCHOR-OP.

In other Kaili-Pamona languages, only local (i.e. 1SG or 2SG) pronominals become pre-verbal while the rest remain post-verbal. This is implemented here by positing person-specific constraints to be in a fixed relationship to each other within a sub-hierarchy (cf. Aissen 1999).

Fixed sub-hierarchy: ALIGN CLITIC 3sg > ALIGN CLITIC 2sg > ALIGN CLITIC 1sg

Other constraints may be interpolated between the person constraints, yielding differential treatment of salient persons along a variety of dimensions. For the Central Sulawesi languages under discussion, the relevant dimension is proclitic versus enclitic attachment. The regular enclitic alignment of the GEN clitics can only be violated for those persons whose ALIGN constraint is dominated by the interpolated ANCHOR-OP. In Da'a, ANCHOR-OP is inviolable; when the irrealis verb occurs with a non-local pronominal, the 3pl. clitic is recruited to become proclitic while the argument pronominal remains enclitic. This is analyzed here as morphological epenthesis but an analysis where *ra* is treated as an irrealis morpheme in competition with the local pronominals is also possible. The ranking in (18) derives the relevant Da'a facts.

(18)   Da'a constraint ranking:
        ALIGN CL 3SG > **ANCHOR-OP** > DEP *ra* > ALIGN CL 2SG > ALIGN CL 1SG

(19)   Bara   ku=po-balu   ri   potomu.   (Da'a)
        perhaps 1SG.GEN=CAU-buy OBL market
        'Perhaps I'll sell it in the market.'   [Abas and Friberg 1989: 47]

(20)   Nuapa ra=powia ira   ri   saa?   (Da'a)
        what *ra*-do   3PL.GEN OBL there
        'What are they going to do there?   [Abas and Friberg 1989: 104]

24

To summarize this section, we have shown some common constraints in the second position systems of the Philippines and some of the major issues regarding verb-adjacent clisis in Sulawesi. Proclisis was shown to be derivable from an ANCHORING constraint which compares irrealis and realis forms and enforces a kind of uniformity between them. This constraint is furthermore tempered by animacy hierarchy effects which determine what clitics are available for procliticization.

3.0 Ancillary issues

This last section sketches several remaining issues. These are portmanteau forms, co-occurrence restrictions, clitic doubling, and non-pronominal clitics.

To begin, quite a number of Austronesian languages have developed fused forms standing in place of certain combinations of pronominal clitics. Most commonly such a portmanteau pronoun represents the combination of a 1SG.GEN and a 2SG.NOM pair, as in (21a). Languages of this kind range from every Atayalic language (all the ones mentioned above plus both major Seediq dialects), every Central Luzon language we have looked at (Kapampangan and Sambalic), nearly every member of Central Philippines, and even languages as far flung as Kimarangang Dusun (Sabahan). The Central Luzon subgroup appears to attest the most such forms. The following are from Kapampangan:

(21)	Features involved	Clitics in isolation	Potmanteau form
a.	1SG.GEN + 2SG.NOM	ko + ka	daka
b.	1SG.GEN + 2PL.NOM	ko + kamo	dakayú
c.	1SG.GEN + 3SG.NOM	ko + ya	ke
d.	1PL.GEN + 2SG.NOM	mi + ka	daka
e.	1PL.GEN.EX + 2PL.NOM	mi + kayu	dakayu
f.	1DL.GEN.IN + 3SG.NOM	ta + ya	te
g.	2SG.GEN + 3SG.NOM	mo + ya	me
h.	2PL.GEN + 3SG.NOM	moyo + ya	ye
i.	3SG.GEN + 1PL.NOM.IN	na + tá:mu	nakatamu
j.	3SG.GEN + 3SG.NOM	na + ya	ne
k.	3PL.GEN + 1PL.NOM.IN	da + tá:mu	dakatamu
l.	3PL.GEN + 3SG.NOM	da + ya	de

[Gonzalez 1981: 177, among others]

This is the most extensive list from a single language. In other language groups, most notably Northern Philippines, pronominal forms are fused to the verb (Reid 2001). The primary significance of such forms is what they can tell us about the syntactic history of the languages. For example, at least two Subanun languages (spoken in Western Mindanao), instead of exhibiting a 1SG.GEN + 2SG.NOM portmanteau, as in (21a), attests a 2SG.GEN + 1SG.NOM pair (William Hall and Ryan Galorport, p.c.): Northern and Central Subanen *mau*; cf. Western Subanon *mu akon* '2SG.GEN 1SG.NOM'. In addition, Northern Subanen attests 3SG.GEN + 1SG.NOM *naun* and 2PL.GEN + 1SG.NOM *niyau*. Why this group of languages would have portmanteaux involving the 1SG.NOM rather than the 1SG.GEN may help explain the history of its morphosyntax.

Another type of phenomenon occasionally found in Austronesian clitic systems in co-occurrence restrictions. Here we demonstrate using the E.

Formosan language Kavalan, as reported in Huang et al. (1999: 193–96). Like quite a number of languages (cf. §1.2 above), in Kavalan, the only possible clitic-pronoun cluster is that containing a GEN Actor and a NOM non-Actor (in that order, as in the Central Luzon group and Mamanwa). The verb is typically first with the cluster following directly. Details like these are found in numerous other Austronesian languages. The unique co-occurrence restriction in Kavalan becomes evident when a non-clitic element (NEG or an auxiliary) aside from the verb is initial. In such environments, although it is apparently possible for the two to remain after the verb, one of the two clitic pronouns can appear in front of the verb. The choice of which pronoun moves is based on case and person. Huang et al. (1999: 195) write that "when the two bound pronouns are the first and second person, the Genitive pronoun manifesting the Agent participant may be added to the negator/auxiliary. However, when one of the two bound pronouns is the third person, then the first/second person pronoun gets to attach to the negator/auxiliary." Co-occurrence restrictions of a different kind in N. Tsou (Tsouic) and Agutaynen (Calamianic, Philippine) have reached our attention (Elizabeth Zeitoun and J. Stephen Quakenbush, p.c., resp.). In these languages, the presence of one pronominal form can cause deletion of a second one or replacement by an OBL pronoun. In N. Tsou there can be no overt cluster. Even if a NOM-case pronoun is possible with a verb, the presence of a GEN-case pronoun displaces the other pronoun. Similarly, in Agutaynen, a first- or second-person NOM-case Patient causes its otherwise-GEN counterpart to appear in the OBL case instead.

One important issue which we are not able examine here in any depth is that concerning the parameters of clitic doubling. We only mention here two general observations. There appears to be a close connection between verb-adjacent clisis and clitic doubling. We have uncovered only two examples of clitic doubling in second position systems: Kapampangan and Mandar. Sasak (doubled) clitics appear to show some second position effects but the available data is insufficient. In Kapampangan, clitics double the subject and object of non-Actor voice verbs but only the subject of Actor voice verbs. Abstract nouns are also typically not doubled.

(22) a. E   ya   masanting ing   igu.   (both from Kapampangan)
        NEG 3SG.NOM pretty   NOM rattan.basket
        'The rattan basket is not pretty.'   [Gonzalez 1981:161]
    b. Seli   ne   nitang   tau ing bale.
       buy.PV 3SG.GEN+3sg.NOM that.GEN-LNK man NOM house
       'The house was bought by that man.'   [Gonzalez 1981:168]

In Mandar, the Actor of intransitive and transitive verbs is doubled by a second position clitic as in (23). Patient voice verbs in Mandar seem to not require clitic doubling although more information is needed here.

(23) a. Manarang=i   ma'-elong i Murni.   (Mandar)
        skilled-3SG.NOM VRB-sing PM Murni
        'Murni is good at singing.'   [Abd. Muis Ba'dulu 1990:69]
    b. Andiang=i   pole   i Ali.
       NEG-3SG.NOM return PM Ali   (Mandar)
       'Ali isn't arriving.'   [Abd. Muis Ba'dulu 1990:70]

   c. Ma'ande=i loka   i Kaco'   (Mandar)
      eat-3sg.NOM banana PM Kaco   [Abd. Muis Ba'dulu 1990:71 ex.2b]
      'Kaco eats a banana.'

      Kulawi offers interesting evidence for the second generalization concerning clitic doubling. In Kulawi only proclitics (appearing in the irrealis patient voice) may double arguments. Enclitics are not indexed with full NPs in the same clause. The extent of this pattern requires more investigation.
      One final issue that bears on the typology of Austronesian clitic systems is clitics that are not personal pronouns. These include demonstrative pronouns, which only optionally appear in the clitic cluster in several Central Philippine languages (Lee 2004), and adverbial clitics (Kaufman to appear). Optional clitics tend to appear at the end of the cluster and therefore are not of as much interest as adverbial clitics, which are often positioned according to rather complicated and less than categorical rules. For example, in Tagalog the adverbial clitics—aside from four optional ones (McFarland 2001: 8)—all appear after a monosyllabic clitic pronoun, as in (6b)/(15) above, and before any disyllabic clitic pronoun, as in (6a). If there are both types of pronouns, the adverbial appear in between. (See also the reference above to Kavalan in this regard.) However, as (14) shows, it is occasionally possible for adverbial and pronominal clitics in the same clause not to cluster. How to position adverbial and pronominal clitics together in the same cluster has remained a challenge to generative linguistic theories for decades (Schachter 1973, Billings and Konopasky 2003, and Anderson to appear). Schachter points out that any purely syntactic theory of clitic placement would have trouble handling the portmanteau 1SG.GEN + 2SG.NOM form in Tagalog. Instead of 1SG.GEN *ko* + 2SG.NOM *ka*, the disyllabic form *kita* is used; cf. (21a) above. The problem, as Schachter points out, is that both *ko* and *ka*, being monosyllabic appear before adverbials, but disyllabic *kita* appears after adverbials. Purely syntactic models would likewise have difficulty handling such data. Our database of adverbial clitics is relatively small at this point, so we cannot make claims about distribution. Still, any theory of Austronesian clitics must consider adverbials.

References

Abas, Husen, and Barbara Friberg. 1989. Tesa-tesa eo-eo ri ara basa Daa. Palu: Summer Institute of Linguistics.
Adriani, Nicolaus, and Samuel Esser. 1939. Koelawische taalstudien. Vol. 1. Overzicht der spraakkunst, gesprekken en verhalen met vertaling. Bandoeng: Nix.
Aissen, Judith. 1999. Markedness and subject choice in Optimality Theory. Natural Language and Linguistic Theory 17. 673-711.

Anderson, Stephen R. To appear. An a-morphous account of Tagalog second position clitics. The Nature of the Word: Essays in Honor of Paul Kiparsky, ed. by Sharon Inkelas and Kristin Hanson. Cambridge, MA: MIT Press.

Ba'dulu, Abd. Muis. 1990. Frase verbal bahasa Mandar Majene. Bahasa-bahasa daerah Sulawesi dalam konteks bahasa nasional, ed. by Husan Abas and T. David Andersen, 69–84. Ujung Pandang: Universitas Hasanuddin and Summer Institute of Linguistics.

Berg, René van den. 1996. The demise of focus and the spread of conjugated verbs in Sulawesi. Papers in Austronesian linguistics 3, ed. by Hein Steinhauer, 89–114. Canberra: Pacific Linguistics.

Billings, Loren A. 2002. Phrasal clitics. Journal of Slavic Linguistics 10.53–104.

Billings, Loren A., and Abigail Konopasky. 2002. The role of morphology in ordering verb-adjacent clitics: from syntax to prosody in Bulgarian and Tagalog. Linguistics in Potsdam 19.1–26.

Billings, Loren A., and Abigail Konopasky. 2003. Reassessing the role of syntax inside the morphological word: verb-adjacent clitics in Tagalog and Bulgarian. Cornell Working Papers in Linguistics 19.20–34.

Chung, Sandra. 2003. The syntax and prosody of weak pronouns in Chamorro. Linguistic Inquiry 34.547–99.

Collins, Millard A., Virginia R. Collins, and Sulfilix A. Hashim. 2001. Mapun-English dictionary. Manila: Summer Institute of Linguistics.

Dubois, Carl D. 1976. Sarangani Manobo: an introductory guide. Manila: Linguistic Society of the Philippines.

Fukuda, Takashi. 1997. A discourse-oriented grammar of Eastern Bontoc. Manila: Linguistic Society of the Philippines.

Gonzalez, Andrew B. 1981. Pampangan: towards a meaning-based description. Canberra: Pacific Linguistics.

Himmelmann, Nikolaus P. 1996. Person marking and grammatical relations in Sulawesi. Papers in Austronesian Linguistics 3, ed. by Hein Steinhauer, 115–36. Canberra: Pacific Linguistics.

Huang, Lillian M. 1995. A study of Mayrinax syntax. Taipei: Crane.

Huang, Lillian M., Elizabeth Zeitoun, Marie M. Yeh, Anna H. Chang, and Joy J. Wu. 1999. A typological study of pronominal systems in some Formosan languages. Selected papers from the Fifth International Conference on Chinese Linguistics, ed. by Wang Hsu, Feng-fu Tsao, and Chin-fa Lien, 165–98. Taipei: Crane.

Itô, Junko, and Mester, Armin. To appear. Systemic markedness and faithfulness. Chicago Linguistic Society 39.

Kaufman, Daniel. 2003. Paradigm effects and the affix-shape/position generalization. Proceedings of the 22nd West Coast Conference on Formal Linguistics, ed. by Gina Garding and Mimu Tsujimura, 273–86. Somerville, MA: Cascadilla Press.

Kaufman, Daniel. To appear. Rigidity versus relativity in adverbial syntax. Clause Structure and Adjuncts in Austronesian Languages, ed. by Hans-Martin Gärtner, Paul Law, and Joachim Sabel. Berlin: Mouton de Gruyter.

Klavans, Judith L. 1985. The independence of syntax and phonology in cliticization. Language 61.95–120.

Kurisu, Kazutaka. 2001. The phonology of morpheme realization. Santa Cruz: University of California dissertation.

Lee, Celeste Chia-Yen. 2004. The ordering of clitic pronouns in the languages of Southeast Mindanao. Shalu: Providence University M.A. thesis.

Liao, Hsiu-chuan. 2004. Transitivity and ergativity in Formosan and Philippine languages. Honolulu: University of Hawai'i dissertation.

McCarthy, John J. 2002/2003. Optimal paradigms. Amherst: University of Massachusetts, ms. [Rutgers Otimality Archive 485.]

McFarland, Curtis D. 2001. Filipino Enclitics. Philippine Journal of Linguistics 32. 1–15.

Mead, David. 2002. Proto-Celebic focus revisited. The history and typology of Western Austronesian voice systems, ed. by Fay Wouk and Malcolm Ross, 143–77. Canberra: Pacific Linguitics.

Rau, Der-Hwa Victoria. 1992. A grammar of Atayal. Taipei: Crane.

Reid, Lawrence A. 2001. On the development of agreement markers in some Northern Philippine languages. Issues in Austronesian morphology: a focusschrift for Byron W. Bender, ed. by Joel Bradshaw and Kenneth L. Rehg, 235–57. Canberra: Pacific Linguitics.

Rudin, Catherine, Christina Kramer, Loren Billings, and Matthew Baerman. 1999. Macedonian and Bulgarian *li* questions: beyond syntax. Natural Language and Linguistic Theory 17. 541–86.

Schachter, Paul. 1973. Constraints on clitic order in Tagalog. Parangal kay Cecilio Lopez, ed. by Andrew B. Gonzalez, 214–31. Manila: Linguistic Society of the Philippines.

Schachter, Paul, and Fe T. Otanes. 1972. Tagalog reference grammar. Berkeley: University of California Press.

Weaver, Dan, and Marilou Weaver. 1964. Ranking of personal pronouns in Agusan Manobo. Oceanic Linguistics 3.161–70.

Wolff, John U. 1996. The development of the passive verb with pronominal prefix in Western Austronesian languages. Reconstruction, classification, description: Festscrift in honor of Isidore Dyen, ed. by Bernd Nothofer, 16–40. Hamburg: Abera.

Zobel, Erik. 2002. The position of Chamorro and Palauan in the Austronesian family tree: evidence from verb morphosyntax. The history and typology of Western Austronesian voice systems, ed. by Fay Wouk and Malcolm Ross, 405–34. Canberra: Pacific Linguitics.

Loren Billings
Department of Foreign Languages and Literature
National Chi Nan University
Puli, Nantou County 545
Taiwan
billings@ncnu.edu.tw

Daniel Kaufman
Department of Linguistics
Cornell University
Ithaca, NY 14853–4701
U.S.A.
dak37@cornell.edu

Phonetic Structures of Paiwan

Chun-Mei Chen
University of Texas at Austin

1. Introduction
1.1 Goals of the research
This study focuses upon a detailed description and analysis of the phonetic structures of Paiwan, an aboriginal language spoken in Taiwan, with around 53,000 speakers. Paiwan, a member of the Austronesian language family, is not typologically related to the other languages such as Mandarin and Taiwanese spoken in its geographically contiguous districts. Earlier work on phonological features of Paiwan (Chang, 1999; Tseng, 2003) sought an account in terms of segments and isolated facts about reduplication and stress, without accounting for the possible roles of phrase-level and sentence-level prosodic structures. Government Teaching Material (1993) listed 25 consonants and 4 vowels, without any description of phonetic features and phonological rules. Chang's (2000) reference grammar included 22 consonants and 4 vowels, with a very brief description of 5 phonological rules on single words. Regional diversity and 25 consonants have been mentioned in Pulaluyan's (2002) teaching material; however, no description of phonological rules was found in his material.

Maddieson (2001) has pointed out that the majority of field reports on languages give rather minimal details on their phonetic properties, sometimes nothing more than a list of symbols. Though the phonological inventory of Paiwan has been constructed in recent work (Chang, 2000; Pulaluyan, 2002), it varies from one fieldwork documentation to another. Verbal arts play an important role in communication among the ethnic members, but none of the existing documentation provides phonetic evidence for the prosodic structures. Consider the sentences listed in (1), (2) and (3).

(1)   a.   timaɟu mamazaŋinan.
                'He is a chieftain.'
          b.   timaɟu mamazaŋinan.
                'Is he a chieftain?'

(2)   a.   aitʃiu ini ka suntʃiu
                'This is not the village head.'
          b.   aitʃiu ini ka suntʃiu
                'Isn't this the village head?'

(3)   a.   timaɟu <u>vuluŋvuluŋ</u> ŋaŋa mamazaŋinan.
                'He is an old chieftain.'
          b.   timaɟu <u>vuluŋvuluŋ</u> ŋaŋa mamazaŋinan.
                'He is a very old chieftain.'

The distinction between (a) and (b) in (1)-(3) is based on prosodic structures. Work based on prosodic theory (Beckman, 1986; Pierrehumbert & Beckman, 1988; Pierrehumbert & Talkin, 1992; Pierrehumbert, 1995; Fox 2000) has shown that prosodic prominence is an important factor in the description of many phonetic phenomena. The phonetic manifestation of prosodic contrasts usually involves characteristic fundamental frequency patterns, and interactive contribution from stress, duration, pitch and syllables.

This study provides description and examination for the following issues: (i) consonant inventory; (ii) vowel inventory, (iii) (C)V(V)(C) syllable structure; (iv) stress and the interaction between vowel length, contrastive ratios, pitch and accent; (v) intonation, pitch patterns in sentences with different syntactic structures. The fieldwork reported here assesses the nature of the sounds of the Austronesian language, though it may be the last to record systematic data from a reasonable sample of Paiwan speakers. Voice data is digital recording of single words, phrases, sentences and discourse. Phonetic Measurements include vowel-quality plots, f0 and relative formant locus, vowel duration, phonetic correlates of stress, pitch track of intonation and accentuation. The goals of the study are to provide a detailed description and analysis of the phonetic structures of Paiwan and to provide empirical evidence for descriptive and theoretical documentation of the Austronesian language.

1.2 Background information

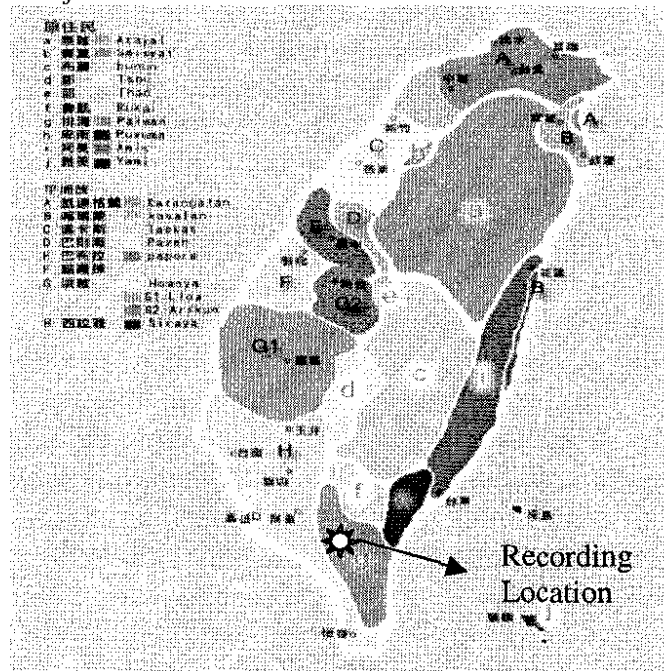Figure 1: Distribution of Austronesian Languages in Taiwan
Source: http://www.ocac.net/newocac/taiwan/taiwan_6-3.htm#

The language community where we collected data is Piuma tribe, Pingtung County. The tribe is surrounded by mountains. More than 90% of the

31

residents in the community are Paiwan aborigines. Non-Paiwan residents in the community are mainly spouses of Paiwan aborigines. Paiwan is the primary communication language in the tribe. However, the young generation speaks Mandarin in public educational institutions and Taiwanese outside of the tribe. Shown in Figure 1, the recording location is marked by an asterisk.

The informants are two male and one female Paiwan speakers, ages 48-68. All of them speak very fluent Paiwan. All the oral data in this study is based on natural speech, elicitation, narratives, and conversation.

2. Consonants

2.1 Descriptive consonant inventory

The consonant inventory of Paiwan is shown in Table 1. Note that /ts/ is the only affricate found in Paiwan. /ts/ may be a phoneme in contact with Mandarin and Taiwanese. Many other Austronesian languages such as Fijian, Javanese, Malay and Tagalog do not have this affricate phoneme. Wolff (1988) has claimed that in the languages outside Formosa *t and *C merged. The phoneme /ts/ is found only in Formosan Atayalic and Paiwanic languages.

Table 1: Paiwan consonants

	Bilabial	Alveolar	Retroflex	Palatal	Velar	Uvular	Glottal
Plosive	p b	t d	ɖ	c[1] ɟ	k g	q	ʔ
Fricative	v	s z					h
Affricate		ts					
Trill		r					
Nasal	m	n			ŋ		
Approximant	w	l	ɭ	j			

Shown in Table 1, Paiwan has alveolar, palatal, velar and uvular stops. Quite a few glottal stops [ʔ] at word-initial position were also found. There is dialectal variation between [q] and [ʔ]. For instance, *qiɭas* 'moon' is pronounced as *ʔiɭas* 'moon' in Northern Paiwan villages. Fieldwork based on Northern Paiwan (Chang, 2000) does not include the phoneme /q/. On the other hand, Government Teaching Material (1993) does not include glottal stop /ʔ/, only /q/ is found. Many words starting with uvular /q/ have free alternation of /k/ in the other dialects (Pulaluyan, 2000). In Piuma tribe, although both /q/ and /ʔ/ are found, they are free alternation.

---

[1] [c] and [ɟ] are transcribed as [tj] and [dj] in earlier Austronesian literature.

2.2 Palatalization

Palatalization is found in Paiwan. This phonological rule is represented in (4).

$$(4) \quad \begin{bmatrix} s \\ ts \\ z \end{bmatrix} \rightarrow \begin{bmatrix} \int \\ t\int \\ 3 \end{bmatrix} / \underline{\quad} i$$

It is interesting that palatalization in the Piuma tribe is optional in many cases. For instance, 'human beings' in Paiwan is '*tsaotsao*'. The female informant always pronounces it as '*tʃiaotʃiao*', in which case palatalization occurs. It is possible that /ts/ is the reflex of /c/.

2.3 Retroflexion

Paiwan has alveolar [d], [l] and retroflex [ɖ], [ɭ], as shown in Table 1. A minimal pair for /l/ and /ɭ/ was found, as in *alu* 'honey' and *aɭu* 'eight'. However, we did not find any minimal pairs for /d/ and /ɖ/, though *ɖusa* 'two' is a basic numeral in Paiwan. Ferrell (1980) has argued that only one Formosan language, Paiwan, has a phoneme inventory directly comparable to the Proto-Austronesian (PAN) inventory proposed by Dahl (1973). All of the Formosan languages, including Paiwan, have made extensive mergers and splits among sonorants and laterals; and all Formosan languages except Paiwan have also made extensive mergers and splits among non-peripheral obstruent segments. Paiwan is unique in being the only Formosan language which does not merge any of the segments *t and *d under his examination. On the other hand, though earlier study (Wolff, 1988) has proposed that retroflexion may not be a distinctive feature in Paiwan, we did find the minimal pair for /l/ and /ɭ/. Thus, we claim that retroflexion is a distinctive feature in Paiwan.

3. Vowels

3.1 Descriptive vowel inventory

The vowel inventory of Paiwan is shown in (5).

$$(5) \qquad i \qquad\qquad u$$
$$\qquad\qquad ɔ \quad o$$
$$\qquad\qquad\quad a$$

Paiwan has a five-vowel system, though the back mid vowel /o/ may be a product of language contact. The occurrence frequency of vowel phoneme /o/ is only 3%, whereas the relative occurrence of /a/ is the most frequent, as high as 39%. Blust (1988) has noted that almost all Austronesian specialists admit just four Proto-Austronesian vowels: a, ə, i and u. Though many of the words with phoneme /o/ are associated with Mandarin, Taiwanese or Japanese loanwords, /o/

also serves for basic lexicon of body parts such as throat '*li?o*', waist '*fioc*', kidney '*pulo*', and ankle '*tfimomuwu*' in Piuma Paiwan. Therefore, we include /o/ in our vowel inventory. On the other hand, we did not find any vowel length as a distinctive feature in our corpus. Paiwan vowels after voiceless stops at four places of articulation are shown in Table 2. All tokens selected were extracted from elicitation. Each token was repeated twice. There were a total of forty tokens measured to get the formants of the main vowels.

Table 2: Words exemplifying contrasts among Paiwan vowels

Bilabial		Alveolar		Velar		Uvular	
pi	pida 'how many'	ti	tisun 'you'	ki	kina 'mother'	qi	qiḷas 'moon'
pu	pumaya 'it doesn't matter'	tu	tutu 'breasts'	ku	kuɖakuɖaḷ 'big'	qu	qulivaŋrau 'rainbow'
pa	pavai 'to give'	ta	takit 'sword'	ka	kama 'father'	qa	qarəpus 'cloud'
pə	pənana? 'shoot'	tə	təməkəḷ 'to drink'	kə	kəmələŋ 'to know'	qə	qəmuɖaḷ 'to rain'
po	kipo[2] 'soil'	to	tola 'eel'	ko	kolo[3] 'head'	qo	qolivaivai 'Korean oriole'

A plot of formant values of Paiwan main vowels is illustrated in Figure 2. Tongue position for /u/ is slightly lower than that for the front high vowel /i/.

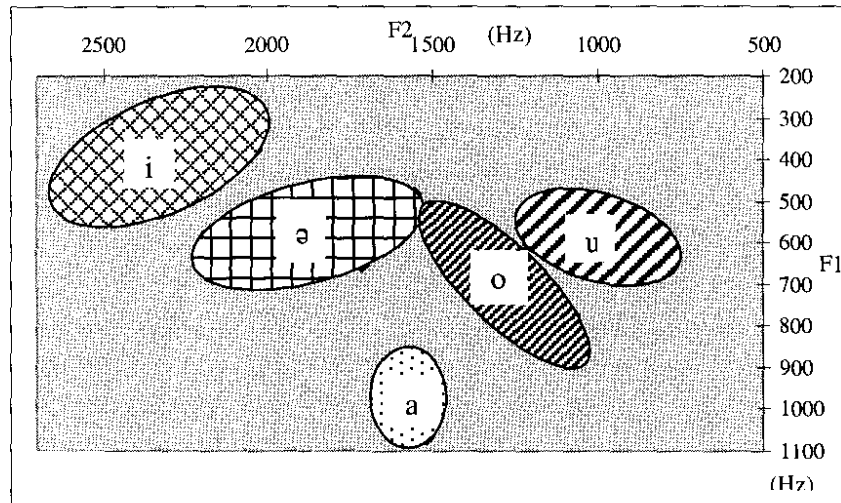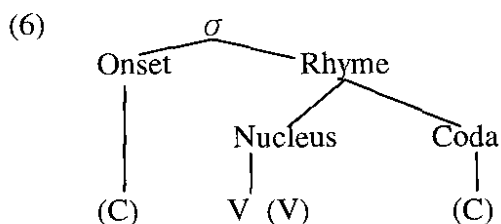Figure 2: Formant Plot of Paiwan Main Vowels

---

[2] We could not find any token with the CV combination 'po' at word-initial position. The only token we found in the corpus is the word 'kipo' ('soil').

[3] 'head' can be 'kolo', '?uḷu', or 'ulu'.

The length of a diphthong is not longer than that of a single vowel. Neither does a diphthong affect the assignment of stress.

4. Syllable Structure

Paiwan syllable structure is (C)V(V)(C). The metrical syllable structure of Paiwan is illustrated in (6).

(6)

```
                σ
         ┌──────────┴──────┐
      Onset            Rhyme
        │            ┌────┴────┐
        │         Nucleus    Coda
        │            │         │
       (C)         V (V)      (C)
```

Words with multiple syllables are not rare in Paiwan. Syllabification is based on syllable structure (C)V(V)(C). In the long word *mamazaŋinan* 'chieftain', for example, syllabification was made as shown in Table 3. The penult syllable was found to be longer and with higher pitch than the other syllables.

Table 3: Phonetic realization of *mamazaŋinan* 'chieftain'

mamazaŋinan = CV.CV.CV.CV.CVC					
Elicitation	CV	CV	CV	CV	CVC
Duration (msec)	173	136	185	216	157
Percentage	20%	16%	21%	25%	18%
Pitch (HZ)	132	122	121	149	117
Feature	H*	Low	Low	Peak	Low

Reduplication is also a morphological process to make long syllables. But it does not affect syllabification. On the other hand, affixation of vowel [a] may result in re-syllabification. Phonetically, liaison occurs whenever vowel [a] is inserted between lexical words, regardless of the preceding segments. Even if the preceding segment is also an [a], there is no lengthening. The example of numeral syllabification is shown in Table 4, in which liaison of affix [a] is always found.

Table 4: Syllabification of Paiwan numerals

Paiwan	Gloss	Paiwan	Gloss	Syllabification
ḍusa CVCV	two	ḍusa_a puḷu? CVCV+V+CVCVC	twenty	ḍu.sa.pu.ḷu? CV. CV. CV. CVC
unəm VCVC	six	unəm_a puḷu? VCVC+V+CVCVC	sixty	u.nə.ma.pu.ḷu? V.CV.CV.CV.CVC
pitu CVCV	seven	pitu_a_ puḷu? CVCV+V+CVCVC	seventy	pi.tua.pu.ḷu? CV.CVV.CV.CVC

35

daidai CVVCVV	one hundred	pitu a idai CVCV+V+VCVC	seven hundred	pi.tua.i.dai CV.CVV.V.CVC

5. Stress

5.1 Descriptive Stress Patterns

Paiwan stress falls on penultimate syllables in general. However, there are a few exceptions. Our description of Paiwan stress is in (7).

(7) a. In CV or CVCV syllables, stress falls on the first V ($\sigma \leq 2$)
     b. Stress falls on the penultimate syllable elsewhere ($\sigma > 2$)
     c. Stress patterns are fixed in content words but not fixed in function words, reduplicated and affixed forms.
     d. If the nucleus of the penultimate syllable is a weak schwa [ə], stress shifts to the last syllable; when both penultimate and last nuclei are schwa [ə], stress falls on the last [ə].

Earlier work on Proto-Austronesian (PAN) stress mainly focuses on comparison and reconstruction. Wolff (1993) has proposed that PAN roots had a stress contrast in the final two syllables of the root. He argues that in PAN the stress patterns fell on the penult of the root if it was long and on the final syllable of the root if the penult was short. Nevertheless, no further description of the interaction between vowel quality and stress patterns was found in his study. In fact, vowel length is a phonetic realization or a correlate of stress, rather than a prerequisite to trigger stress. On the other hand, affixation has been found to affect stress in earlier PAN studies (Wolff, 1993; Zorc, 1993). There was the phenomenon of accent shift whereby the affixed form had the accentual pattern opposite to the pattern of the root. This is true for Paiwan stress, as we have described in (7c) that stress in affixation is not fixed.

Syntactic classes in Paiwan, as far as we have observed, do not affect the assignment of stress in Piuma Paiwan. Stress falls on the penult in numerals and free pronouns. As we have described in (7d), stress never falls on a schwa, unless the schwa is the final syllable. Sample words are shown in (8). Stress shifts to the last syllable when the penultimate syllable is a schwa. Shown in (8b) and (8c), when schwa [ə] is in both penultimate and final nuclei, stress falls on the last [ə].

(8)   Paiwan   Gloss   Stress

     a. kəməlaŋ   'to know'   (x)
         [CVCVCVC]   [σ σ σ́]

     b. masəŋsəŋ   'to work'   (x)
         [CVCVCCVC]   [σ σ σ́]

     c. mipərəpər   'to fly'   (x)
         [CVCVCVCVC]   [σ σ σ σ́]

36

5.2 Phonetic Correlates of Paiwan Stress

Acoustic correlates of stress in English have been widely studied. Fry (1955, 1958) investigated the acoustic and perceptual correlates of lexical stress in English noun/verb word-pairs. Following his hierarchy, duration is a more effective cue to stress than intensity, and pitch is an even more effective cue than duration. However, as early as Berinstein (1979), it has been disclosed that languages with phonemic length do not use duration as a correlate of stress. Moreover, pitch-accent languages such as Japanese use pitch as a correlate of stress to a greater extent than English (Beckman 1986). While whether the conclusions of these studies can apply to Paiwan is unknown, it is clear that phonetic cues to stress in different languages can vary.

As we have seen in section 3.1, vowel length is not phonemic in Paiwan, and tonal feature is not distinctive in this language. Three phonetic cues to stress were selected here: vowel duration, pitch and intensity. We propose the following hypotheses of Paiwan stress in (9).

(9)   a.   Stressed syllables have longer DURATION
      b.   Stressed syllables have higher PITCH
      c.   Stressed syllables have greater INTENSITY

In the following section, contrast of stressed and unstressed syllables will be examined to test the hypothesis of Paiwan stress correlates.

5.3 Contrast of Stressed and Unstressed Syllables

Figure 3 illustrates the duration contrast between stressed vowel and unstressed vowels. In syllables more than two morae (>2 σ s), the duration was measured from penultimate vowel and final vowel in stressed and unstressed positions. In three and five-mora syllables, when stressed, duration of a vowel is more than twice that of an unstressed vowel. The duration of a stressed vowel is overall longer than an unstressed one, more than 1.7 times that of an unstressed vowel. It is clear that duration is a correlate of Paiwan stress.
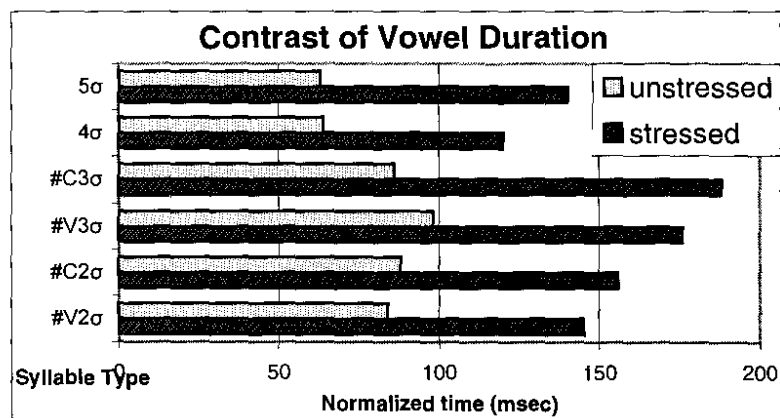

Figure 3: Vowel Length Contrast of Stressed and Unstressed Syllables

As far as pitch is concerned, stressed vowels were perceived to have higher pitch. Figure 4 summaries the means for pitch values in stressed and unstressed positions. Pitch values were measured at the midpoint of each syllable. In syllables with more than two moraes, pitch contrast was measured at penultimate and final vowels of each token.
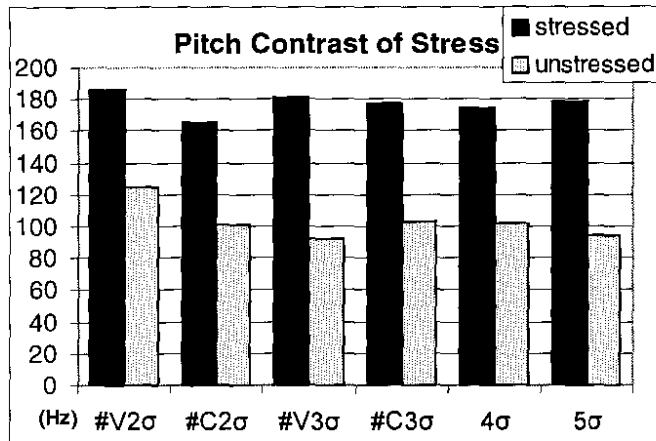

Figure 4: Pitch Contrast of Stressed and Unstressed Syllables

All the stressed vowels measured have higher pitch than the unstressed vowels, as shown in Figure 4. The hypothesis that stressed syllables have higher pitch is therefore verified. Note that in multiple syllables (>5 σ s), stress is more correlated with stress than duration. Multiple syllables are usually compound words, such as 'eleven': *dapuḷuʔ saga ita*. Vowel contrast in compound words is not as apparent as that in single words. Vowel duration in the final position of compound words is usually longer, regardless of stress. Even though duration contrast is neutralized in multiple syllables, pitch contrast in stressed and unstressed positions is retained.

However, the intensity contrast between stressed and unstressed syllables is not consistent in all the syllable types. This indicates intensity may not be a robust correlate of Paiwan stress, at least, less correlated than duration and pitch. Thus, our hypothesis in (9c) may not be true. Unstressed final syllables may have greater intensity than stressed penultimate syllables.

5.4 Emphatic Degree Accentuation

Emphatic degree accentuation was found in Paiwan discourse. This type of accentuation can be found in many other languages. Emphatic accentuation in Paiwan is used not only to signal the degree of intensity but also to function as different degree of semantic lexicon. For instance, 'very old' has longer duration and higher F0 than 'old'. There is no degree adverb 'very' or 'extremely' in Paiwan lexicon. The most common way to denote the meaning of 'very' is either to reduplicate stems or by means of prosodic prominence. In the case of *vuluŋvuluŋ* 'old', since it is already a reduplicated form, the distinction between

38

'old' and 'very old' is thus imposed on prosodic elements. The contrast of pitch track between *vuluŋvuluŋ* 'old' and *vuluŋvuluŋ* 'very old' is shown in Figure 5.
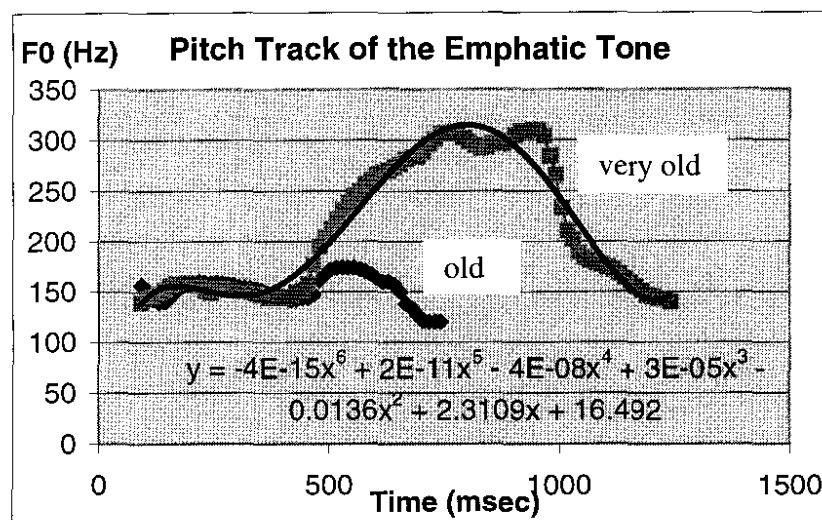

Figure 5: Pitch Contrast of *vuluŋvuluŋ* 'old' and *vuluŋvuluŋ* 'very old'

Emphatic accentuation results in longer duration of the reduplicated vowels, as shown in Figure 5. The reduplicated vowel RV1 at penultimate position has the most significant lengthening, around 1.35 times that of an unaccented vowel. However, accentuation does not lengthen the relative duration of the other vowels in stems. Lengthening is most apparent in reduplicated vowels at syllable-penultimate position.

Emphatic accent has higher pitch than the control one. Higher pitch is most apparent in the penultimate-syllable position. On the other hand, amplitude is not correlated with accentuation. Emphatic vowels do not have greater intensity than the control ones. This indicates intensity is not a correlate of Paiwan accent.

In summary, emphatic accentuation signals both degree of intensity and distinctive levels of degree. Emphatic accent in Paiwan is realized as higher pitch and longer duration. Intensity is not a correlate of emphatic accentuation.

6. Intonation

6.1 Descriptive Intonation

We described the intonation types of Paiwan in terms of features. The distinctive features in our description include shape of pitch contour, starting/ending points of pitch contour, pitch range, transition, rising and falling.

The most distinctive intonation in Paiwan is the feature at the ends of a declarative and a yes/no question. Declarative sentences in Paiwan usually have falling ending pitch. Yes/no questions, on the other hand, are marked by high rising pitch. Neither does a question marker nor movement of auxiliaries form a yes/no question in Paiwan. Though there are some WH-words in Paiwan such as *ima* 'who', *nəma* 'what', *inu* 'where/which', *kəmuda* 'how', *aku* (a zua) 'why',

there is no question marker except for one particle borrowed from Taiwanese. The differences between declaratives and yes/no questions are in the H*L higher starting points and the LH% high pitch at the final-nuclear position of the questions. Intonation in declarative and yes/no question is correlated with the presence of the final rising. Peak at sentential-penultimate position is retained in the declaratives with emphatic accent. Note that questions in Paiwan are correlated with a rising ending, excluding WH-Questions. The typological description of question intonation in Paiwan is summarized in (10).

(10)   a.   Types of questions with a rising ending:
                  (i) Yes/No: Is he a chieftain?
                  (ii) Rhetorical: He is a chieftain, isn't he?
                  (iii) Alternative: Is he a chieftain or village head?
        b.   Types of questions with a falling ending: WH-Questions
             Where, When, What, Who, How, Why

In the following section, four types of intonation will be analyzed: declarative, yes/no question, WH-questions and negation. Gaps caused by voiceless segments in pitch tracks were not avoided, because the recording process was not experimental-designed. All the sentences were collected from natural speech.

6.2 Pitch tracks of sentences

The typical contrast between declaratives and yes/no questions is shown in (1) and (2). A pitch track of the sentences in (1) shows the distinctive prosodic structures of intonation. Declaratives have a distinct pitch contour from yes/no questions. The pitch contrast of the sentences in (1) is illustrated in Figure 6.
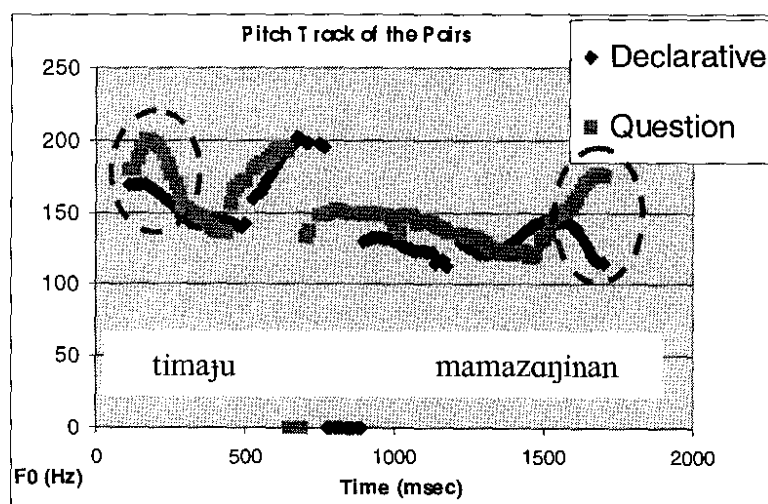

Figure 6: Pitch Tracks of *timaju mamazaŋinan*

Shown in Figure 6, yes/no question has a higher starting point and a rising ending LH%. Notice the slight dropping before the final rising. All the yes/no

questions examined have this dropping before the final rising. The rising ending feature of yes/no question is retained even in embedded questions. For instance, in the embedded yes/no questions such as 'he asked me if he is a chieftain (is he a chieftain)' or 'can you tell me if people in the village know he is a chieftain', the final content word *mamazaŋinan* 'chieftain' still keeps the rising ending. Prosodic features do play an important role in Paiwan discourse and conversation.

Negation sentences have very similar pitch contours to declarative intonation, with a falling ending. The negation word *ini-ka* 'not' does not have any particular rising or falling prosodic feature. When the negation word *ini-ka* 'not' is placed at the end of a declarative sentence, it forms a rhetorical question, such as *timaju mamazaŋinan ini* 'he is a chieftain, isn't he?' The intonation of a rhetorical question is very different from negation intonation. A pitch track of a rhetorical question is illustrated in Figure 7.
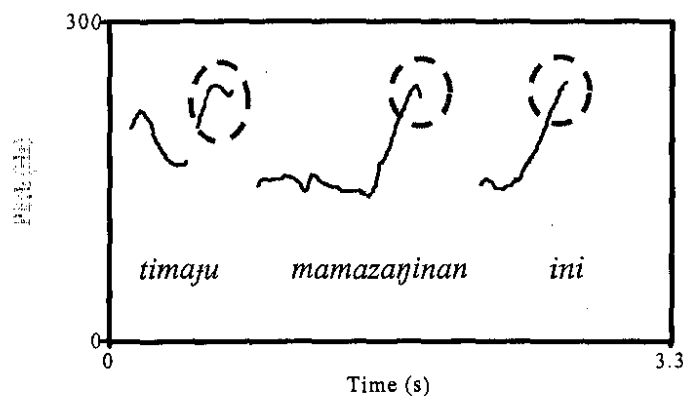
Figure 7: Pitch Track of *timaju mamazaŋinan ini*

Shown in Figure 7, not only the negation word *ini* 'not' but also the other content words have rising boundary tone H%. Pitch range is not a correlate of intonation typology, as it varies from one pair to another.

As shown in (10a), the third type of question with a rising ending is alternative. In Paiwan, alternative word *manu* 'or' is placed between two alternatives. All the content words in the alternative sentence have a rising boundary tone LH%, which indicates all the content words are affected by the interrogative tone, and the sentence ends with a rising tone. The intonation peak usually falls on the second nucleus of the alternative word *manu* 'or'.

Finally, we analyzed the intonation pattern of WH-Questions. As we have described in (10b), WH-questions end with a falling tone. Pitch tracks of WH-questions show that intonation peak falls on WH-words, with or without affixation. Examples of WH-questions are shown in (11). Pitch tracks of the sentence are illustrated in Figure 8.

(11) a.   a nǝma hokan
          'What do you like to eat?'

41

b.   inu a gakiduluwan
          'Where is the school?'
    c.   ma-inu-sun
          'Where are you going?'

Shown in Figure 8, intonation peaks fall on the second syllables of WH-words *nəma* 'what' and *inu* 'where'. Unlike other types of questions with a rising ending, WH-questions end with a falling tone. The features of WH-questions are summarized as follows: a falling ending and intonation peak on WH-words.
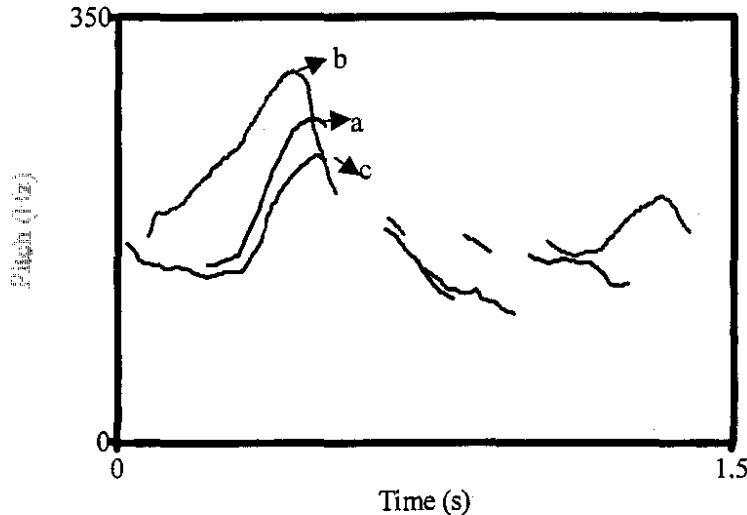
Figure 8: Pitch Tracks of WH-Questions

7. Conclusion

In this study, consonants, vowels, syllable structure, stress and intonation in Paiwan were examined. Consonant inventory, five-main vowels, formant plot of Paiwan main vowels have been illustrated. Paiwan has bilabial, alveolar, palatal, velar, uvular and glottal stops. Syllabification in Paiwan is based on syllable structure (C)V(V)(C). The penult syllable was found to be longer and with higher pitch than the other syllables in a prosodic word.

It has been shown that phrase-level and sentence-level prosodic features play important roles in Paiwan. In phrase-level of prosody, penultimate stress and non-stress accent have been investigated. In sentence-level, both declarative intonation and interrogative intonation have been examined. Paiwan stress falls on penultimate syllables, if the prosody word is a stem. Stress patterns are fixed in content words but not fixed in function words. Weak schwa [ə] may result in stress shift to the final syllable, if the final syllable is not a schwa.

Vowel length is not phonemic in Paiwan. The hypotheses of Paiwan stress were verified. Duration and pitch are correlates of Paiwan stress. Stressed syllables have longer duration. All the stressed vowels measured have higher

42

pitch than the unstressed vowels. However, the intensity contrast between stressed and unstressed syllables is not consistent. Emphatic accentuation in Paiwan signals both degree of intensity and distinctive levels of degree. Emphatic accent in Paiwan is realized as higher pitch and longer duration. Intensity is not a correlate of emphatic accentuation.

Intonation in Paiwan questions has been categorized into two types, with a rising ending or a falling ending. Types of questions with a rising ending include yes/no questions, rhetorical questions and alternative questions. Questions with a falling ending include WH-Questions. Declarative and negation sentences have a falling ending. Negation sentences have very similar pitch contour to declarative intonation. Questions in Paiwan are correlated with a rising tone; pitch peak falls on WH-words. Yes/no question has higher starting points and a rising ending LH%. The rising ending feature of yes/no question is retained even in embedded questions. In rhetorical questions, not only the negation word *ini* 'not' but also the other content words have rising boundary tone H%.

Thus far, empirical study of Austronesian prosodic structure is relatively rare, and the discussion on the typology of Formosan prosody is even less. We need more detailed description and studies on Formosan languages, not only Paiwan but also other endangered aboriginal languages.

Endnotes

References
BECKMAN, MARY E. 1986. Stress and Non-Stress Accent. The Netherlands: Foris Publications Holland.
BERINSTEIN, A. 1979. A Cross-Linguistic Study on the Perception and Production of Stress. UCLA Working Papers in Phonetics 47, UCLA.
BLUST, ROBERT A. 1988. Austronesian Root Theory. Philadelphia: John Benjamins Publishing Company.
CHANG, XIU-JUAN. 2000. Paiwan Yu Cankao Yufa [Paiwan Reference Grammar]. Taipei: Yuanliu Quban ShiyeYiouxian Gongse [Yuanliu Publishers].
DAHL, OTTO C. 1973. Proto-Austronesian. Lund, Sweden: Scandanavian Institute of Asian Studies, Monograph Series No.15.
FERRELL, RALEIGH. 1980. Phonological Subgrouping of Formosan Languages. Austronesian Studies: Papers from the Second Eastern Conference on Austronesian Languages, ed. by Paz B. Naylor, 241-254. Michigan: Ann Arbor.

FOX, ANTHONY. 2000. *Prosodic Features and Prosodic Structure*. Oxford: Oxford University Press.

FRY, D. B. 1955. Duration and Intensity as Physical Correlates of Linguistic Stress. Journal of the Acoustical Society of America 27: 155-58.

——. 1958. Experiments in the Perception of Stress. Language and Speech 1: 126-152.

MADDIESON, IAN. 2001. Phonetic fieldwork. Linguistic Fieldwork, ed. by Paul Newman and Martha Ratliff, 211-229. Cambridge: Cambridge University Press.

PIERREHUMBERT J. & M. BECKMAN. 1988. Japanese tone structure. Cambridge, MA: MIT Press.

PIERREHUMBERT J. & TALKIN. 1992. Lention of /h/ and glottal stop. Papers in laboratory phonology II: gesture segment prosody, ed. by G. Doherty and D. R. Ladd, 90-117. Cambridge: Cambridge University Press.

PIERREHUMBERT J. 1995. Prosodic effects on glottal allophones. Vocal fold physiology: voice quality control, ed. by O. Fujimura and M. Hirano, 39-60. San Diego: Singular Publishing Group.

PINGTUNG COUNTY GOVERNMENTT, 1993. Pingtung Xian Muyu Jiaocai—Paiwan Yu [Native Language Teaching Material of Pingtung County—Paiwan]. Taiwan: Pingtung.

PULALUYAN, TALIGU. 2000. Pinaiwanan [An Introduction to Paiwan]. Printed in Taipei.

ROSS, MALCOLM. 2002. The History and Transitivity of Western Austronesian voice and Voice-Marking. The History and Typology of Western Austronesian Voice Systems, ed. by Fay Wouk and Malcolm Ross, 17-62. Australia: The Australian National University.

TSENG, MEYLYSA. 2003. Reduplication as Affixation in Paiwan. Presented in Austronesian Formal Linguistics Association Tenth Annual Meeting, University of Hawaii at Manoa.

WOLFF, JOHN U. 1988. The PAN Consonant System. Studies in Austronesian Linguistics, ed. by Richard McGinn, Monographs in International Studies Southeast Asia Series No. 76, 125-147. Ohio: Ohio University Center for Southeast Asian Studies.

——. 1993. Proto-Austronesian Stress. Tonality in Austronesian Languages, ed. by Jerold A. Edmondson and Kenneth J. Gregerson, 1-15. University of Hawaii Press.

ZORC, DAVID. 1993. Overview of Austronesian and Philippine Accent Patterns. Tonality in Austronesian Languages, ed. by Jerold A. Edmondson and Kenneth J. Gregerson, 17-24, University of Hawaii Press.

CHUN-MEI CHEN
University of Texas at Austin
Department of Linguistics, University of Texas at Austin, Austin, TX78741
USA
chench@mail.utexas.edu

Vowel Dispersion in Truku

Wen-yu Chiang and Fang-mei Chiang
National Taiwan University

0. Introduction

This study investigates the dispersion of vowel space in Truku, an endangered Austronesian language in Taiwan. Adaptive Dispersion (Liljencrants and Lindblom, 1972; Lindblom, 1986, 1990) proposes that the distinctive sounds of a language tend to be positioned in phonetic space in a way that maximizes perceptual contrast. For example, languages with large vowel inventories tend to expand the overall acoustic vowel space. Adaptive Dispersion predicts that the distance between the point vowels will increase with the size of a language's vowel inventory. Thus, the available acoustic vowel space is utilized in a way that maintains maximal auditory contrast.

Flemming (1995, 1996, 2001) introduced another version of Dispersion Theory, which incorporates the mechanisms of dispersion into current phonological theories of constraints on development of phonological inventories. Expressed in terms of Optimality Theory, this version of Dispersion theory claims that constraints favoring less perceptually confusable contrasts are ranked higher than constraints that would favor more confusable contrasts. Thus, the markedness of a sound would depend on the sounds that it contrasts with in a particular inventory. Flemming proposes three central constraints: i) Maximize the distinctiveness of contrasts; ii) Minimize articulatory effort, and iii) Maximize the number of contrasts. The particular requirements of each language's inventory determine the order in which these three constraints are ranked, which results in divergent language-specific vowel systems.

The current study provides unique data in Truku, which is particularly suited to testing dispersion theory, because they reveal conflicting opinions as to the number of vowels in the Truku inventory. In the scant literature currently available, Truku is inconsistently claimed to contain three vowels (/i/, /u/ and /a/), four vowels (/i/, /u/, /o/ and /a/) or five vowels (/i/, /e/, /u/, /o/ and /a/), not including the schwa. Moreover, many native speakers of Truku also differ with respect to their perception of the number of vowels in its inventory. The current study investigates two research questions: 1) whether dispersion theory can account for variations in the reported number of Truku vowels and 2) how syllabic stress may interact with the vowels' distribution and contrastiveness.

1. Background

In terms of linguistic typology, Truku is described as a descendant of proto-Atayal. Due to its long alienation from Atayal, Truku diverges considerably from its prototype. Truku had no writing system, which is common to all Formosan aboriginal languages, so evidence of diachronic change cannot be determined from written records. The Atayal tribe can be divided into two subtypes: Atayal and Seediq (Li *et al*, (1963) and Hung (1993)). The latter, further divided into Eastern and Western Seediq, are dispersed over Taiwan's Hualien and Nantou Counties. The languages spoken

by the Seediq can be grouped into three dialects: Teuda, Tkdaya and Truku, which belonged to the Eastern Seediq. The Eastern Seediq population of Hualien is 24,183, which represents around 84% of the entire Seediq population (Huang, 1990).

In recent years, debate has arisen as to the composition of the Truku vowel inventory. The Truku vowel system, in the divergent opinions of researchers, consists of three, four or five vowels.[1] Hu (2003) claims that the Truku inventory consists of the three vowels /i/, /u/ and /a/. The ALCD (Center for Aboriginal Languages Cultures Education, 2001) classifies Truku as having the four vowels /i/, /u/, /o/ and /a/. Li (1992) claimed that all three dialects of Seediq have /i/, /u/, /o/ and /a/, while the occurrence of /e/ or /ə/ depends on regional variation.[2] Yang (1976) and Chang (2000) consider all dialects of Seediq to have a five-vowel system, consisting of /i/, /e/, /u/, /o/ and /a/.

The writing system of Truku language is currently undergoing the process of construction through the discussion of its native speakers. Disagreement concerning the number of phoneme emerges as an unresolved problem. The current study thus aims to investigate the relationship between the number of vowel inventory and the distribution of vowel space.

2. Method

Materials

A list of seventy-seven disyllabic words containing the vowels [i], [I]/[E], [u] and [U]/[O] were selected from our field recordings. Forty-four were presented in stressed conditions and thirty-three in unstressed conditions.[3] The wordlist, including vowel occurrence tabulations, are given in Table 1. The words were selected specifically to include as many contexts as possible, in order to include the possibility of coarticulation effects. Gaps in this list result from either the lack of lexical item to fulfill that condition, or the absence of an acceptable token of such an item in our field recordings[4]. Forty disyllabic words containing the vowels [a] and [ə] were added to test the integrity of the vowel chart. The column of unstressed [ə] in Table 1 is empty; our consultations with informants yielded no example in which [ə] can appear as the second syllable of a disyllabic word.

We observed that [i], [u] and [a] were the distinctive vowels upon which all speakers agreed, unlike [e] and [o]. The controversial nature of [e] and [o] was marked by the alternating use of capital "I" and "E" for the former, and "U" and "O" for the latter, the capital letters representing the underlying form as recognized by different groups of speakers.

Table 1a. Items representing *stressed* vowels in Truku

Preceding Consonants	[i]	[I]/[E]	[u]	[U]/[O]
p	*pila* 'money'	*pedang*'[name]'	*pungu* 'knee'	*poli* '[name]'
b	*birat* 'ear'		*bunga*'sweet potato'	*boyak* 'boar'
m	*mirit* 'goat'	*meyaw* 'awake'	*muhing* 'nose'	*moda* 'to pass'
t	*tcimu* 'salt'	*temu* '[name]'	*tunux* 'head'	
d				*dohong*'mortar'

46

n	nisu 'your'		nunuh 'a fig'	noxay 'because'
k	kingal 'one'		kumay 'bear'	
g	gimi 'to find'		gupun 'teeth'	gomuk 'a lid'
ng	ngiraw 'mushroom'		ngudus 'beard'	
q			quyu 'snake'	qomi 'needle'
r	risaw 'boy'	rema 'five'		rodox 'chicken'
l	lingis 'to cry'		lungaw 'bottle'	lomun 'liver'
y		yeku '[name]'	yumaw '[name]'	yona '[name]'
w	wihe 'spoon'			
h	hiya 'he'	heni 'here'		holing 'dog'
s	cida 'branch'		sudu 'grass'	soki 'knife'
z	zima 'banboo'			
Total	**14**	**6**	**12**	**12**

Table 1b. Items representing *unstressed* vowels in Truku

Preceding Consonants	[i]	[I]/[E]	[u]	[U]/[O]
p	rapit 'flying squirrel'		gupun 'teeth'	
b	labis 'mosquito'		bubu 'mother'	
m	qomi 'needle'		lomun 'liver'	
t			utux 'ghost'	
d			sudu 'grass'	
n	heni 'here'		munuh 'to breast-feed'	
k	daking 'to grow up'		yaku 'I'	
g	ləgi 'to shake'		megun 'to tie'	
ng	dangi 'lover'		pungu 'knee'	
q		laqe 'child'	uqun 'sth to eat'	
r	mirit 'goat'		paru 'big'	tro 'three'
l	holing 'dog'		malu 'good'	
y			quyu 'snake'	
W	tciwin 'little'			
H	muhing 'nose'	wihe 'spoon'		dohong 'mortar'
X	ləxi 'bamboo shoot'		kuxul 'to like'	
S	nasi 'if'		nisu 'your'	
Z	buzi 'arrow'			
Total	**14**	**2**	**15**	**2**

Table 1c. Items containing the vowels [a] and [ə] in Truku

Preceding Consonants	Stressed [a]	[ə]	Unstressed [a]
p	paru 'big'		sapah 'house'
b	baraw 'upper'		yabas 'guava'
m	maxal 'ten'	məkan 'to eat'	tama 'father'
t	tama 'father'	təro 'three'	watan '[name]'
d	daha 'two'	dəmux 'grains of'	idas 'moon'
n	naku 'my'		kana 'all'
k	kari 'language'	kəla 'to understand'	

47

g	gaga 'that'	gəhak 'seed'	baga 'hand'
ng	ngali 'to take'		bunga 'sweet potato'
q	qalux 'balck'	qəpi 'to squeeze'	niqan 'to own'
r	rawa 'bascket'		
l	laqe 'child'	ləxi 'bamboo shoot'	alang 'village'
y	yamu 'you (pl)'		sayang 'now; today'
w	wadu 'honey'		rawa 'bascket'
h	habuk 'weast belt'	həngak 'air'	daha 'two'
x			maxal 'ten'
s	sari 'taro'	səpac 'four'	məsa 'this way'
Total	**16**	**9**	**15**

Speakers

Three male and three female native speakers of Truku participated in this experiment. The participants were divided into "5-vowel", "3-vowel" and "4-vowel" groups, according to the number of vowels they perceived to be in their language; each group included one male and one female speaker. The 5-vowel group perceived the vowels [i], [e], [a], [u] and [o] in Truku, while the 3-vowel group perceived [i], [u] and [a]. The speakers in the 4-vowel group perceived [i], [u], [o] and [a]. Thus, the vowel inventory for the 5-vowel group was: [i], [E], [a], [u] and [O]; for the 3-vowel group: [i], [I], [a], [u] and [U]; and for uncertain group, [i], [I], [u] and [O].

All of the speakers lived in Hualien County, where they spoke Truku on a daily basis before coming to Taipei at the age of twenty or above.

Recordings

Recording sessions consisted of two parts. The first familiarized the informants with the items on the randomized list of words; the second directed the informants to read each word once in Truku after hearing the corresponding Mandarin gloss. Some items contained two experimental vowels; these appeared twice in the wordlist, but never adjacently. The recordings were conducted in quiet rooms using a Sony TCM 5000-EV portable tape recorder and an electric condenser microphone located approximately 20 centimeters away from the mouth.

Measurements and Analysis

The signals were later transferred to the Kay Elemetrics Computerized Speech Lab (CSL) in National Taiwan University's Phonetics Laboratory, using a 10-kHz sampling rate. Each word was displayed on a wideband spectrogram with a formant history, using a 300 kHz bandwidth cutoff. Using both visual and auditory cues, we removed the preceding and following consonantal transitions around the experimental vowel, so that only the steady state of each vowel remained. Then, five points within this stable range were extracted at equal distances: 0%, 25%, 50%, 75% and 100% of the steady-state duration. F1 and F2 values were obtained for each of these points by means of an LPC analysis. However, since the vowels [e] and [o] are often diphthongized, only the portion before their offglide was extracted. If formant values were unavailable for any of the default five points (0%, 25%, 50%, 75% and 100%),

the value of an adjacent point was adopted, which was determined by examination of the LPC formant history.

3. Result

The following sections will present vowel distribution across groups in terms of the factors vowel identity and gender, as produced in the stressed condition. First, Figures 1a to 1f show the vowel distribution of each speaker in the stressed condition. F1/F2 plots reveal that except for the 5-vowel female speaker (Fig. 1a), who has clear distinctions for all the vowels, the other speakers show overlap among vowels. This is especially apparent in Figure 1b; for the 3-vowel female speaker, the controversial and non-controversial vowels occupy almost the same space. Furthermore, the male speakers' overall vowel space appears to be articulated further back than the female speakers'.

a. 5V-F's vowel distribution

b. 3V-F's vowel distribution

c. 5V-M's vowel distribution

d. 3V-M's vowel distribution

49
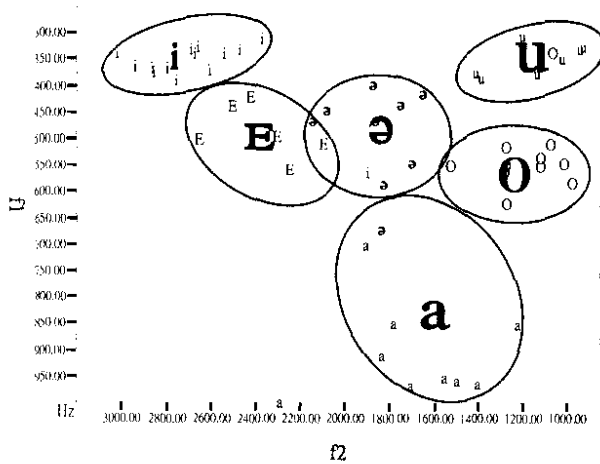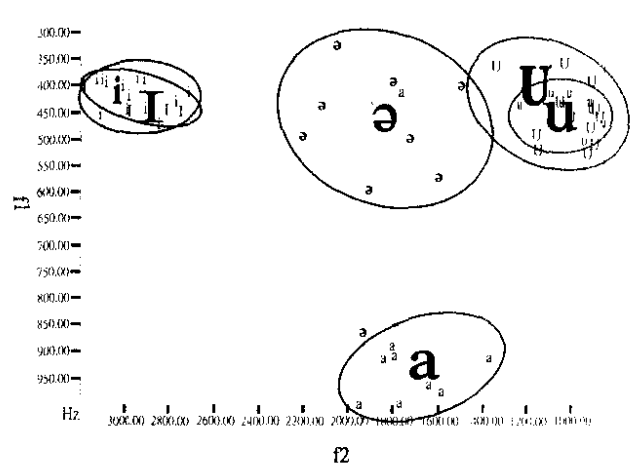
e. 4V-F's vowel distribution
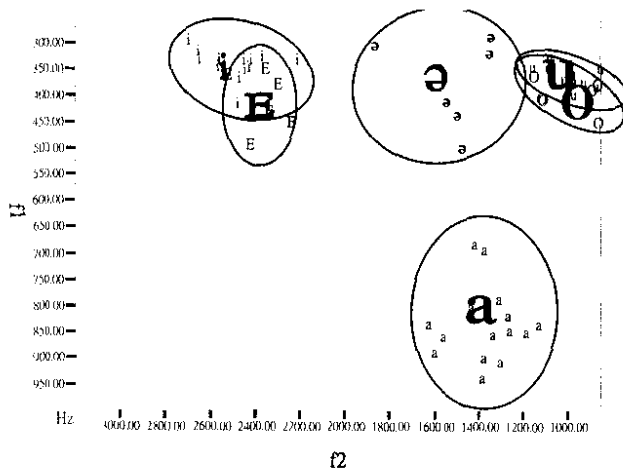
f. 4V-M's vowel distribution
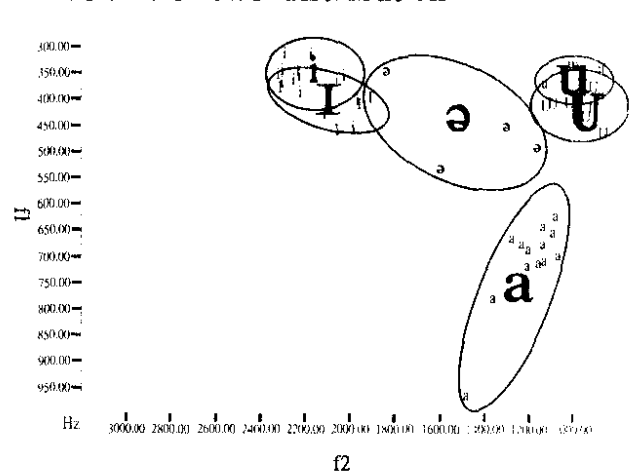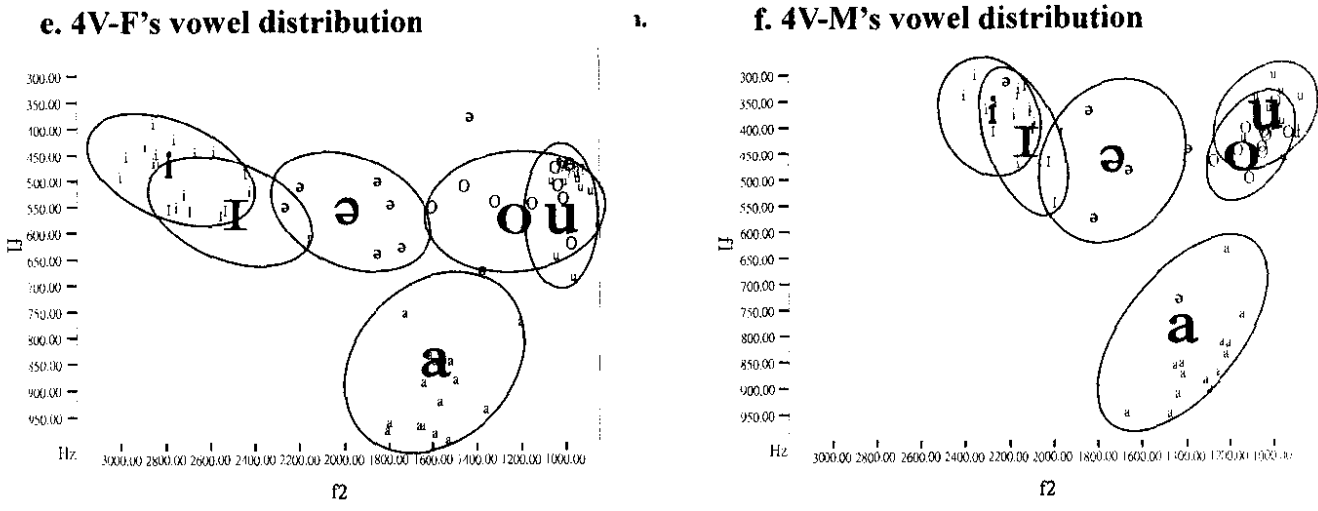
Figure 1. The six speakers' individual vowel distributions in the stressed condition

Front vowels: [i] vs. [I/E]
Vowel and Gender factors.
The data were analyzed using a MANOVA with factors Vowel and Gender. With the F1 and F2 values as the dependent variable, there were significant effects for both factors. (For [i]-[I/E] difference: $F(1, 1051) = 146.2, p < .01$; For female/male difference: $F(1, 1051) = 653.2, p < .01$). That is to say, the non-controversial [i] differ from the controversial [I/E] with respect to both F1 and F2 values despite the individual differences in perception. On the other hand, the factor of Gender is also significant for front vowels. The females' articulation of front vowels is lower and more fronted than the males'.

Within-subject variation.
We will now turn to the aspect of within-subject variation in comparing [i] and [I/E]. Table 2 provides the formant values and the ANOVA results for the comparison of the two front vowels [i] and [I/E].

Table 2. Formant values (Mean/SD) of the front vowels

Mean (SD): Hz		[i]	[I/E]	ANOVA
5V-Female	F1	364 (61)	486 (51)	$F(1, 98) = 61.53, p < .001$
	F2	2649 (280)	2371 (196)	$F(1, 98) = 22.04, p < .001$
4V-Female	F1	464 (44)	556 (41)	$F(1, 98) = 66.11, p < .001$
	F2	2784 (153)	2529 (217)	$F(1, 98) = 26.86, p < .001$
3V-Female	F1	416 (29)	433 (22)	$F(1, 88) = 3.32, p = .072$
	F2	2951 (129)	2928 (194)	$F(1, 88) = .314, p = .577$
5V-Male	F1	337 (27)	408 (57)	$F(1, 98) = 51.81, p < .001$
	F2	2490 (127)	2360 (99)	$F(1, 98) = 19.69, p < .001$
4V-Male	F1	366 (41)	425 (75)	$F(1, 93) = 23.99, p < .001$
	F2	2220 (104)	2076 (113)	$F(1, 93) = 24.12, p < .001$
3V-Male	F1	359 (35)	412 (28)	$F(1, 98) = 40.43, p < .001$
	F2	2171 (123)	2069 (140)	$F(1, 98) = 9.893, p < .005$

According to ANOVA, most speakers' controversial [I/E] was significantly different from their non-controversial [i] in both F1 and F2 values; the only exception was 3-vowel female speaker, who made nearly no difference between the two vowels.

Back vowels: [u] vs. [U/O]
Vowel and Gender factors.
Likewise, there were significant effects for both factors Vowel and Gender on F1 and F2 values (For [u]-[U/O] difference: $F(1, 1216) = 176.3$, $p < .01$; For female/male difference: $F(1, 1216) = 97.5$, $p < .01$). Therefore, the non-controversial [u] differ from the controversial [U/O] with respect to F1 and F2 values. This indicates that [U/O] is lower and more fronted than the [u]. As for the factor of Gender, the females' articulation of back vowels is lower and more fronted than the males'.

Within-subject variation.
As far as the within-subject variation is concerned, Table 3 provides the formant values and the ANOVA results for the comparison of the two back vowels [u] and [U/O].

Table 3. Formant values (Mean/SD) of the back vowels

Mean (SD): Hz		[u]	[U/O]	ANOVA
5V-Female	F1	366 (44)	535 (69)	$F(1, 118) = 469.82, p < .001$
	F2	992 (240)	1190 (154)	$F(1, 118) = 24.87, p < .001$
4V-Female	F1	512 (76)	527 (45)	$F(1, 118) = 1.40, p = .239$
	F2	956 (75)	1065 (217)	$F(1, 118) = 5.57, p < .05$
3V-Female	F1	434 (28)	452 (58)	$F(1, 118) = 3.109, p = .08$
	F2	940 (133)	1012 (144)	$F(1, 118) = 6.275, p < .05$
5V-Male	F1	364 (24)	377 (41)	$F(1, 118) = 3.261, p = .073$
	F2	912 (140)	1077 (444)	$F(1, 118) = 7.727, p < .01$
4V-Male	F1	377 (44)	433 (40)	$F(1, 113) = 41.294, p < .001$
	F2	944 (101)	993 (102)	$F(1, 113) = 5.446, p < .05$
3V-Male	F1	368 (38)	409 (30)	$F(1, 118) = 32.365, p < .001$
	F2	889 (113)	900 (127)	$F(1, 118) = .123, p = .727$

In the back vowels' case, most speakers' controversial [U/O] and non-controversial [u] were significantly different mainly in F2 values, but not consistently in F1 values.

The results showed that despite the individual difference in perception, the controversial vowels ([I/E] and [U/O]) were different from their non-controversial vowels ([i] and [u]). Moreover, the factor Gender also contributed to the variations among front and back vowels.

Stress factor
Figure 2 represents the speakers' vowel distribution in the unstressed condition. Although fewer tokens representing the controversial groups [I/E] and [U/O] were available, Figure 3 shows that some speakers separate the two pairs of vowels more clearly than others. The 5-vowel female, 4-vowel female and 4-vowel male made clearer distinctions than did the 5-vowel male, 3-vowel

male and the 3-vowel female.[5] In addition, the male speakers have a more back and contracted vowel space than the female speakers do, which was also found in the stressed condition.[6]

Results of MANOVA showed that for the [i]-[I/E] pair, stress does not have a significant effect on the formant values produced, nor did stress interact with the factors Gender and Vowel Identity. In contrast, the effect of stress on formant production was found to be significant for the [u]-[U/O] pair ($F(1, 1216) = 69.5, p < .01$). We also found significant interaction of Stress with Gender and Vowel Identity for the [u]-[U/O] pair [Gender*Stress: $F(1, 1216) = 4.8, p < .01$; Vowel Identity*Stress: $F(1, 1216) = 39.8, p < .01$]. In sum, the formant values of the [u]-[U/O] pair were found to be influenced by stress, while those of [i]-[I/E] pair were not.

a. 5V-F's vowel distribution
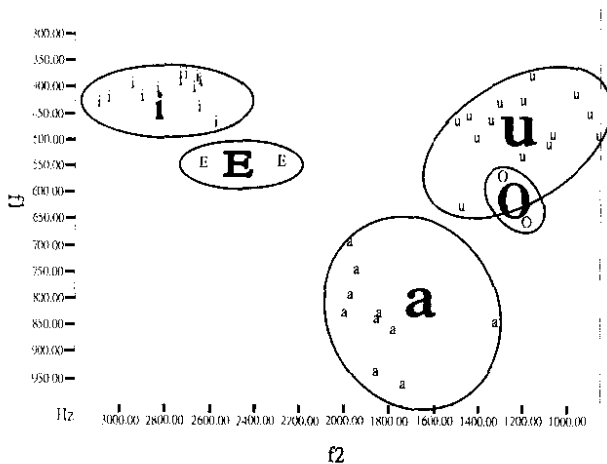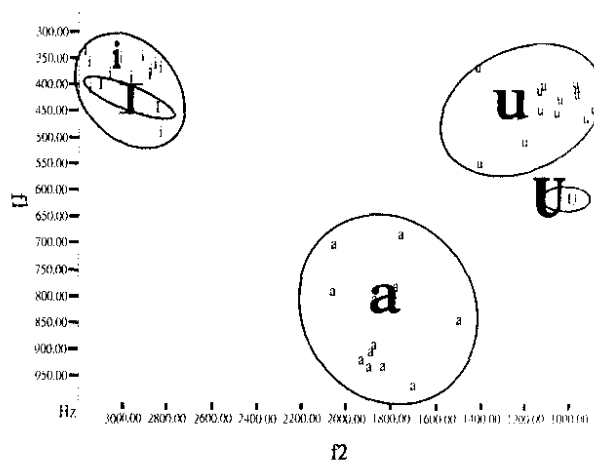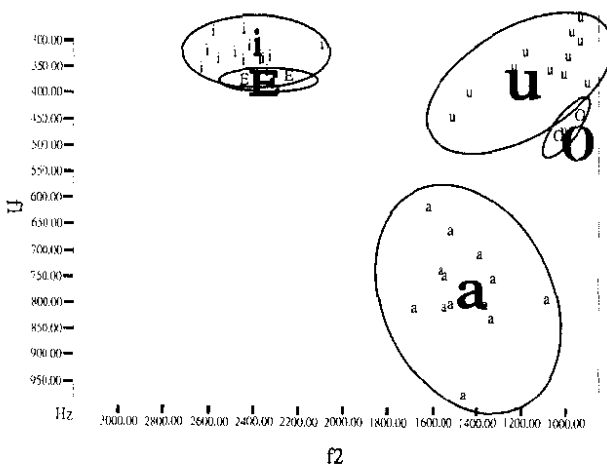
b. 3V-F's vowel distribution

c. 5V-M's vowel distribution

d. 3V-M's vowel distribution

52
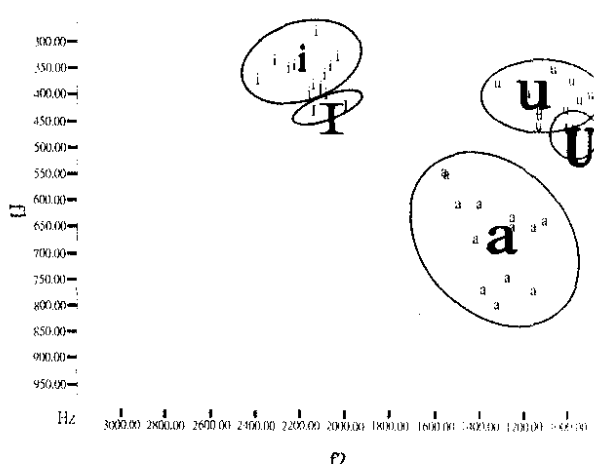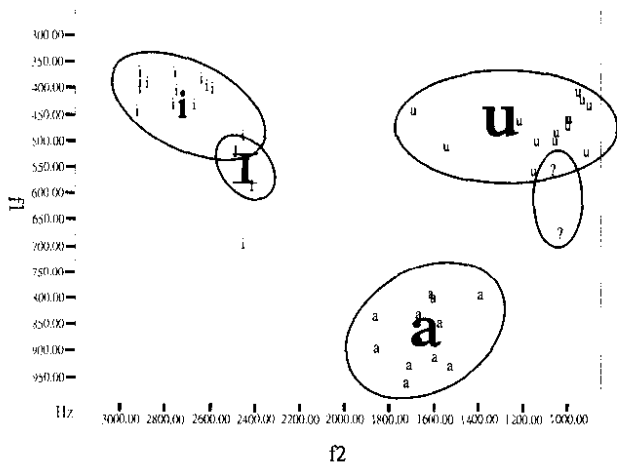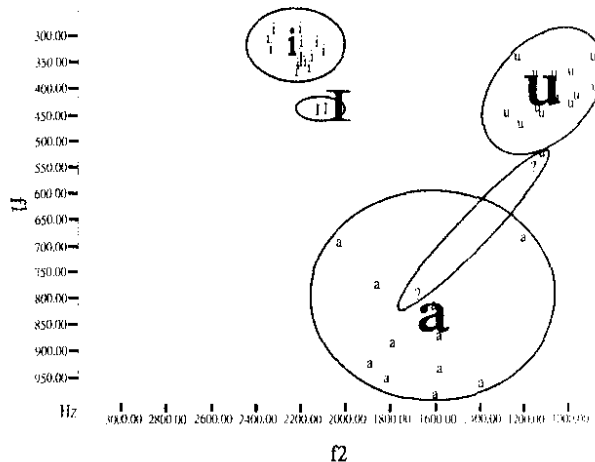
Figure 2.   The six speakers' vowel distributions in the unstressed condition.

Vowel distance and vowel space

Figures 3a-f display each speaker's vowel space in both stressed and unstressed conditions. From these displays, we can observe the following: (1) The vowel space seems slightly larger in the stressed than in the unstressed condition, (2) The distance between the vowels [i] and [I/E], or between [u] and [U/O] is greater in the unstressed than in the stressed condition.

a. 5V-F   **b. 3V-F**

c. 5V-M

f2

d. 3V-M

f2

e. 4V-M

f2

f. 4V-M

f2
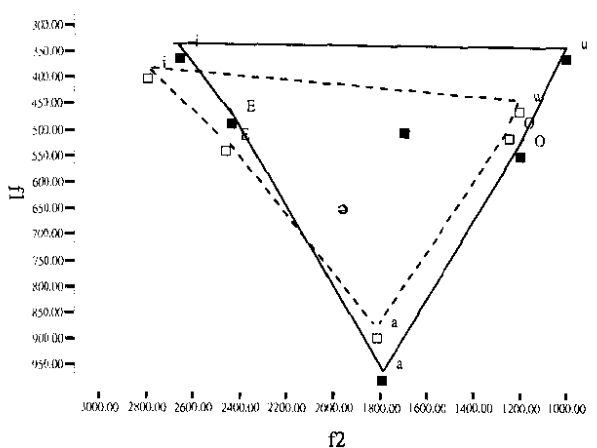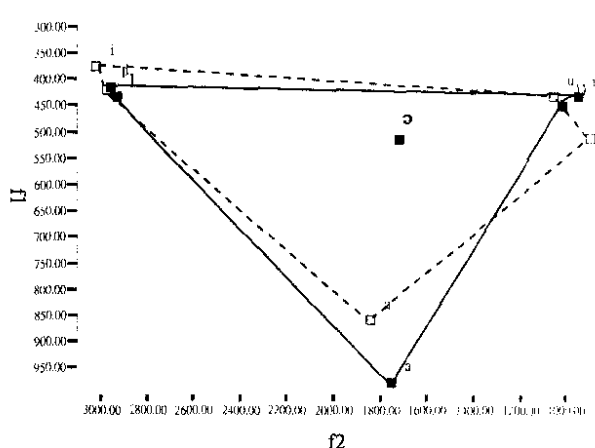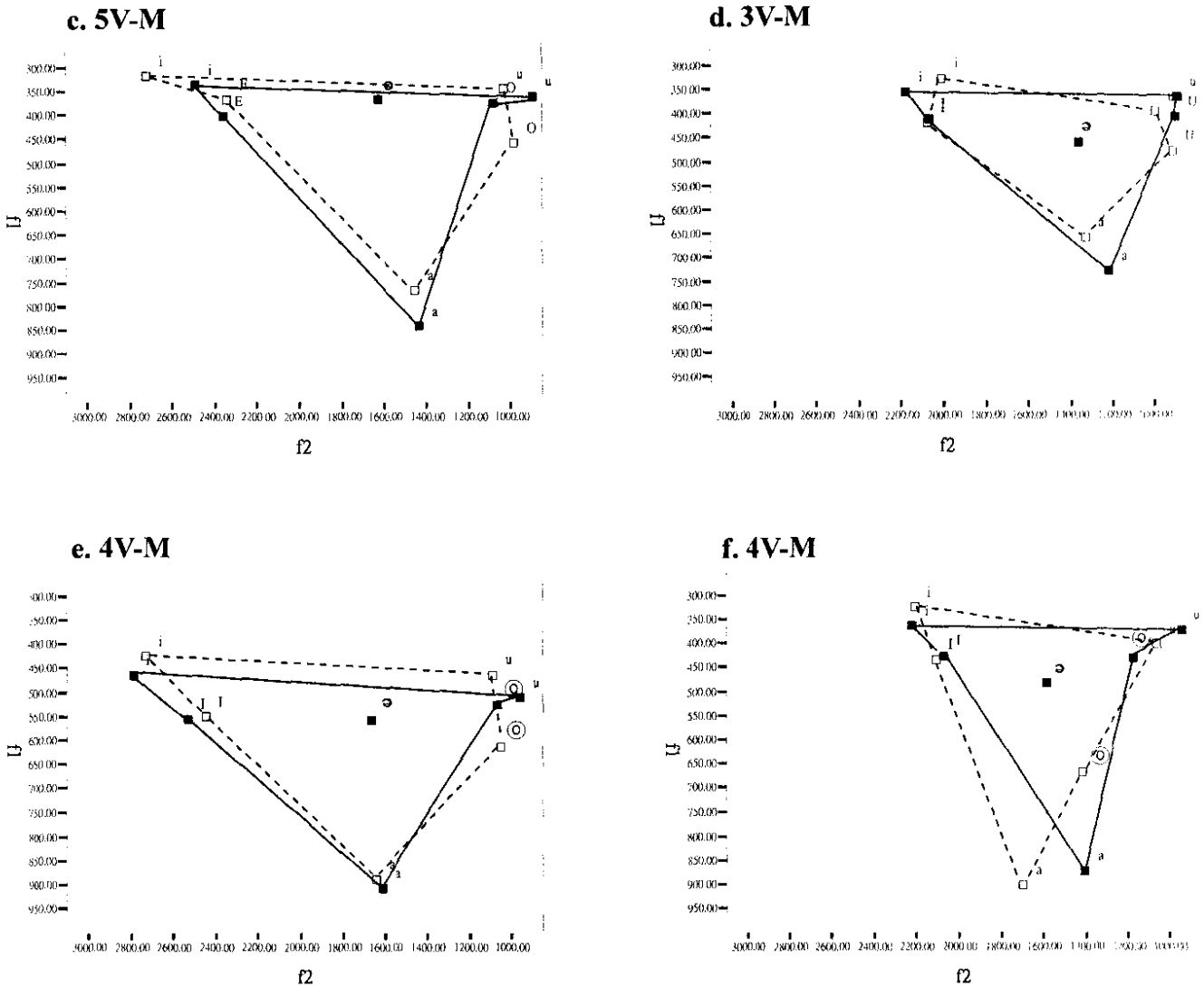
Figure 3. The vowel space of the six speakers. The filled cubes and lines represent the stressed condition, and the hollow cubes and dot lines represent the unstressed condition.

Table 4 and 5 provide further details about the influence of stress on vowel space and vowel distance. The area of the vowel space was calculated by a program[7] designed to divide the entire space into three triangles; the value of the whole space was derived by adding up the areas of those three triangles. Note that the areas of these spaces (in Hz^2) do not have any absolute significance, but can be used in a relative sense to compare one vowel space with another. Table 4 shows that the 5-vowel and 3-vowel groups exhibit a shrinking vowel space in the unstressed condition. In contrast, the two speakers in the 4-vowel group exhibit an expanding vowel space in the unstressed condition.

Table 4.   Stress-related variations in vowel space and the stressed/unstressed vowel space ratio. (Unit: Hz^2)

	Stressed	Unstressed	Shrinkage	Ratio (stressed: unstressed)
5V-F	506303	354415	Shrink	**1: 0.70**
3V-F	544587	451806	Shrink	**1: 0.83**
4V-F	357832	411029	Expand	**1: 1.15**
5V-M	350324	340722	Shrink	**1: 0.97**
3V-M	244194	208259	Shrink	**1: 0.85**
4V-M	262916	304063	Expand	**1: 1.16**

Table 5 displays the vowel distances between [i] and [I/E], and between [u] and [U/O]. The distances were calculated using the Pythagorean Theorem, and vowel distance was derived by a radical expression.

Table 5. Stress-related variation in vowel distance and stressed/unstressed vowel distance ratio

	[i]-[I/E]			[u]-[U/O]		
Subject\distance	Stressed	unstressed	Ratio	Stressed	unstressed	Ratio
5V-F	252.23	358.93	**1: 1.42**	270.77	67.29	**1: 0.25**
3V-F	29.21	62.68	**1: 2.15**	74.58	179.30	**1: 2.40**
4V-F	270.82	306.24	**1: 1.13**	110	155.88	**1: 1.42**
5V-M	146.07	381.79	**1: 2.61**	182.26	122.16	**1: 0.67**
3V-M	115.63	111.26	**1: 0.96**	41.95	114.32	**1: 2.73**
4V-M	159.61	150.44	**1: 0.94**	238.26	444.13	**1: 1.86**

In general, the distance between [i]-[I/E] and between [u]-[U/O] was observed to increase in the unstressed condition. Thus, the two tables indicate an inverse variation between vowel space and vowel distance -- when vowel space shrinks in the unstressed condition, vowel distance increases.

4. General discussion

This section will summarize the three major observations of this study.   First, we found the following asymmetry between perception and production: With the exception of the 3-vowel female speaker, even the speakers who did not perceive [I/E] to be a distinct vowel in their inventory produced a distinction between [i] and [I/E] in terms of F1 value. Second, stress was found to influence the formant values that constitute the [u]-[U/O] distinction. Vowel space shrunk in the unstressed condition, while vowel distance expanded. Finally, subject-based comparisons across gender groups showed that vowel inventory variation was realized along different parameters for male and female speakers. Female speakers tended to rely on vowel formant values to differentiate vowels in production, while male speakers relied more on vowel space.

Vowel dispersion
Despite variation in the number of vowels informants perceived, the

controversial groups [I/E] and [U/O] exhibited significant formant differences from their uncontroversial counterparts [i] and [u], particularly with respect to vowel height. Individual comparisons showed that for female speakers, vowel height accounted for most of the variation in the number of vowels perceived in the stressed condition. In contrast, male speakers tended to use the parameter of vowel space to determine the number of vowels perceived. Although Table 5 shows a vowel distance distinction between 5-vowel and 3-vowel male speakers, this distinction is not as great as that between 5-vowel and 3-vowel female speakers. Instead, variation in the number of vowels perceived depends mostly upon the size of that speaker's vowel space. Table 4 shows that 5-vowel male speaker has larger vowel space than 3-vowel male speaker does, and the size of the 4-vowel male's vowel space is in-between.

We conclude from these observations that vowel dispersion operates on two levels: vowel distance and vowel space. The two levels conflict somewhat with previous accounts of vowel distribution, such as Quantal Theory, since the female speakers' high vowels [i] and [u] do not remain at the same height as vowel inventory size changes. In fact, the height of the high vowels seems to be in direct variation with inventory size. The 3-vowel female speaker lowers her [i] and [u], which contributes to the decrease in contrast between the controversial and uncontroversial high vowels. Our observations on vowel distance distinction by female speakers also violate a basic assumption of Adaptive Dispersion -- the acoustic vowel space does not expand as the number of vowels increases. This suggests that maximal perceptual contrast may not necessarily be accomplished by expanding vowel space and vowel distance simultaneously. Together, the two levels provide a strategy for achieving maximal contrast; but in this case, one level seems to provide sufficient contrast to identify vowel inventory members.

Influence of stress on vowel contrastiveness
It was observed most speakers' vowel spaces shrink when moving from the stressed to unstressed condition, whereas their vowel distances expand. This may be attributed to the insufficient sample size of items containing the controversial vowels [I/E] and [U/O] in unstressed positions. More extensive data collection is necessary to conduct further research the influence of stress on vowel distance.

In addition, stress had no significant effect on vowel duration. It seems that the influence of stress is confined to shrinkage of the vowel space. A possible phonotactic explanation might involve absence of the schwa in last unstressed syllable position. The lack of the schwa obviates the need to form a contrast with [ə], so according to the principle of minimizing articulatory effort, the vowel space would decrease.

5. Conclusion
The findings presented in this paper lead us to propose that the goal of achieving maximal perceptual contrast can be reached in either of two ways: by expanding the vowel space or by increasing the vowel distance. These two levels do not necessarily have to operate in conjunction. The Truku data illustrate both the interaction and the independence of these two levels. Further research is needed, both in terms of more extensive data collection in Truku,

and in the investigation of other Formosan aboriginal languages, to provide support for these initial findings.

Endnotes

1. The vowel inventory discussed here does not include the schwa.
2. Li chose /e/ to represent the fifth vowel, for the purpose of creating symmetry in the vowel inventory.
3. Stress in Truku falls on the penultimate syllable. Chang (2000) described this stress rule as: V→ V/__ (C)V(C)#. This rule can be applied even at the morphological level, since morphological change does not affect the assignment of stress.
4. Yang (1976) and Chang (2000) both indicated a pair of phonological rules in Seediq in which the diphthongs [aw] and [ay] are the ptoto-form of [o] and [e]:
aw → o / __#   ay → e / __#
Our observation showed that Truku preserved these two diphthongs in a consistent way compared with Seediq. This may account for the sparse occurrence of unstressed [e] and [o] word-finally in our wordlist:
(i)   *sinaw* "wine" (Truku) > *sino* "wine" (Seediq)
(ii)   *walay* "thread" (Truku) > *wale* "thread" (Seediq)
5. There is only one sample of the controversial vowel U/O on the 3-vowel female's and 3-vowel male's vowel distributions, because each of them has an outlier which falls outside the range. The 3-vowel female's outlier is 408 Hz F1 and 792 Hz   F2, while the 3-vowel male's outlier is 500 Hz F1 and 826 Hz F2. Both are located far back in the vowel space.
6. Due to the non-occurrence of the schwa /ə/ in unstressed syllables in Truku, these plots do not show schwa distribution.
7. We would like to thank Cheng Chung-ping, who wrote the vowel space measurement program.

References

Aboriginal Languages Cultures Education. (2001). *Textbook: Truku*. Taipei: Council of Indigenous Peoples, Executive Yuan.
Chang, Yung-li. (2000). *Seediq Reference Grammar*. Taipei: Yuan-Liu.
Flemming, Edward. (1995). *Auditory Representations in Phonology*. Ph.D. dissertation, UCLA.
Flemming, Edward. (1996). Evidence for constraints on contrast: The dispersion theory of contrast. Chai-Shune K. Hsu (ed.) *UCLA Working Papers in Phonology* 1, 86-106.
Flemming, Edward. (2001). *Contrast and Perceptual Distinctiveness*. (ms.)
Hu, Chen-Wei. (2003). *Premis on Multimedia Hypertext for Truku Seediq*. M.A. thesis, Graduate Institute of Linguistics, Fu Jen Catholic University.
Hung, Ying-sheng. (1993). *Taiwan early inhabitants' footstep: the culture & legend about ten minority tribes*. Taipei: China Times Publishing Co.
Huang, Chang-hsing. (2000). Dong sai de ke qun de shou lie wen hua (The culture of hunting in Eastern Seediq). *IOE Edited Compilations*. 15:

1-104. Taipei: Institute of Ethnology, Academia Sinica.

Li, Paul Jen-kuei. (1992). *Orthographic Systems for Formosan Languages*. Taipei: Ministry of Education, Republic of China.

Li, Yi-Yuan *et al.* (1963). The Atayal of Nan-ao : an ethnological investigation. *Institute of Ethnology, Academia Sinica, Monograph: 5-6.* Taipei: Institute of Ethnology, Academia Sinica.

Liljencrants, Johan and Lindblom, Björn. (1972). Numerical simulation of vowel quality systems: the role of perceptual contrast. *Language.* 48(4): 839-862.

Lindblom, Björn. (1986). Phonetic universals in vowel systems. J.J. Ohala and J.J. Jaeger (eds.) *Experimental Phonology.* Academic Press.

Lindblom, Björn. (1990). Explaining phonetic variation: a sketch of the H&H theory. In W.J. Hardcastle and A. Marchal (eds), *Speech Production and Speech Modeling.* Dordrecht: Kluwer. 403-39.

Yang, Hsiu-fang. (1976). *Phonological Structure of Paran: A Dialect of Sediq.* Bulletin of the Institute of History and Philology Academia Sinica. 47: 611-706.

Wen-yu Chiang
Graduate Institute of Linguistics, National Taiwan University
No.1, Sec. 4, Roosevelt Road, Taipei, 106
Taiwan
wychiang@ntu.edu.tw

Fang-mei Chiang
Graduate Institute of Linguistics, National Taiwan University
No.1, Sec. 4, Roosevelt Road, Taipei, 106
Taiwan
r90142010@ntu.edu.tw

Prosodic Realization of Negation in Saisiyat and English

Wen-yu Chiang & I Chang-Liao
National Taiwan University

1. Introduction

Cross-linguistically, negation may be realized by means of syntactic marking, prosodic marking or a combination of the two, depending on the prosodic characteristics of the language in question. Syntactically, a language can use a single negator to express negation; such as *not* in English, or use various negators for different syntactic structures, such as those in Austronesian languages of Taiwan. Negation can also be marked prosodically. For example, Yaeger-Dror (2002) found that the F_0 (fundamental frequency) of English *not* and French *pas* are higher than surrounding words. The current literature suggests that such prosodic differences accompanying the syntactic marking of negation are relatively common; the use of only prosodic marking to realize negation, in contrast, is relatively rare.

The question of whether a negator is invariably acoustically more prominent than its surrounding words, however, remains unresolved. Semantically, negators bring new information to a sentence; for this reason, they assume 'focal prominence'. Yaeger-Dror (2002) proposes the 'Cognitive Prominence Principle', according to which cognitively prominent items, such as negators, should be prosodically marked. According to this principle, acoustic prominence enhances discourse participants' attention to focused items, which maximizes the effectiveness of communication.

Acoustic evidence for the 'Cognitive Prominence Principle' was found in Allen (1973). In this study, participants were required to utter a set of negative sentences. Measurement of f0 values within those sentences determined that the pitch of negators was usually higher than that of contiguous lexical items. According to Yager-Dror (1985, 2002), the unstable behavior of negators stems from the conflict between the 'Cognitive Prominence Principle' and what they have termed the 'Social Agreement Principle'. On the one hand, a negator, due to its sentential prominence, must be prosodically marked in order to comply with the Cognitive Prominence Principle. On the other hand, the Social Agreement Principle discourages emphasis on any objection to a conversation partner's previous assertion, which would effectively eliminate any prosodic prominence assigned to negators. Yager-Dror's research suggests that the choice to assign prosodic prominence to negators may be sensitive to the dynamics of culture and discourse.

Whether the prosodic prominence on negators found in English negative sentences can be found across a range of language types has yet to be explored.

The current study focuses on the prosodic realization of negators in Saisiyat[1], an endangered aboriginal language of Taiwan, and compares its prosodic realization of negation with that of English. The results of this study indicate that sentential subjects are the most acoustically prominent items in the Saisiyat negative sentences measured. This contrasts sharply with the English experimental sentences, in which the negator itself was the most acoustically prominent item. These findings suggest that Saisiyat is a pitch-accent language; thus, the presence of negators does not significantly change the prosodic

parameters of surrounding words. English, in contrast, is an intonation language, so the presence of negation results in substantial prosodic modification. This suggests that the phenomenon of negation is universally prominent; however, languages with different prosodic systems will adopt different strategies for realizing prominence.

2. The structure of Saisiyat

This section will provide a brief introduction to the phonology, morphology and syntax of Saisiyat. This information was collected from three sources: Yeh (2000), our own fieldwork and a database of Saisiyat discourse collected at the Graduate Institute of Linguistics of National Taiwan University.

2.1 The phonemic inventory of Saisiyat

The phonemic inventory of Saisiyat consists of seventeen consonants and six vowels; these appear in the tables below:

Table 1 Consonants in Saisiyat[2]

		Bilabial	Alveolar	Palatal	Velar	Glottal
Stop		P	T		K	ʔ (')
Nasal		M	N		ŋ (ng)	
Fricative	Voiceless		S	ʃ (S)		h
	Voiced		Z			
Lateral			L	ʟ (L)		
Thrill				R		
Glide		W		j (y)		

Table 2 Vowels in Saisiyat

	Front	Central	Back
High	I		
Mid	œ (oe)	ə (e)	o
Low	æ (ae)	A	

Pitch accent within words usually falls on the last syllable, except for function words and place names. The most prevalent syllable structure is CV and CVC (Chiang and Chiang, 2004).

2.2 Saisiyat case marking

In contrast with other Austronesian languages of Taiwan, which exhibit VOS word order, Saisiyat has SVO word order. Case markers usually occur before nouns to mark their syntactic function. According to Yeh (2000), there are six case markers in Saisiyat, each of which is divided into two categories: persons (not including pronouns) and common nouns, as shown in Table 3:

Table 3: The Case System of Saisiyat (Yeh, 2000)

	Nominative case	Accusative case	Genitive case	Possessive case	Dative case	Locative case
Persons	¢ hi	Hi	Ni	ʔan-a	ʔini	Kan Kala
Common nouns	¢ ka	Ka	noka no	ʔan noka-a	No	Ray

2.3 Saisiyat focus markers

Like other Austronesian languages, Saisiyat uses focus markers to highlight the focused constituent in a sentence. Yeh (2000) divided these into two groups and four different kinds of focus markers within those groups, classified according to the relationship between the verb and the semantic role of the arguments. Focus is morphologically affixed to the verb, and any constituent can receive Focus: an agent, a patient or a location, among other arguments. This is illustrated in Table 4:

Table 4: The Focus Maker System In Saisiyat (Yeh, 2000)

Focus markers	Group I	Group II
Agent focus (AF)	m-, -om-, ma-, φ	ψ
Patient focus (PF)	-ən	-l
Locative focus (LF)	-an	
Referential focus (RF)	si-	-ani

The choice between Group I and Group II focus markers is determined by negator type, which will be explained in section 2.4. Examples (1) through (4) illustrate the use of different types of focus markers in positive sentences.

<Example of AF>

(1) ¢   ʔojaʔ ʃ-om-βət ka korkoriŋ
     Nom mother AF-beat Acc child
     "Mother beats the child" (data from our field work)

<Example of PF>

(2) korkoriŋ noka ʔojaʔ ʃəβət-en
     child Gen mother beat-PF
     "The child is beaten by mother." (data from our field work)

<Example of LF>

(3) ahœ'   k-om-ita ka takəm   ray katis-ralom-an
     dog see-AF   Acc frog Loc KATIS-water-LF
"The dog was looking at a frog inside a container." (data from Frog Story 7[3], IU 1-3)

<Example of RF>

(4) kahœj   si-səβət ni βakiʔ   ka korkoriŋ
     stick   RF-beat Gen grandfather   Acc child
     "The grandfather uses a stick to beat the child" (data from Yeh, 2000)

In (1), adding the AF infix -*om*- to the verb *ʃəβət* 'beat' marks focus on the agent *ʔoja ʔ* 'mother', which creates a sentence similar in meaning to the English active voice, hence 'Mother beats the child'. When the verb takes the PF suffix *-ə* in (2), the patient becomes the focus of the sentence, which creates a meaning similar to the English passive voice, hence 'The child is beaten by mother'. In (3), the LF suffix *-an* is added to the noun *ralom* 'water', creating the inflected word *katisraloman* 'container', which is perceived as sentential focus on the location of the water. In (4), the use of the RF (referential) suffix places narrow focus the instrument *kahœj* 'stick'.

The range of pragmatic uses for focus markers in Saisiyat is actually much more complex than the examples illustrated above would suggest, and an exhaustive discussion of them is beyond the scope of this paper. Thus, we limit our current discussion of focus to the hypothesis that Saisiyat's ability to mark focus syntactically allows it to minimize reliance on prosodic cues for marking focal prominence. Since focus is marked using inflectional morphology, the word undergoes minimal changes from the lexical to the postlexical level in Saisiyat. Thus, the intonational pattern of Saisiyat remains relatively stable and predictable, like that of other pitch accent languages such as Japanese. In addition, since negators marked even more specifically by being matched with one of eight different focus markers in Saisiyat, we expect the negator and its surrounding words to undergo fewer prosodic changes than those in English do.

2.4. Negators in Saisiyat
Saisiyat has an inventory of eight negators: *ʔokik, ʔokaj, ʔamkik, ʔamkaj, kayni ʔ, ʔoka ʔ, ʔizi ʔ* and *ʔin ʔini ʔ* (Yeh, 1998, 2000). Negators are chosen from this inventory according to the syntactic constructions in which they appear, and they are followed by verbs with focus markers chosen according to sentential focus. Table 5 provides glosses, distributions and co-occurrence restrictions of negators in Saisiyat.

Table 5: Negators in Saisiyat[4]

Negator	Translation	Construction	Co-occurrence restrictions	Focus marker
ʔokik	Not		Nouns [+stative]verbs verbs with aspect markers	
ʔokaj	Not		[-stative]verbs	Group II
ʔamkik	will not	The same as *ʔokik*		
ʔamkaj	will not	The same as *ʔokay*		
Kajni ʔ	not want (to)	Volitional	nouns verbs	Group I
ʔoka ʔ	There is not	Existential	Nouns	
ʔizi ʔ	Don't...	Imperative	Verbs	Group I
ʔin ʔini ʔ	not yet		verbs adverbs	Group II

3. Experiment 1: Saisiyat expression of prosodic prominence in negative sentences

First, we investigated the relative prosodic prominence of negators in Saisiyat and compared them to those occurring in English negative sentences. Allen's (1973) investigation of English negative sentences demonstrated that negative modals are prosodically prominent, specifically by being produced with raised pitch. The current experiment tested this observation on Saisiyat negative sentences, as well as on a set of English sentences, in order to confirm Allen's findings.

3.1.1   Methodology

The Saisiyat materials designed for this experiment consisted of 15 negative sentences, an example of which is given in (5). [5]

<Saisiyat sentence >

(5) ʔœβaj ʔokik minatiniʔ
     ʔœβaj Neg brother
     "ʔœβaj is not a brother"

Each negative sentence contains a single subject, negator and a predicate. The predicate consists of either a noun or a stative/dynamic verb, depending on the negator chosen. The grammatical subject of each sentence was *ʔœβaj*, a common male name. All types of negator were included in the sentences except for *ʔin ʔini ʔ* 'not yet', because its relatively complex syntactic structure would have made the sentence in which it appeared differ substantially from the other experimental items. Words appearing after negators were controlled for segmental content; these were limited to words composed exclusively of sonorants and vowels, so that an uninterrupted pitch track could be extracted from that area. Our Northern Saisiyat informant confirmed that all experimental sentences were grammatical and acceptable. English sentences used in the study were direct translations of the Saisiyat sentences. In all cases, the sentential subject was the English male name "Bob". An example is given in (6):

<English sentences>

(6) **Bob is not my brother.**

Three informants participated in the experiment: two native Saisiyat informants and one native British English informant. One of the Saisiyat informants speaks Northern Saisiyat and the other speaks Southern Saisiyat[6]. Both are male, and between 50-60 years of age. They also speak Japanese and Hakka, a Chinese dialect spoken in Taiwan. The British English-speaking informant is a 24-year-old male. By self-report, none of the informants had a problem related to either hearing or articulation.

Recordings were made in the speech lab at the Graduate Institute of Linguistics at National Taiwan University, using a Kay Elemetrics CSL 4400. A condenser microphone was placed approximately 10 centimeters away from the informants' mouths for the duration of the recording. Total recording time for each participant was approximately one hour.

Saisiyat is a language without a writing system, so it was not possible to

elicit the negative sentences using written materials. Instead, informants were asked questions by the researcher, and instructed to answer each of the questions in the negative, an example of which is given in (7).

<Elicitation of negative sentences>

(7) Researcher: **ʔœβaj ʔokik minatiniʔ ay?**
(Is not ʔœβaj a brother?)

Informant: **ʔœβaj ʔokik minatiniʔ.**
(ʔœβaj is not a brother.)

After all recordings had been completed, Praat 4.1.19 signal processing software was used to measure the pitch, amplitude and duration of the vowels and coda consonants appearing in stressed syllables which, in the Saisiyat materials, were invariably final syllables. Syllable onsets were excluded from these measurements to avoid the confound of microprosodic variation that would have been introduced by different initial consonants. For the same reason, the accusative marker *ka*, which occurs in existential sentences, was also excluded from measurement.

3.1.2 Results

Table 6 below summarizes comparison among the informants of Northern Saisiyat, Southern Saisiyat, and English with respect to the parameters of both pitch and intensity peaks. The letter X represents the word following the negator in any given sentence.

Table 6 Differences between Subject, negator, and X with respect to pitch peak and intensity peak for three informants

			Subject	Negator	X	ANOVA
Northern Saisiyat	Pitch Peak	Mean	134.89	127.1	121.29	$F_{(2, 42)}=7.27$, $p<.01$
		SD	8.17	12.29	8.39	
	Intensity Peak	Mean	78.43	75.35	75.12	$F_{(2, 42)}=9.986$,
		SD	2.69	2.34	1.64	$p<.01$
Southern Saisiyat	Pitch Peak	Mean	115.73	107.21	100.43	$F_{(2, 42)}=12.772$,
		SD	7.08	7.23	10.23	$p<.01$
	Intensity Peak	Mean	76.4	72.35	70.83	$F_{(2, 42)}=7.551$,
		SD	3.61	4.44	4.09	$p<.01$
English	Pitch Peak	Mean	130.77	142.73	110.69	$F_{(2, 42)}=122.816$,
		SD	4.67	5.17	6.89	$p<.01$
	Intensity Peak	Mean	77.24	72.63	68.5	$F_{(2,42)}=46.147$, $p<.01$

Table 6 shows that among subject, negator and X, both pitch peak and intensity peak (taken to be indicators of prominence) are realized on the sentential subject for the Northern Saisiyat informant. The mean difference between item categories was found to be significant in a one-way ANOVA ($F_{(2,42)}=7.27$, $p<.01$ for pitch peak, and $F_{(2,42)}=9.986$, $p<.01$ for intensity peak). A post-hoc

test indicated that the difference between ?æβaj and X contributes to the contrast in pitch peak mean, while the contrast in intensity peak mean is created by ?æβaj being significantly higher than both ?æβaj and X.

The Southern Saisiyat informant's data displays patterns similar to those of the Northern Saisiyat informant; both pitch and intensity peaks are realized on the sentential subject (F(2,42)=12.772, p<.01 for pitch peak and F(2,42)=7.551, p<.01 for intensity peak). A post-hoc test revealed that the significant difference emerges as the result of differences between item types, i.e. the subject and negator, and between the subject and X.

The English informant's data with respect to intensity peak does not differ significantly from Saisiyat informants' data; the highest mean of peak intensity occurs at the sentential subject (F(2,42)=46.147, p<.01). However, the English informant's data does differ with respect to peak f0. In the English data, the highest pitch peaks were realized on negators (F(2,42)=122.816, p<.01). This result confirms Allen (1973), in which pitch peaks were realized most often on negators in English. A post-hoc test showed that the mean differences of pitch and intensity peak among the three item types are all significant.

3.2 Experiment 2: –aj sentences

The results of Experiment 1 suggest that sentential subjects represent the most acoustically prominent constituents in Saisiyat negative sentences. To investigate the possibility that microprosodic effects of the 'aj' coda in the baj subject have skewed the results of Experiment 1, another 15 pairs of sentences were constructed, each word of which has an –ay coda, an example of which appears in (6) [7]. If we obtain similar results holding vowel quality constant across item types, it diminishes the likelihood that /aj/ is simply intrinsically higher in pitch and/or amplitude than other vowels.

3.2.1 Results

As can be seen in table 7, ?æβaj still remains the most prominent item, even when the last three items have all been controlled for vowel quality.

Table 7: Differences between subject, negator, and X with respect to pitch peak and intensity peak for the three informants (three item codas controlled as –ay)

			Subject	Negator	X	ANOVA
Northern Saisiyat	Pitch Peak	Mean	134.99	118.16	111.2	F(2, 42)=25.496,
		SD	9.57	11.95	5.45	p<.01
	Intensity Peak	Mean	65.94	63.34	63.05	F(2, 42)=12.185,
		SD	1.27	1.39	2.41	p<.01
Southern Saisiyat	Pitch Peak	Mean	136.93	121.49	122.49	F(2, 42)=23.084,
		SD	4.71	7.2	8.45	P<.01
	Intensity Peak	Mean	71.69	67.95	69.37	F(2, 42)=15.812,
		SD	1.51	1.46	2.4	p<.01

A post-hoc test showed that the significance of both pitch and intensity peaks stems from the mean difference between ?æβaj and ?okaj, ?æβaj and X-ay. That is, the pitch and intensity peak of ?æβaj is significantly higher than ?okaj and X-ay.

The Southern Saisiyat informant's data are similar to those of the Northern Saisiyat informant. The most prominent item is still the subject *ʔæβaj* in terms of both pitch and intensity peaks. In addition, a post-hoc test revealed that the subject *ʔæβaj* is significantly higher than both *ʔokaj* and X-ay.

3.3 Experiment 3: Sentences starting with *ʔalʔalak* 'young person'

The results of Experiments 1 and 2 support our conclusion that the subject is the most acoustically prominent item in Saisiyat negative sentences. To strengthen this conclusion, we constructed another set of sentences, the aim of which was to decrease the acoustic prominence of the sentential subject by using the word *ʔalʔalak* 'young person', which ends with a voiceless stop consonant 'k', rather than any open syllable. If we obtain similar results after deliberately decreasing the acoustic prominence of the subject by choosing a syllable composed of segments that are intrinsically lower in f0 and amplitude, then our conclusion will be reinforced. In Experiment 3, the negator and "X" consisted of the same items that appeared in Experiment 2. An example is given below in (8).

(8) ʔalʔalak   ʔokaj   βœaj

young_person   Neg   give
'The young person does not give'

3.3.1 Results

Table 8 shows that there is no significant difference from the previous experiments in terms of pitch peak location. Subject *ʔalʔalak* was shown to be the location of pitch peak in a one-way ANOVA.

Table 8: Differences between Subject (with ak coda), negator, and X with respect to pitch peak and intensity peak for the three informants

			Subject	Negator	X	ANOVA
Northern Saisiyat	Pitch Peak	Mean	124.47	108.01	100.92	F(2, 42)=20.25,
		SD	8.15	7.41	7.08	p<.01
	Intensity Peak	Mean	62.31	61.32	61.81	F(2, 42)=1.074,
		SD	1.76	1.68	2.09	p>.01
Southern Saisiyat	Pitch Peak	Mean	119.39	114.29	110.93	F(2, 42)=6.125,
		SD	7.05	5.89	7	p<.01
	Intensity Peak	Mean	57.09	60.04	59.6	F(2, 42)=.491, p>.01
		SD	14.81	1.32	3.36	

However, in these sentences, the subject is no longer the location of peak intensity, and there is no significant difference among the three item types (F(2,42)=.491, p>.01). A post-hoc test showed that the mean pitch peak of *ʔalʔalak* is significantly higher than those of *ʔokaj* and X-ay. It should be noted that in this experiment, the two Saisiyat speakers both realized pitch peak on the sentential subject, but differed in their placement of the intensity peak. The Southern Saisiyat informant realized the intensity peak on the negator, while the Northern speaker's intensity peak remained on the sentential subject. Results of a one-way ANOVA indicate that the three item types are significantly different with respect to mean pitch peak. In contrast, there is no significant difference

among the mean intensity peaks of the three items. The mean intensity peak of the subject *ʔalʔalak* is even a little lower than other two items, which is quite different from the results from previous experiments.

From the analysis of the post-hoc test, the mean pitch peak is significantly higher than that of X-ay. There is no significant difference between the mean pitch peaks of *ʔalʔalak* and *ʔokaj*.

3.4 Summary and conclustion

The English informant's data from Experiment 1 confirm Allen (1973), in the sense that negators were the most prominent items (as measured by peak f0 and intensity) in English negative sentences. However, this phenomenon is absent in Saisiyat negative sentences; the most prominent item in these sentences are the sentential subjects.

To further test the acoustic prominence of Saisiyat sentential subjects, we designed Experiment 2, in which the coda of each item was held constant as /aj/; still, sentential subjects remained the most prominent item. Experiment 3 decreased the sonority of the subject by changing it to a closed syllable with a voiceless stop coda. Under those conditions, sentential subjects remained the locus of pitch peak, but not uniformly of peak intensity. The results of Experiments 1, 2 and 3 provide strong evidence for the claim that the most acoustically prominent item in Saisiyat negative sentences is the sentential subject.

In addition, a disjunction between mean pitch peak and intensity peak can be seen in the English data in Experiment 1 and in the Saisiyat data in Experiment 3. In Experiment 1, the mean pitch peak falls on English negators, but the mean intensity peak falls on sentential subjects. In Experiment 3, we can see that the subject remains the locus of the mean pitch peak, but not peak intensity.

Since the post-hoc test in Experiment 1 showed that there are significant differences among all three item types in English negative sentences, we suggest that there is a principled difference in prominence ranking orders between English and Saisyat negative sentences. The ranking order for English would be negator> subject> X. and for Saisiyat, subject > negator > X.

The different prosodic characteristics of Saisiyat and English suggest that a variety of intonational strategies is available for the linguistic realization of negation. For example, negators are cognitively prominent, but they are not necessarily the most acoustically prominent items in a negative sentence. Thus, the Cognitive Prominence Principle can be applied to English negators but not to those in Saisiyat. In fact, Saisiyat informants realize more acoustic prominence on sentential subjects than on any other items in a negative sentence. Moreover, in English negative sentences, various prosodic changes are realized on the "X" constituent, which was not found in the data of Saisiyat informants.

In interpreting our results, it must be noted that unlike Yager-Dror (1985, 2002), the current study does not take pragmatic factors into consideration; the materials used in this study are designed sentences rather than natural discourses, and the recording procedure is strictly controlled.

Finally, the question of why Saisiyat realizes acoustic prominence on sentential subjects rather than negators is not yet clear. A possible explanation is that the prosodic focus in Saisiyat sentences is positionally determined; i.e. the focus falls on the agent as a default strategy. It must also be noted that the

materials in this study were restricted to agent-focused sentences, so it has yet to be determined whether focus in different positions would receive the same sentential prominence. Further studies are needed to test if agents are always the intonational focus in Saisiyat.

This paper provides a typological comparison of the prosodic realization of negation in Saisiyat and English and provides preliminary evidence that Saisiyat is best classified as a pitch-accent language. Future study should examine the cue-trading relationships of morphosyntax and prosody in the realization of negation as it occurs in natural discourse, across a wide variety of language types.

Endnotes

1. Saisiyat is an endangered language in Taiwan with about 7000 speakers. It belongs to the Austronesian language family and has two dialectal variations—North Saisiyat (Daai dialect) and South Saisiyat (Tungho dialect). North Saisiyat is spoken in Miao-Li and South Saisiyat is spoken in Xin-Zhu. The speaker population of these two dialects of Saisiyat lives in central Taiwan.
2. The alphabets in parentheses represent the ASCII equivalent of the original IPA form, which is the transcription system that appears in Yeh (2000). For consistency, all the Saisiyat words appearing subsequently in our study will be transcribed in standard IPA.
3. Frog Story is one of the texts collected in a research project of Saisiyat in National Taiwan University. Project No.: NSC 91-2411-H-002-088
4. The materials used in this study include all negators in the table above, except for *ʔamkik* and *ʔamkay*. These two are actually negators blended with future tense marker *ʔam* (Yeh, 2000)
5. A complete list of the experimental sentences appears in Appendix A
6. Since Saisiyat is an endangered language, informants who are both proficient enough to participate in the experiment and fluent enough in Mandarin Chinese to comprehend the instructions are very difficult to find.
7. A complete list of sentences is given in Appendix B.

Appendix A: Designed Saisiyat and English sentences used in experiment 1

No.	Saisiyat	English
1	'oebay minatini'	Bob is my brother
	'oebay 'okik minatini'	Bob is not my brother
2	'oebay lalaor	Bob dozes
	'oebay 'okik lalaor	Bob does not doze
3	'oebay liyabo'	Bob is rich
	'oebay 'okik liyabo'	Bob is not rich
4	'oebay m-atol	Bob sings
	'oebay haSa' m-atol	Bob does not sing
5	'oebay miltamako'	Bob hunts
	'oebay haSa' miltamako'	Bob does not hunt
6	'oebay laloehay	Bob has fun
	'oebay haSa' laloehay	Bob does not have fun
7	'oebay mayna:a'	Bob waits
	'oebay 'okay ayna:a'	Bob does not wait
8	'oebay mwa:i'	Bob comes to my place
	'oebay 'okay wa:i'	Bob does not come to my place
9	'oebay miltamako'	Bob smokes
	'oebay okay miltamako'	Bob does not smoke
10	'oebay hayza' ka lapowar	Bob has a guava
	'oebay 'oka' ka lapowar	Bob does not have a guava
11	'oebay hayza' ka laro'	Bob has a persimmon
	'oebay 'oka' ka laro'	Bob does not have a persimmon
12	'oebay hayza' ka mona:	Bob has a snail
	'oebay 'oka' ka mona:	Bob does not have a snail
13	'oebay 'am lapowar	Bob wants a guava
	'oebay kayni' lapowar	Bob does not want a guava
14	'oebay 'am laro'	Bob wants a persimmon
	'oebay kayni' laro'	Bob does not want a persimmon
15	'oebay 'am mona:	Bob wants a snail
	'oebay kayni' mona:	Bob does not want a snail

Appendix B: Designed Saisiyat sentences used in experiment 2

No.	Saisiyat
1	'oebay m-obay
	'oebay 'okay be:ay
2	'oebay ma-sangay
	'oebay 'okay sangay
3	'oebay romamramay
	'oebay 'okay ramramay
4	'oebay so-mo-way
	'oebay 'okay soway
5	'oebay kipazay
	'oebay 'okay kipazay
6	'oebay kiboway
	'oebay 'okay kiboway
7	'oebay 'aelipoway
	'oebay 'okay 'aelipoway
8	'oebay inkonkonay
	'oebay 'okay inkonkonay
9	'oebay rasiwazay
	'oebay 'okay rasiwazay
10	'oebay tisko-aewhay
	'oebay 'okay tisko-aewhay
11	'oebay masay
	'oebay 'okay pasay
12	'oebay omas'asay
	'oebay 'okay 'as'asay
13	'oebay maStalay
	'oebay 'okay 'iStalay
14	'oebay min-balbalay
	'oebay 'okay balbalay
15	'oebay kakoway
	'oebay 'okay kakoway

Appendix C: Designed Saisiyat sentences used in experiment 3

No.	Saisiyat
1	'al'alak m-obay
	'al'alak 'okay beːay
2	'al'alak ma-sangay
	'al'alak 'okay sangay
3	'al'alak romamramay
	'al'alak 'okay ramramay
4	'al'alak so-mo-way
	'al'alak 'okay soway
5	'al'alak kipazay
	'al'alak 'okay kipazay
6	'al'alak kiboway
	'al'alak 'okay kiboway
7	'al'alak 'aelipoway
	'al'alak 'okay 'aelipoway
8	'al'alak inkonkonay
	'al'alak 'okay inkonkonay
9	'al'alak rasiwazay
	'al'alak 'okay rasiwazay
10	'al'alak tisko-aewhay
	'al'alak 'okay tisko-aewhay
11	'al'alak masay
	'al'alak 'okay pasay
12	'al'alak omas'asay
	'al'alak 'okay 'as'asay
13	'al'alak maStalay
	'al'alak 'okay 'iStalay
14	'al'alak min-balbalay
	'al'alak 'okay balbalay
15	'al'alak kakoway
	'al'alak 'okay kakoway

Reference

Allen, J. (1973). Pitch peaks on negatives. *Journal of the Acoustical Society of America*, 55, S77 (Abstract)

Chiang, Wen-yu, & Chiang, Fang-mei. (2004) *The lexicon-level pitch accent in Saisiayt*. Unpublished manuscript

Cruttenden, A. (1997). *Intonation* (2nd ed.). Cambridge: Cambridge University Press.

Yaeger-Dror, M. (2002). Register and prosodic variation, a cross language comparison. *Journal of Pragmatics*, 34, 1495-1536.

Yaeger-Dror, M.(1985). Intonational prominence on negatives in English. *Language and Speech*, 28, 197-230.

Yeh, Marie M, Lilian M. Huang, Elizabeth Zeitoun, Anna H. Chang, Joy J. Wu. (1998).A Preliminary Study On Negative Constructions In Some Formosan Languages. *Selected Papers From The Second International Symposium On Languages In Taiwan*.(ISOLIT-II), 81~111

Yeh, Marie M 葉美利, (2000). *sai siya yu can kau yu fa [A Reference Grammar of Saisiyat]*. Taipei: Yuan-Lio.

Wen-yu Chiang
Graduate Institute of Linguistics, National Taiwan University
No.1, Sec. 4, Roosevelt Road, Taipei, 106
Taiwan
wychiang@ntu.edu.tw

I Chang-Liao
Graduate Institute of Linguistics, National Taiwan University
No.1, Sec. 4, Roosevelt Road, Taipei, 106
Taiwan
r91142007@ntu.edu.tw

A Voice Opposition without Voice Morphology

Mark Donohue
National University of Singapore

1. Voice systems in Austronesian languages

The symmetrical voice systems of the western Austronesian languages are well known; Schachter (1976, 1977) provides a succinct summary of the issues, and numerous works before and since have argued variously for passive and ergative (and other) interpretations of the data. It is widely supposed that the Central-eastern Malayo-Polynesian languages lack voice oppositions: such mechanisms are widely reported to be absent in Papuan languages, and are generally assumed to be absent from the Austronesian languages of the area.[1] The complex voice systems of the western Austronesian languages are present in only relic form (see Wouk and Ross 2002), and the Oceanic passives (*ta- and *ma-) are only weakly attested (Donohue 2004).

Another pattern of voice marking is found in many eastern Indonesian and western Melanesian languages, separated from the more well-recognised western languages that have overt, dedicated voice marking morphology, and separate from the well-described passive and passive-like morphemes (*ta or *ma) that are found in the Pacific. Tukang Besi, the south-easternmost Western Malayo-Polynesian language, is one such language. Like other Muna-Buton languages, Tukang Besi has pronominal agreement on the verb to index the core arguments of the clause. The morphological template for the pronominal inflection on the verb is shown in (1) (the same template holds for other pronominal voice languages):[2]

Morphological template for verbal inflection
(1)   Verb = Agreement $_{S,A}$ - Verb root - (Agreement $_P$)

While agreement for the S,A argument is obligatory agreement for the P is not. The conditions that determine its appearance are complex, and shall be mentioned here only as far as they concern the presentation of this form of variation as part of the grammatical voice system. The bulk of this paper deals with an analysis of the voice system of Tukang Besi, which, has both a complex verbal agreement system as well as the last fully developed (and obligatory) case marking system among Austronesian languages with an increasingly head-marking trend to the east (case marking of core constituents only becomes functional again in Vanuatu and the Solomons, and is well-developed in Polynesia). For that reason, as well as personal acquaintance with the language, it is a sensible starting point.

2. Tukang Besi voice: multifunctional morphology

The voice system in Tukang Besi is marked only by variation in the verbal agreement system, as described in (1), not by the presence of any distinct voice-marking morphology. In Tukang Besi agreement for the S,A argument is obligatory, while agreement for a P is by enclitic, and is for most lexemes

73

generally optional. When it occurs there are further morphosyntactic changes in the sentence, affecting word order possibilities, case marking coding, and most relevantly grammatical function assignment. What does NOT change is the identity of the argument indexed by the agreement prefix, indicating that the voice system is not one that involves demotion of one argument in order to promote the other, such as familiar passive or antipassive systems from languages like English or the Mayan languages, but is more similar to the Austronesian voice systems from the west, as might be expected. This also indicates that when the voice changes the categories that determine which agreement markers are used to index an argument do not, showing that 'subject' and 'object' are not appropriate labels to apply to the prefixes and enclitics.

2.1. The voice system: basics and argumentation

In both voices the verb is prefixed to index the values of the S or A, and in the non-active voice there is an enclitic which bears the pronominal features of the P. When the P is indexed no fixed order holds between the DPs representing the postverbal A and P arguments, while a clause that lacks P agreement is only grammatical when the P precedes the A, in a VP constituent.[3] These facts are summarised in table 1, and illustrated in (2) - (6).

(2) *Ku-'ita='e*   (*te*   *iaku*) *na*   *ana.*
     1SG-see=3P   CORE   1SG   CORE   child
     'I saw the child.'

(3) ~ *Ku'ita'e na ana te iaku.*

(4) *Ku-'ita te*   *ana*   (*na iaku*).
     1SG-see   CORE   mother   NOM   1SG
     'I saw a child.'

(5) * *ku'ita na iaku te ana.*

(6) *No-waliako=mo*   *na*   *anai*   *kampo.*
     3R-return=PF   NOM   child   OBL   village
     'The child returned to the village.'

Table 1. Agreement, word order and case marking in Tukang Besi

Clause type	variant	Agreement	word order	case marking		
				A	S	P
Bivalent	P indexed on V	A-V=P	V A P (~ VPA)	*te*		*na*
	P not indexed on V	A-V	V P A	*na*		*te*
Monovalent		S-V	V S		*na*	

The relevant testable phrase-structure configuration of the clause, with no preverbal elements, is shown in (7), which also shows the positional possibilities for adjunct material of different classes. Adverbs appear within the VP, left- or right- adjoined to any element. Locative adjuncts appear following the VP, while time adjuncts are constrained to appear following the S or A of the clause, that is, right of the IP. Table 2 shows the pronominal forms.[4] The free pronouns are listed to show the degrees of relative grammaticalisation between them and the bound verbal forms, the P clitics clearly being a more recent grammaticalisation than the prefixes.

(7)

```
                              CP
                    _____/‾‾_____
                  IP              te DPS,A, na DPP
           ____/‾‾\____            |_____|
         VP        na DPS,A      Time expressions
      __/‾‾\__
    V      te DPP      |_____|
    /\                      Locative adjuncts
 S/A-V
|_____|
  Adverbs
```
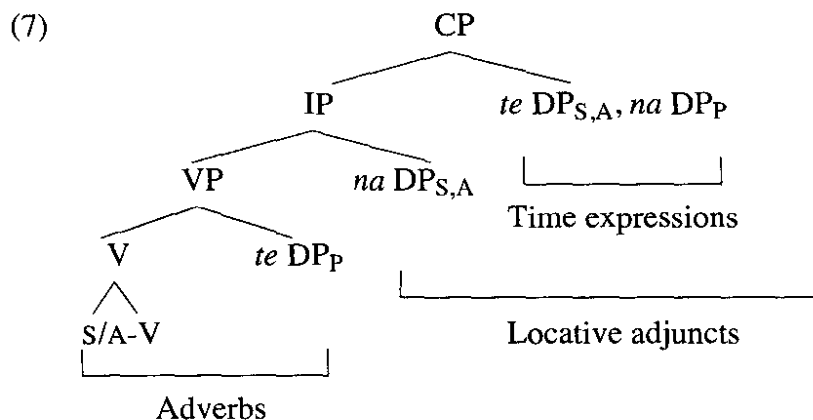
Table 2. Pronominal forms in Tukang Besi

Position: Role: (mood):	Pre-root S,A realis	irrealis	Post-root P	Post-root nominal POSS'R, subordinate A, S, P	Independent (any)
1SG	*ku-*	*ku-*	*=aku*	*=su*	*iaku*
2SG	*'u- / nu-*	*ko-*	*=ko*	*='u*	*iko'o*
3SG	*no- / o-*	*na- / a-*	*='e*	*=no*	*ia*
1PA	*ko-*	*ka-*	*=kami*	*=nto*	*ikami*
1PL	*to-*	*ta-*	*=kita*	*=mami*	*ikita*
2PL	*i-*	*ki-*	*=komiu*	*=miu*	*ikomiu*
3PL	*no- / o-*	*na- / a-*	*='e*	*=no*	*amai*

We shall return to the restrictions on the appearance of free and bound pronouns in more detail in 2.3.

2.2. The syntax of pronominal voice in Tukang Besi

While the use of S,A prefixes is obligatory, the use of P clitics is not. This has been seen in (2) and (4). The use of P-coding clitics on the verb has consequences for the assignment of properties associated with grammatical functions. This has been described elsewhere (Donohue 1999a, 1999b, 2002), but the essential properties are:

- the restriction of floated quantifiers;
- conjunction reduction;
- eligibility to appear as the head of an internal relative clause;
- external possession (possessor raising, possessor ascension).

Floated quantifiers can only refer to a nominative argument; this will be the single argument of a monovalent clause, and for a bivalent clause it will be the P, if there are P-coding clitics on the verb, or the A if there are not.

Floated quantifier referring to an A

(8)　[QUANT *Saba'ane*]　　*no-lemba*　*te*　　*kaluku*　　[*na*　　*amai*].
　　　　　　all　　　　　　3R-carry　CORE　coconut　　NOM　3PL
　　　'All of them carried coconuts.'
　　　* 'They carried all of the coconuts.'

75

Floated quantifier referring to a P

(9) [QUANT *Saba'ane*] *no-lemba='e* [*na* *kaluku*] *te* *amai*.
        all   3R-carry=3P   NOM coconut   CORE 3PL
   'They carried all of the coconuts.'
   * 'All of them carried coconuts.'

Conjunction reduction preferentially applies between nominative arguments in adjacent clauses. A textual example is presented in section 3, but the following sentences illustrate the principle. In both sentences the preferred controller for the third person argument of the monovalent clause *nowaliakomo di kampo* is the nominative argument of the following clause, even if there are no overt nominals in the second clause.

Coreference between S and A, both nominative

(10) *No-waliako=mo* *di* *kampo,* *maka* *no-'ita* *te* *ana.*
        3R-return=PF   OBL village   and.then 3R-see   CORE child
   'Øᵢ returned home, and then Øᵢ saw a childⱼ.'

Coreference between S and P, both nominative

(11) *No-waliako=mo* *di* *kampo,* *maka* *no-'ita='e* *(te* *ana).*
        3R-return=PF   OBL village   and.then 3R-see=3P   CORE child
   'Øᵢ returned home, and then (a child/Ø)ⱼ saw Øᵢ.'

Internal relative clauses show complex restrictions; only an absolutive argument (an S or a P) may appear as head, and moreover that argument must be nominative. This is trivial for an S, but for the P we can see that this restriction forces P-voice.

Internal relative clauses referring to a nominative P

(12) *No-waliako=mo* [RC *to-siasia='e* *na* *mia* *i* *aba*].
        3R-return=PF   1PL.R-beat.up=3P NOM person OBL previous
   'The person we beat up before has gone home.'

Internal relative clauses referring to a non-nominative P

(13) * *no-waliako=mo* [RC *to-siasia* *te* *mia* *i* *aba*]
          3R-return=PF   1PL.R-beat.up NOM person OBL previous
   'The person we beat up before has gone home.'

External possession is restricted to Ps and patient-like Ss (unaccusative subjects) (Donohue 1999b), in non-active, if bivalent, thus establishing a requirement that the external possessor must be nominative.

External possession with a nominative P

(14) *No-siasia=ko=mo*   *na* *ikaka='u.*
        3R-beat.up=2SG.P=PF   NOM eSi=2SG.GEN
   'They beat up your big brother.'

External possession with a non-nominative P

(15) * *no-siasia=moq* *(na / te)* *iko'o* *(na / te)* *ikaka='u*
          3R-beat.up=PF   NOM / CORE   2SG   NOM / CORE   eSi=2SG.GEN
   'They beat up your big brother.'

Clearly there are far-reaching grammatical consequences to the clause, in terms of constructional restrictions, that correlate with the choice of the verb

76

appearing with P-enclitics or not. We have, thus, established a voice system in which the identity of the grammatical subject changes (see Falk 2000 for arguments on why these constructions are suitable to establish grammatical subject status, and are not simply ones that identify an argument structure position). Evidence that the A is a core argument in both the P-agreement and the non-P-agreement constructions is easy to find. The A participant:

- remains indexed on the verb by invariant prefix regardless of voice, as seen in the invariant use of *ku-* to show agreement with the S,A in both (2) and (4);

  This shows that the determination of prefixal verbal agreement is based on lexically-assigned argument structure relationships, and not the voice-sensitive grammatical functions.

- is marked with a core case, *na* 'nominative' or *te* 'generic core argument', depending on the voice (again, see examples (2) and (4) above);

  Neither the general oblique *i*, nor any of the more specialised oblique prepositions, may be used to mark an A.

- is available to bind a reflexive regardless of its status with respect to the voice system (cf. Arka and Manning 1998).

The reflexive facts can be seen in the following two sentences (Donohue 1999a: 418), in which the marking or absence of marking for the P on the verb does not affect the ability of the A to bind a reflexive.

(16) *O-pepe='e*   *na*   *karama=no*   *te*   *ana.*
      3R-hit=3P   NOM   self=3GEN   CORE   child
      'The child hit her/himelf.'

(17) *O-pepe*   *te*   *karama=no*   *na*   *ana.*
      3R-hit   CORE   self=3GEN   NOM   child
      'The child hit her/himelf.'

A sentence with a reflexive agent, analogous to the English 'He was promoted by himself', is ungrammatical in Tukang Besi. This is most apparent with a non-third person S,A. (18) has a verb with prefixal agreement for the S,A, and (19) additionally shows agreement with the reflexive P, just as in (2) and (4). (20) is ungrammatical because it assigns the reflexive *karama'u* to the A position, grammatical in English since it is an oblique in the passive construction. The ungrammaticality of sentences such as (20) shows that the P-voice construction cannot be interpreted as a passive with an oblique A. This leads us to assume that the voice system involves an effective reversal of the assignment of grammatical functions to the two arguments of the bivalent verb. (Examining three or more place predicates shows that the voice system deals only with the highest two arguments for the purposes of voice assignment. Further arguments (such as the theme of a verb like *hu'u* 'give') are syntactically inert 'second objects'.)

(18) *'U-tulumi*   *te*   *karama='u.*
      2SG.R-help   CORE   self=2SG.GEN
      'You helped yourself.'

(19) 'U-tulumi='e   na   karama='u.
      2SG.R-help=3P   NOM   self=2SG.GEN

(20) * no-tulumi=ko   te   karama='u
      3R-help=2SG.P   CORE   self=2SG.GEN
      'You were helped by yourself.'

This voice system cannot be characterised as active/passive, or active/antipassive: in both of the bivalent coding options there are two core arguments. The voice system then represents more closely those of the Philippine languages, or the inverse systems of some Cariban or Algonquian languages (though not all; see Dahlstrom 1991). This can be shows with the following crude mapping of argument structure positions to grammatical functions in Tukang Besi (shown using the conventions from Alsina 1996) (Arguments against an ergative-with-antipassive analysis are given in Donohue 1999a: 160-166).

A-structure / GF-mapping for the pronominal voice alteration

(21) a. No enclitics ⟨ __A , __P ⟩   b. P enclitics   ⟨ __A , __P ⟩
                          | | SUBJ OBJ
                         SUBJ OBJ

An approach such as that advocated in Sells (2001), in which the P is raised in a clause with P clitics on the verb, would more formally capture the generalisations illustrated here. Note that, in Tukang Besi, rather than being covert the pronominal incorporation on the verb posited by Sells is overt, in the form of the P agreement clitics.

2.3. Interactions with person

It is not ungrammatical for an indexed pronominal to also be represented by a free pronoun in the same clause; this is extremely rare, but is attested in texts. Such constructions are usually associated with pragmatically marked information, and so are judged as being more natural if the case-marked pronoun appears preverbally, as in (23) and the textual (25).

(22) No-'ita=aku   te   ana   (#! na   iaku).
      3R-see=1SG.P   CORE   child   NOM 1SG
      'The child saw me.'

(23) Te   iaku   no-'ita=aku   te   ana.
      CORE   1SG   3R-see=1SG.P   CORE   child
      'The child saw me.'

(24) # To-waliako=mo   i   kampo   na   ikita.
      1PL.R-return=PF   OBL   village   NOM   1PL

(25) ...te   iaku   habuntu   (')u-hu'u=aku   te   kuli=no ...
      CORE   1SG   in.fact   2SG.R-give=1SG.P   CORE   skin=3GEN
      '...in fact you've just given me the peel (of the bananas) ...'

Sentences such as (26), while consistently judged to be grammatical possible, are at best marginally felicitous, and are not part of any corpus of naturally-occurring speech. This implies a constraint requiring local person Ps to be the subject of their clause; morphologically, local persons should be

78

marked on the verb. This does not apply to all local person, regardless of syntactic role; (27) shows an instance of a non-nominative local person A.

(26) *∕ # *no-'ita*   *te*   *iaku*   *na*   *ana*
          3R-see   CORE   1SG   NOM   child
          'The child saw me.'

(27)   *"Oho,*   *'u-pidi='e*   *te*   *iko'o*   *na*   *loka*   *saba'ane='e."*
          yes   2SG.R-rubbish=3P CORE   2SG   NOM   banana   all=3P
          'Yes, you've rubbished all of the bananas.'

2.4. Grammatical voice, not (just) pragmatic topic

Given what is known about the evolution of agreement systems in language (for instance, Givón 1976), we might ask whether the construction with P-clitics involves a change in the pragmatic function status of the P, but not the grammatical function status. This is a position that would be hard to sustain given the data on clause-internal constructions presented in the previous section, detailing the syntactic correlates of P-agreement, and the interaction with the person hierarchy.

Further evidence against a topic analysis involves the regular attestation of pre-verbal topics. Expanding the structure seen in (7), we might describe a monoclausal sentence in Tukang Besi as follows (after Donohue 1999: 80xxx). This accords well with what is described for other Austronesian languages.

(28)
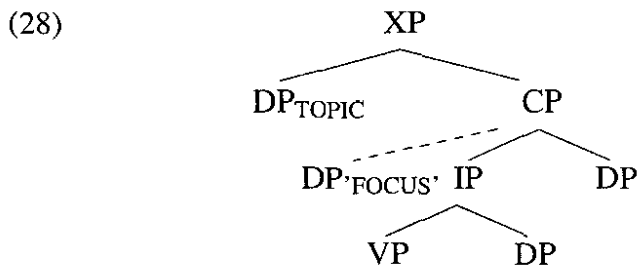
Examples of these structures can be found in the following textual extracts. The first shows that a 'true' topic, in preverbal position, need not have any grammatical connection with the arguments of the clause at all. In (29), in which two characters discuss the yield of a banana tree, the topic of the sentence, *te iaku*, is not the subject of the predicate. The textual extracts shown in (30) and (31) demonstrate that when a topic does serve as a term in the clause it is possible for it to be realised in the topic position alone, as in (30), or in both the preclausal topic position and overtly in the normal IP-internal position.

          TOPIC   SUBJECT
(29)   *"E*   *iaku*   *iso*   *no-ha'a*   *na*   *iso!*
          CORE   1SG   yon   3R-why   NOM   yon

          No-'ido   *sa-ro'o,*   *o-mate*   *sa-ro'o."*
          3R-live   1-leaf   3R-die   1-leaf
          'And me, it's just like that (ie., nothing at all). One leaf lives, and one leaf dies (ie., it's not doing too well).'

TOPIC
(30) ... *te*   *iaku*   *ku-[m]e-mbula*   *te*   *hu'u=no* ...
      CORE   1SG   1SG-VRB.SI-plant   CORE   trunk=3GEN
      'I will plant the trunk.'

TOPIC
(31) *Po'oli*   *te*   *ia*   *iso*   *no='eka=mo*
      finish   CORE   3SG   that   3R-ascend=PF

      na   *La bela kape'ingkape'i*   *kua*   *wunua-no.*
      NOM   Mr dear Fool   ALL   house=3GEN
      'And then he, The Fool, went (back) up to his house.'

On the other hand, it remains true that nominative encoding is not compatible with pragmatically focussed question words; questions can be formed of and A or a P only if it is non-nominatively coded, or in a cleft; an S may only be questioned in a cleft.

(32)   *'U-ala*   *te*   *paira?*   (33) * *'u-ala='e na paira?*
      2SG.R-fetch   CORE   what   2SG.R-fetch=3P
      'What did you fetch?'

(34)   *No-ala='e*   *te*   *emai?*   (35) * *no-ala na emai?*
      3R-fetch=3P   CORE   who   2SG.R-fetch
      'Who fetched it?'

(36)   *Te*   *emai*   *na*   *[m]aliako?*   (37) * *no-waliako na emai?*
      CORE   who   NOM   return.SI   2SG.R-fetch
      'Who returned?'

(38)   *Te*   *paira*   *na*   *ni-ala='u?*
      CORE   what   NOM   PP-fetch=2SG.GEN
      'What did you fetch?'

(39)   *Te*   *emai*   *na*   *[um]ala*   *te*   *loka?*
      CORE   what   NOM   fetch.SI   CORE   banana
      'Who fetched the banana(s)?'

This is not completely surprising, given the well-known correlations between the relative topicality of an A and a P and the choice of voice. This is summarised in table 3.

Table 3. Topicality and voice choice

Relative topicality of A and P	Preferred voice	Subject =
A » P	antipassive	A
A > P	active	A
A < P	inverse	P
A « P	passive	P

(Rarely are more than three of these voices instantiated for any one language. See Payne (1994), or Sells (2001) for the use of these labels as a combined voice+pragmatic function descriptors.)

There are, nonetheless, good reasons to consider the alternations to represent a voice system, and not simply an artefact of topicalisation (such as has been demonstrated for, amongst others, Chichewa – see section 3). The evidence involves the syntactic tests shown in 2.2, and the semantic correlates described in the following section.

2.5. Semantic correlates of the pronominal voice system

Other reasons for assuming that the pronominal voice system is a real voice system, and not simply an extended topicalisation phenomenon, become apparent when we examine the necessary semantic correlates of the voice choice made with this paradigm. In addition to the change in status of the grammatical functions, other effects associated with the use of P-clitics when there is an alternation include the following specification of the P argument:

- greater specificity or definiteness associated with the P;

- more referential or retrievable P.

The contrast between the following clauses shows the instantiation of these semantic parameters.

(40) *Ku-ala   te   loka.*   (41) *Ku-ala='e   na   loka.*
1SG-fetch CORE banana   1SG-fetch=3P NOM banana
'I fetch some/a banana(s).'   'I fetched the banana(s).'
(irrealis, indefinite, nonspecific)   (perfective, specific)

The fact that local persons are almost invariably marked on the verb, including when they have the role of the P in a clause (see section 2.3) also matches what is known about correlations between animacy and voice choice/subject selection.

There are also correlations between aspectual interpretation and the choice of pronominal voice, something that we would not expect to find with a system that simply marked incorporate pronominal status and relative topicality. A P-voice clause is more telic/punctual/realis that an otherwise equivalent clause in active voice. Relative clauses with an *-um-* infixed verb (necessarily with an S or A as head), in which the normal perfective clitic *=mo* cannot be used, use P-voice as one way to indicate a perfective aspect.

The telicity parameter has been illustrated in (40) and (41). This contrast in relative clauses can be seen in (42) and (43) (which, with (44), also demonstrate the continued function of the older Austronesian voice morphology *-um-* 'actor voice' and *-in-* 'PV' in subordinate clause, and show that in Tukang Besi they have become disassociated from the determination of grammatical functions).by a verb infixed with *-um-*; independently, the P clitics may be used to indicate a nominative P in the relative clause, and a level of perfectivity not otherwise associated with this construction. Interestingly, relative clauses are one of the few cases in which local person Ps do not have to be coded on the verb, though that is still the preferred option.

(42) *Te   wowine   [RC s[um]ampi   te   loka]*
CORE woman   pick.SI   CORE banana
'the woman who pick(ed, s) bananas'

(43) *Te   wowine   [RC s[um]ampi='e   na   loka]*
CORE woman   pick.SI=3P   NOM banana
'the woman who has picked the banana(s)'

(44) *Te   loka   [RC ni-sampi   nu   wowine]*
CORE banana   PP-pick   GEN woman
'the bananas which were/are being picked by the woman'

81

All of the semantic correlates illustrated here are factors that Hopper and Thompson associate with higher transitivity, and all are also variables commonly associated with the alternations of a voice system: voices that assign high grammatical status to the P of a clause are more likely to appear in past or perfective clauses, and to refer to more specific referents. Again, the agreement alternation in Tukang Besi functions semantically as a voice system can be expected to function.

3. The genesis of a pronominal voice system

While I have argued that the P-voice in Tukang Besi is not (simply) a case of a clause reflecting incorporated pronominal material and a topical P argument, there are striking similarities with such phenomena as reported in, for instance, Chichewa (Bresnan and Mchombo 1987). In both languages a P DP in a clause *with* P-coding agreement is outside the VP, while a P DP in a clause *without* agreement is inside the VP. We shall propose formal representation of the structure of clauses with and without P agreement on the verb in (99) - (99).

Having two transitive verb forms differentiated by the presence or absence of P-agreement is not limited to Tukang Besi. Georgopoulos (1985, 1998) discusses similar correlations of 'object agreement' in Palauan, as does Kissock (2003) on Rotuman, and there is evidence that other languages of Melanesia also show pronominal voice alternations (section 4). In all cases these are languages that no longer display the original Austronesian voice morphology in its original function.

The historical loss and development of Austronesian voice morphology is summarised in (45), with considerable simplification.

(45)	Main Clause		Subordinate clause	
	S,A subject	P subject	S,A subject	P subject
PAN	V-*um*-	V-*in*-	V-*um*-	V-*in*-
(Celebic)	V-*um*-	S,A-V	V-*um*-	V-*in*-
P M-B	S,A-V	S,A-V-P	V-*um*-	V-*in*-

The Tukang Besi system reflects the proto-Muna Buton system; we have seen the alternation in the main clause correlating with the presence or absence of P enclitics; the subordinate clause morphology has been seen in (36) and (38) - (39), as well as in (42) - (44). By comparison in Tagalog the original Austronesian voice morphology (Ross 2002) is used in a voice function in both main clause and subordinate clause positions.

(46)   *P[um]unta*   *ang*   **bata**   *sa*   *pulo.*
      examine.AV   NOM   child   DAT   island
      'The child went to the island.'

(47)   *S[um]uri*   *ng*   *kaibigan*   *niya*   **ang**   **bata.**
      examine.AV   GEN   friend   3SG.GEN   NOM   child
      'The child examined her/his friend.'

(48)   *S[in]uri*   *ng*   *bata*   **ang**   **kaibigan**   *niya.*
      examine.PV   GEN   child   NOM   friend   3SG.GEN
      'A child examined her/his friend.'

We can see that there are good reasons to consider that the morphological alternation associated with the presence or absence of P clitics has replaced, in main clauses, the encoding of the voice by means of the infixes. The use of this voice alternation in discourse has been hinted at in 2.2 can be illustrated with a real textual example in (49). Here we can see that the preferred target of a clause-chaining construction is the subject of that clause, and the voice in that clause is selected to guarantee such coreference.

(49)   *Ara*   *ku-[m]o-busu*   *na-t[um]alo=aku,*   *kene*   *te*
      if   1SG-REC.SI-forward.fist   3I-win.SI=1SG.P   and   CORE

       ia   *no-pande di*   *lola-'a,*   *jari*   *labi*   *ku-akala='e.*
       3SG   3R-clever OBL   fly-NL   so   better   1SG-trick=3P
       'If I₁ want to fight heⱼ'll beat me₁, and heⱼ's good at flying, so
       it'd be better if I₁ tricked himⱼ.'   (Reiger & Aap: 24)

(49)'   1SG: S → P   A | SUBJ → SUBJ   OBJ
       him: A → S → P | OBJ → SUBJ → SUBJ

In (50) we have an alternative version of the same passage from the story, in which the syntactic roles of the participants are the same, but in which the grammatical function status of those arguments has been changed to deny the SUBJ → SUBJ preference in coreference. This retelling is uniformly judged to be much less felicitous than the original in (49).

(50) */# *Ara*   *ku-[m]o-busu*   *na-t[um]alo*   *te*   *iaku,*
        if   1SG-REC.SI-forward.fist   3I-win.SI   CORE   1SG

        kene   *te*   *ia*   *no-pande*   *di*   *lola-'a,*
        and   CORE   3SG   3R-clever   OBL   fly-NL

        jari   *labi*   *ku-akala*   *te*   *ia.*
        so   better   1SG-trick   CORE   3SG
        'If I₁ want to fight him heⱼ'll beat me₁, and heⱼ's good at flying,
        so it'd be better if I₁ tricked himⱼ.'   (retold, badly)

(50)'   1SG: S → P   A | SUBJ → OBJ   SUBJ
       him: A → S → P | SUBJ → SUBJ → OBJ

Following Bresnan and Mchombo (1987) I assume that the difference between clauses with P agreement on the verb and those without, in Chichewa, can be modelled as in (51) and (52). In (51) we see the DP representing the P internal to the VP external to the VP, since it is already present in the VP in the form of the object agreement prefix on the verb. Furthermore, as described in Bresnan and Mchombo, there are positional freedoms associated with the P in this clause that are not found in a clause in which the verb lacks agreement for the P. In a clause with no P clitics, the DP representing the P must appear inside the VP, and is not eligible for any special positional privileges. This is shown in (52).

(51) Chichewa: Clause with P marking, enclitic interpreted as an anaphoric pronoun, nominal P topical and VP-external

f_1:

$$\begin{bmatrix} f_2\colon \text{TOP} = f_7\colon & [\text{PRED} = \text{`...'}\,] \\[4pt] f_6\colon \text{PRED} & \langle\_\,,\_\rangle \\[4pt] f_2\colon (\text{A}) & \begin{bmatrix} \text{PRED} = \text{`...'} \\ \text{PERS} = \alpha \\ \text{NUM} = \beta \end{bmatrix} \\[12pt] f_7\colon (\text{P}) & \begin{bmatrix} \text{PRED} = \text{`PRO'} \\ \text{PERS} = \gamma \\ \text{NUM} = \delta \end{bmatrix} \end{bmatrix}$$

```
            CP f1
          /      \
      DP f2        IP f3
  (f2 TOP = f7)  /      \
            DP f4        VP f5
                          |
                        V f6, f7
```

$(f_4 \text{ A PERS}) = \alpha$
$(f_4 \text{ A NUM}) = \beta$
$(f_2 \text{ P PERS}) = \gamma$
$(f_2 \text{ P NUM}) = \delta$
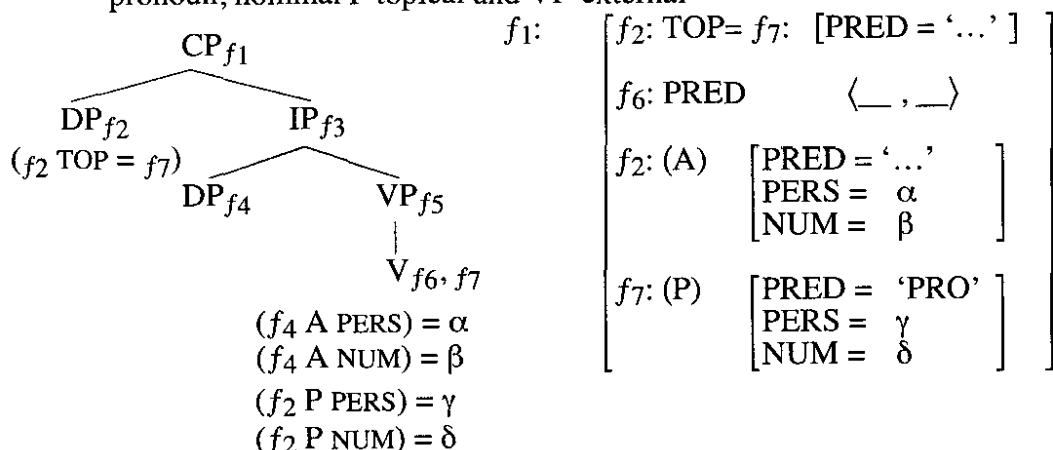
(52) Chichewa: Clause without P-enclitic, nominal P fills argument position, VP-internal

f_1:

$$\begin{bmatrix} f_4\colon \text{PRED} & \langle\_\,,\_\rangle \\[4pt] f_2\colon (\text{A}) & \begin{bmatrix} \text{PRED} = \text{`...'} \\ \text{PERS} = \alpha \\ \text{NUM} = \beta \end{bmatrix} \\[12pt] f_5\colon (\text{P}) & \begin{bmatrix} \text{PRED} = \text{`...'} \\ \text{PERS} = \gamma \\ \text{NUM} = \delta \end{bmatrix} \end{bmatrix}$$

```
        IP f1
      /      \
  DP f2        VP f3
            /      \
        V f4        DP f5
```

$(f_2 \text{ A PERS}) = \alpha$
$(f_2 \text{ A NUM}) = \beta$
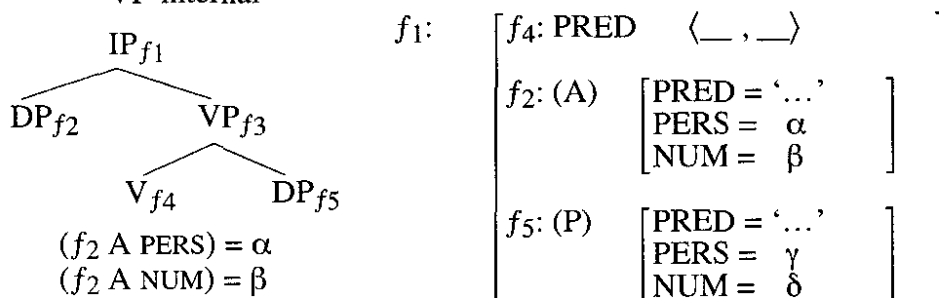
A typical model of the constituent structure and functional structure corresponding to a passive predicate (not in Tukang Besi, but generally) would be similar to that shown in (53) (assuming, for the sake of simplicity, an English-like phrase structure). Although a passive necessarily involves a reassignment of grammatical functions, unlike the case for pronominal agreement marking in Chichewa, the commonality with pronominal agreement systems is that the pragmatic function of topicality is assigned to the P argument based on the morphological choice on the verb.

(53) Generic language: Agreement for P subject, passive voice: P topical, A core

f_1:

$$\begin{bmatrix} f_2\colon \text{TOP} & [\text{PRED} = \text{`...'}\,] \\[4pt] f_4\colon \text{PRED} & \langle\text{SUBJ}\rangle, \langle\text{OBL}\rangle \\[4pt] f_2\colon \text{SUBJ (P)} & \begin{bmatrix} \text{PERS} = \gamma \\ \text{NUM} = \delta \end{bmatrix} \\[10pt] f_7\colon \text{OBL (A)} & \begin{bmatrix} \text{PRED} = \text{`...'} \\ \text{PERS} = \alpha \\ \text{NUM} = \beta \end{bmatrix} \end{bmatrix}$$

```
        CP f1
      /      \
  DP f2        VP f3
            /      \
        V f4        PP f5
                   /    \
                 P f6    DP f7
```

$(f_2 \text{ P PERS}) = \gamma$
$(f_3 \text{ P NUM}) = \delta$
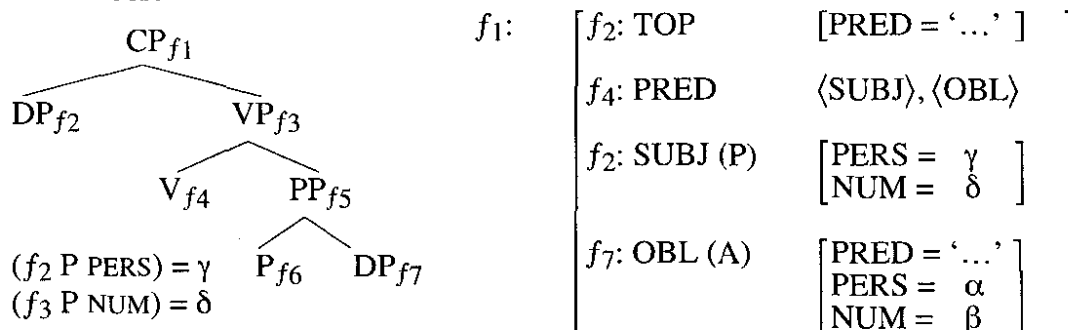
It is not hard to see how an agreement system with an optional pronominal element, associated with some level of increased pragmatic force, can develop into a voice system with grammatical as well as pragmatic status associated with its alternations. The pronominal voice system in Tukang Besi

involves an alternation between a structure such as (54), representing the clause without P-enclitics, and (55).

(54) Tukang Besi: Agreement for A only: A & P core

IP$_{f1}$
VP$_{f2}$ DP$_{f3}$
V$_{f4}$ DP$_{f5}$

(f_3 A PERS) = α
(f_3 A NUM) = β

f_1:
$$\left[\begin{array}{ll} (f_3\text{: TOP} & [\text{PRED} = \text{'...'}\ ]\quad) \\[4pt] f_4\text{: PRED} & \langle \text{SUBJ, OBJ} \rangle \\[4pt] f_3\text{: SUBJ (A)} & \left[\begin{array}{l} \text{PRED} = \text{'...'} \\ \text{PERS} = \alpha \\ \text{NUM} = \beta \end{array}\right] \\[12pt] f_5\text{: OBJ (P)} & \left[\begin{array}{l} \text{PRED} = \text{'...'} \\ \text{PERS} = \gamma \\ \text{NUM} = \delta \end{array}\right] \end{array}\right]$$

(54)'
                                     OBJ   SUBJ

[IP [VP *No-'ita* [DP:P *te* *kene=no*]] [DP:A *na* *ana*]].
3R-see   CORE friend=3GEN   NOM   child
'The child saw her/his friend.'

(55) Tukang Besi: Agreement for A and P: P more topical, A core

CP$_{f1}$
IP$_{f2}$ DP$_{f3}$
VP$_{f4}$ DP$_{f5}$
V$_{f6', = f7}$

(f_5 A PERS) = α
(f_5 A NUM) = β
(f_7 P PERS) = γ
(f_7 P NUM) = δ

f_1:
$$\left[\begin{array}{ll} f_3\text{: 'TOP'} = f_7 & [\text{PRED} = \text{'...'}\ ] \\[4pt] f_6\text{: PRED} & \langle \text{SUBJ, OBJ} \rangle \\[4pt] f_7\text{: SUBJ (P)} & \left[\begin{array}{l} \text{PRED} = \text{'PRO'} \\ \text{PERS} = \alpha \\ \text{NUM} = \delta \end{array}\right] \\[12pt] f_5\text{: OBJ (A)} & \left[\begin{array}{l} \text{PRED} = \text{'...'} \\ \text{PERS} = \alpha \\ \text{NUM} = \beta \end{array}\right] \end{array}\right]$$

(55)'
                                       OBJ   SUBJ

[CP [IP [VP *No-'ita='e*] [DP:A *te* *ana*]] [DP:P *na* *kene=no*]]].
3R-see=3P   CORE child   NOM   friend=3GEN
'The child saw her/his friend.'

Some predicates can be used either monovalently or bivalently. When they are used bivalently they must appear with P-enclitics (Donohue 1999a: 100). This indicates a slight preference for the inverse voice in a bivalent clause. This suggests that the bivalent use is dependent on combination with a predicate, either a causative one that specifies only the linking of the predicates, or a pronominal voice derivation that not only specifies predicate linking but also the voice that assigns the grammatical function subject to the P. See Sells (2000) for a model that is even more plausibly instantiated in Tukang Besi than in the languages for which it was developed.

4. Other attestations of pronominal voice

Similar evidence for pronominal voice systems similar to that described here can be found in other Austronesian languages to the east (Larike, Ansus and Ambai, Misima, Saliba, Tawala; Laidig and Laidig 1995, Callister 1987, Margetts 1999, Ezard 1997); the unrelated (but geographically contiguous) West Papuan languages of North Halmahera (eg., Wimbish 1991) (in an inverted form); Mixe (Oto-Manguean, Mexico; Dietermann 1998), Dullay (Cushitic, east Africa; Sasse 1984: 247) and Lango (xxxx, east Africa; Noonan 1992 and others).

Generally there is less reported morphosyntactic evidence for the pronominal voice system in these languages; in some cases this reflects the level of published documentation (eg., Muna; van den Berg 1989), in some cases it reflects genuine 'fuzzy' language phenomena, in which the level of grammaticisation is not complete, such as in Ansus; the most comparable construction is textual coreference: for a chain of subjects; new participants are introduced as objects, and then (if they stay) become subjects.

Endnotes

1. There are exceptions to this trend of under-reportage; Arka (2000) reports on the behaviour of passives in several languages of eastern Indonesia, where the nominal marking (case and/or position) is overtly present and correlates with demonstrable syntactic behaviour, but there is no dedicated passive morphology on the verb. Palu'e (Donohue 2004) has a similar system.
2. Because of the Philippinist controversy over grammatical functions, I shall use the labels A, S and P (following Comrie 1978) to refer to the syntactic roles, without making claims for their grammatical status at this stage. A, S and P refer to the most agentive argument in a lexically bivalent clause, the single argument in a monovalent clause, and the non-A non-oblique argument(s) in a bi- or trivalent clause, respectively.
3. A further alternative is available in which one nominally appears preverbally (yet clause-internally), with pragmatic focus: this argument can only be the P in a clause with P-enclitics, or the A (or S) in a clause without P enclitics. Additionally, any DP may appear preverbally if it is topical. A preverbal argument is invariably marked with *te*.See 2.3 and 2.4 for examples.
4. The P clitics only show morphophonemic variation: the third person clitic is realised as [kE] following a syllable with a glottal stop onset. The variation in the third person prefixes, and in the 2SG realis prefixes, is unconditioned.

References

Alsina, Alex. 1996. *The role of argument structure in grammar: evidence from Romance*. CSLI publications, Stanford University.

Arka, I Wayan. 2000. *Voice and Being Core: evidence from (Eastern) Indonesian Languages.* Paper presented at the AFLA 7 Conference Amsterdam, May 2000.

Arka, I Wayan, and Christopher D. Manning. 1998. *Voice and grammatical relations in Indonesian: a new perspective.* In Miriam Butt and Tracey King, eds., On-line LFG proceedings. Available at http://www-csli.stanford.edu/publications/ LFG3/lfg98-toc.html.

Berg, René van den. 1989. *A grammar of the Muna Language.* KITLV, Verhandelingen 139. Dordrecht/Providence: Foris.

Bresnan, Joan, and Sam A. Mchombo. 1987. Topic, Pronoun, and Agreement in Chichewa. *Language* 63 : 741-782.

Callister, William. 1987 (for 1985). 'How Austronesian focus works in the Misima language. Language and Linguistics in Melanesia 16 (1-2): 71-99.

Comrie, Bernard. 1978. Ergativity. In Winifred P. Lehmann, ed., *Syntactic typology: studies in the phenomenology of language*: 329-394. Sussex: The Harvester Press.

Dahlstrom, Amy Louise. 1991. *Plains Cree morphosyntax.* Outstanding dissertations in linguistics series, Jorge Hankamer, ed. New York: Garland Publishing.

Dieterman, Julia Irene. 1998. Participant reference in Isthmus Mixe narrative discourse. *Journal of Translation and Textlinguistics* 10: 47-79.

Donohue, Mark. 1999a. *A grammar of Tukang Besi.* Berlin: Mouton de Gruyter.

   1999b. Syntactic roles vs. Semantic roles: External Possession in Tukang Besi. In Doris Payne and Immanuel Barshi, eds., *External Possession.* Typological Studies in Language No. 39: 373-401. Amsterdam: John Benjamins.

   2002. Voice in Tukang Besi and the Austronesian voice system. In Fay Wouk and Malcolm Ross, eds., *The history and typology of western Austronesian voice systems*: 81-99. Canberra: Pacific Linguistics 518.

   2004. The Palu'e passive: from pragmatic construction to grammatical device. In I Wayan Arka and Malcolm Ross, eds., *Voice in Western Austronesian languages.* Canberra: Pacific Linguistics.

Ezard, Bryan. 1997. *A grammar of Tawala, an Austronesian language of the Milne Bay area, Papua New Guinea.* Canberra: Pacific Linguistics C-137.

Falk, Yehuda. 2000. Philippine subjects in a monostratal framework. In Carolyn Smallwood and Catherine Kitto, eds., *The Proceedings of the Austronesian Formal Linguistics Association VI*: 133-136. Toronto Working Papers in Linguistics. Also available at http://pluto.mscc.huji.ac.il/~msyfalk/PhilippineSubjects.pdf.

Georgopoulos, Carol. 1985. Variables in Palauan Syntax. *Natural Language and Linguistic Theory* 3: 59-94.

   1998. Direct object definiteness effects. *Canadian Journal of Linguistics* 43 (3/4): 307-340.

Givón, Talmy. 1976. Topic, Pronoun and Agreement. In Charles Li, ed., *Subject and Topic*: 149-188. New York: Academic Press.

Kissock, Madelyn J. 2003. Transitivity and Objecthood in Rotuman. *Oceanic Linguistics* 42 (1): 145-160.

Laidig, Wyn D., and Carol J. Laidig. 1995. A synopsis of Larike phonology and syntax. In Wyn D. Laidig, ed., *Descriptive studies in languages of Maluku Part II. NUSA: linguistic studies of Indonesian and other languages in Indonesia* 38: 19-42. Jakarta: Universitas Katolik Indonesia Atma Jaya.

Margetts, Anna. 1999. *Valence and transitivity in Saliba: an Oceanic language of Papua New Guinea.* Ponsen & Looijen BV.

Noonan, Michael. 1992. *A grammar of Lango.* Berlin: Mouton De Gruyter

Payne, Thomas. 1994. The pragmatics of voice in a Philippine language: Actor-focus and goal-focus in Cebuano narrative. In Talmy Givón, ed., *Voice and inversion:* 317-64. Amsterdam : John Benjamins.

Sasse, H.J. 1984. The pragmatics of noun incorporation in Eastern Cushitic languages. In Frans Plank, ed., *Objects: towards a theory of grammatical relations:* 243-268. London: Academic Press.

Sells, Peter. 2000. Raising and the order of clausal constituents in the Philippine languages. In Ileana Paul, Vivianne Phillips, and Lisa Travis, eds., *Formal issues in Austronesian linguistics:* 117-143. Dordrecht: Kluwer.

2001. Form and function in the typology of grammatical voice systems. In Géraldine Legendre, Jane Grimshaw and Sten Vikner, eds., *Optimality-theoretic syntax:* 355-391. Cambridge: MIT Press.

Schachter, Paul. 1976. The Subject in Philippine languages: Topic, Actor, Actor-Topic, or None of the Above. In C.N. Li, ed., *Subject and Topic:* 491-518. New York: Academic Press.

1977. Reference-related and Role-related properties of Subjects. In Peter Cole and Jerrold M. Sadock, eds, *Syntax and Semantics 8: Grammatical Relations:* 279-305. New York: Academic Press.

Wimbish, Sandra Gay. 1991. *An introduction to Pagu through the analysis of narrative discourse.* MA thesis, University of Texas at Arlington.

Wouk, Fay, and Malcolm Ross, eds.2002. *The history and typology of western Austronesian voice systems:* 81-99. Canberra: Pacific Linguistics 518.

Mark Donohue
National University of Singapore
Department of English, Language, and Literature
National University of Singapore
Block AS5, 7 Arts Link
Singapore   117570
mark@donohue.cc

Infix Allomorphy in Ida'an-Begak

Nelleke Goudswaard
Vrije Universiteit Amsterdam

1. Introduction

Ida'an-Begak is a Western Malayo-Polynesian language spoken by approximately 6,000 people on the east coast of Sabah, Malaysia, Borneo and belongs to the Sabahan subgroup of the North Borneo subgroup (Blust 1998). Ida'an-Begak has three dialects, Ida'an, spoken in the villages of Segama to the west of Lahad Datu, Ida'an Sungai spoken in the Kinabatangan and Sandakan districts, and Begak spoken in Ulu Tungku, to the east of Lahad Datu (Banker 1984).[1] Moody (1993) deals with Ida'an; this paper concentrates on the Begak dialect. In this paper I will present new data gathered in the field and provide an analysis of the allomorphy. The study is based on spontaneous data as well as examples elicited from my language informants.

Several languages of this group show infixation/prefixation allomorphy, where the vowel of the infix allomorph coalesces with the stem vowel (Blust 1997's 'ablaut'). Ida'an-Begak shows a rather complex instantiation of this phenomenon. In this paper I will present new data gathered in the field and provide an analysis of this allomorphy.

The Ida'an-Begak Past Tense (P) is marked by *ni-*, *-i-* or *-ən-*, depending on the shape of the stem, while the Dependent (D) is marked by *m-*, *-u-* or *-əm-*, depending on the shape of the stem.[2] The distribution of these infixes is as in Table 1. Stems that start with a consonant followed by schwa or /a/ are infixed with *-i-* (P) and *-u-* (D), resulting in vowel coalescence. Stems that start with a consonant followed by a high vowel are infixed with *-ən-* (P) and *-əm-* (D); vowel-initial stems are always prefixed with *ni-* (P) and *m-* (D). Stems starting with a liquid followed by a high vowel are prefixed with *nə-* (P) and *mə-*(D). The result of the affixation is ideally a consonant-initial bisyllabic form, which is the optimal prosodic word of the language.

Stem Starts With	stem	Gloss	P. affix	Past Tense	D. affix	Dependent
V	*abput*	'bite'	*ni-*	*nebput*	*m-*	*mabput*
Ca/Cə	*səgkow*	'call'	*-i-*	*sigkow*	*-u-*	*sugkow*
Ci	*timbak*	'shoot'	*-ən-*	*tənimbak*	*-əm-*	*təmimbak*
Cu	*tunu*	'set on fire'	*-i-*	*tinu*	*-əm-*	*təmunu*
Li	*ligow*	'deceive'	*nə-*	*nəligow*	*mə-*	*məligow*

Table 1: Distribution of the suppletive allomorphs

The morphemes of the Past Tense and the Dependent derived historically from proto-Austronesian *IN and *UM respectively (Blust 1997). Contrary to Ussishkin (2000)'s analysis of the same phenomenon in the related language Mukah Melanau, I will claim that the synchronic allomorphs in Ida'an-Begak are no longer derived from one underlying form. Synchronically,

they must be analysed as suppletive affixes that still bear resemblance to *IN and *UM but never surface as such. The claim of this paper is that allomorphy in Ida'an-Begak is suppletive and that the choice of allomorphs is governed by constraint ranking. The phonology decides which allomorph fits best for which stem.

In the next section I will describe which constraints are needed to obtain this optimal consonant-initial bisyllabic form. In section 3, the interaction between the various constraints is analyzed. In section 4, arguments are presented in favour of an analysis with suppletive allomorphs, and section 5 offers some conclusions.

2. Description of the data and possible constraints

The attachment of the Past Tense and Dependent affixes always results in consonant-initial words. Therefore ONSET must be rather highly ranked in Begak. ONSET is active in various parts of the phonology. The language uses for example glide insertion and vowel coalescence to avoid vowel hiatus. Sometimes, glottal stops are inserted if a word starts with a vowel. Nevertheless, certain words (nouns and stative verbs) may start with a vowel; in that case ONSET is overruled by other constraints.

(1)   ONSET: "A syllable must have an onset" (Prince & Smolensky 1993:99)

The second constraint that we need to adopt is that a word must be a foot, i.e. a word is maximally and minimally two syllables long. This constraint explains why Begak strives towards bisyllabic words.

(2)   PRWD=FT: "a word is a foot / maximally two syllables long"
This constraint is a slight modification of PRWD=BIN: a Prosodic Word must be binary (Prince &Smolensky 1993: 55).[3]

The third constraint that we need to adopt says that only the last two syllables of a word can have a full vowel. Not all words are bisyllabic, but if they are longer, for example after a bisyllabic stem has been prefixed or infixed, the syllables before the penultimate one can only contain an empty vowel. All prefixes have the shape C-, Cə (C)- or CəCə- and the three non-coalescing infixes have the shape -əC-, thereby filling non-footed syllables with schwa[4]:

(3)	*gə-*	Actor Voice	*gə-lindut*	'run'
	bəg-	Actor Voice	*bəg-arab*	'search'
	məng-	Actor Voice	*məng-inum*	'drink'
	kə-	Actor Voice-Stative	*kə-lapas*	'passed by'
	məngə-	Actor Voice causative	*məngə-lawas*	'make clean'
	pə-	Dependent causative	*pə-buay*	'make long'
	nə-	Past Tense causative	*nə-lewas*	'made clean'
	tə-	Intensive	*tə-buat*	'very long'
	məkə-	Petitive	*məkə-ledung*	'ask for help to go down'
	səng-	Nominalization	*səng-əgkot*	'manner of work'

Infixes in Begak have only schwa as their vowel:

(4)	*-ər-*	Reciprocal	*s-ər-agga'*	'fight with each other'
	-ən-	Past Tense	*t-ən-iru'*	'taught'
	-əm-	Dependent	*t-əm-iru'*	'teach'

Just like some other North Bornean languages, Begak does not have any productive suffixes, but even historically suffixed words have schwa in the prepenultimate syllable:

(5)	*turug*	'sleep'	*tərug-an*	'bed'
	tapis	'strain'	*təpis-an*	'strainer'
	?		*kə-təpus-an*	'last'
	?		*kələgbung-an*	'ridgepole'

The Stative prefix *a-* is the only exception; it forms a kind of appendix and appears to the left of all other prefixation.

(6)	*ligot*	'late'	*a-ligot*	'rather late'
	tot	'stuck'	*a-p-ə-tot*	'accidentally stuck'
	luan	'go out'	*a-kə-luan*	'accidentally go out'
	inum	'drink'	*a-kə-p-inum*	'accidentally drink'

Blust (1997:21) formulates the vowel distribution of North Bornean languages as Prepenultimate Neutralization. This can be turned into the following constraint:

(7) PREPENULTIMATE NEUTRALIZATION (PPN): "Every unfooted syllable is reduced to schwa."

PPN is actually a weaker form of the constraint PRWD=FT mentioned above and can therefore perhaps better be split up in two separate constraints: PRWD=FT and a form of Positional Faithfulness (Beckman 1997, Alderete

91

1995) or Licensing (Steriade 1994a,b, Zoll 1998). Positional Faithfulness deal with faithfulness constraints of elements in "strong" positions while Licencing deals with the markedness of elements in "weak" positions. Although PPN can be split up, I will use the constraint PPN to deal with unstressed syllables, for ease of exposition.

The fourth observation on Table 1 is that vowel coalescence is often used to avoid vowel hiatus and to keep words bisyllabic. Malay loan words containing a vowel hiatus or a sequence of a glide and a vowel are adapted by means of vowel coalescence, as in (9). This means in OT terms that UNIFORMITY must be ranked low in Begak.

(8)   UNIFORMITY: "no element of S₂ has multiple correspondents in S₁"
       (McCarthy & Prince 1995)

(9) **Malay**   **gloss**   **Begak**   **gloss**
     wayang   'movie'   *oyang*   'movie'
     wayar   'wire'   *oyar*   'wire'
     tuala   'towel'   *tola'*   'towel'
     pəlihara   'look after'   *(pə)lera'*   'look after'
     jarum [djarum]   'needle'   *derum*   'needle'
     ajaib [adjaib]   'miraculous'   *deip*   'astonished'
     janji [djandji] 'promise'   *dendi*   'vow'

Begak has only four underlying vowels: /a/, /i/, /u/, and /o/ in final syllables and /a/, /i/, /u/ and schwa in penultimate syllables.

(10)

	Front -Round		Round Back
High	i		u
-High -Low	e	ə	o
Low		a	

The two derived vowels /e/ and /o/ can only occur as a result of vowel coalescence in the penultimate syllable. These vowels are actually two vowels realized in one segment and can be represented as in (11). The vowel schwa is always overwritten by /i/ and /u/ after vowel coalescence, because schwa has no features to start with.

(11)

V   +   V   =   V
| [high front -round]   [high front -round]
| | |
/ə/   /i/   /i/

V   +   V   =   V
| [high back round]   [high back round]
| | |
/ə/   /u/   /u/

92
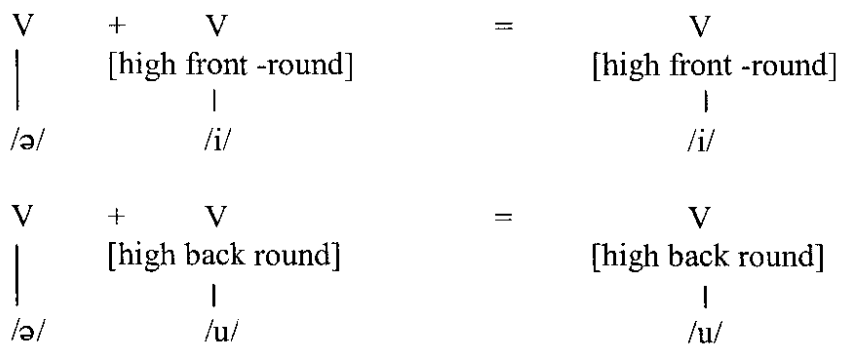
```
V        +     V                    =              V
[low]    [high front -round]                       [-high -low front -round]
 |            |                                            |
/a/         /i/                                          /e/


V        +     V                    =              V
[low]    [high back round]                         [-low-high back round]
 |            |                                            |
/a/         /u/                                          /o/


V        +     V                    =              V
[high    +    [high back round]                    [high front -round]
front
-round]        |                                          |
 |             |                                          |
/i/          /u/                                        /o/
```

Other combinations of vowels are not involved in coalescence because in the case of affixation, an other allomorph is chosen. If the stem vowel /u/ is infixed with -*i*- it is overwritten by it, for example /t-i-unu/ 'burn' becomes *tinu*, but the opposite does not occur: stems with a penultimate /i/ cannot be infixed with -*u*-, for example */t-u-iru'/ but *t-əm-iru'* 'teach'. This is an idiosyncracy that I cannot explain.

The last constraint that needs to be adopted for the analysis of the Past Tense and Dependent allomorphy forbids complex onsets: *COMPLONSET (Prince and Smolensky 1993). Begak native words lack complex onsets and complex onset in loan words are adapted, as is illustrated in (12).

(12) **English**   **Begak**
      tractor   *təlaktul*
      class   *kəlas*
      school   *iskul*

3. Analysis
We can now adopt the following hierarchy of the constraints:

(13)   ONSET, PPN, *COMPLONSET >> PRWD=FT >> UNIFORMITY

As for the choice between infixation and prefixation, it is possible to adopt an alignment constraint ALIGN (affix, L) (Prince&Smolensky 1993, McCarthy 2002) for Begak, because the language has no productive suffixes. All its affixes are prefixes, or infixes that are affixed after the first consonant of the stem. Nevertheless, I will not discuss any constraints that refers to the place of affixes, because Alignment constraints are so low on the hierarchy that they do not influence the choice of the allomorph.[5]

3.1 Consonant-initial verbs

Table (2) shows the Past Tense of the verbs *səgkow* 'call', *dalud* 'wait' and *sukot* 'ask'. The prefixed candidates **nisəgkow*, **nidalud* and **nisukot* have a full vowel in an unfooted syllable, and are therefore ruled out. The candidates infixed with *-ən-* are too long and violate PRWD=FT. The last candidates are the optimal candidates despite the fact that they violate UNIFORMITY because of their vowel coalescence: these candidates are bisyllabic and consonant-initial.

Affix	*səgkow* 'call'/ *dalud* 'wait'/ *sukot* 'ask'	ONSET	PPN	PRWD=FT	UNIFORMITY
ni-	nisəgkow / nidalud / nisukot		*!	*	
-ən-	sənəgkow/ dənalud / sənukot			*!	
☞-*i-*	sigkow / delud / sikot				*

Table 2 Past Tense of verbs starting with Cə,Ca or Cu

The Dependent allomorphy in Table (3) can be explained in almost the same way: the prefixed candidates **msəgkow* and **mdalud* fatally violate **COMPLONSET; the infixed candidates **səməgkow* and **dəmalud* are too long and the candidates *sugkow* and *dolud* win despite vowel coalescence.

Affix	*səgkow* 'call' / *dalud* 'wait'	*COMPLONSET	PRWD=FT	UNIFORMITY
m-	msəgkow/mdalud	*!		
-əm-	səməgkow /dəmalud		*!	
☞-u-	sugkow /dolud			*

Table 3 Dependent of verbs starting with Cə or Ca

Verbs starting with a consonant followed by the vowel /i/ cannot be infixed with /i/ in the Past Tense because vowel coalescence cannot take place. Two identical adjacent vowels are forbidden, therefore **tiiru'* is out. An anti-homophony constraint is needed for the Past Tense of verbs starting with Ci and the Dependent of verbs starting with Cu.[6]

(14) MORPHDIS: "A segment cannot belong to more than one morpheme." (McCarthy & Prince 1995).[7]

The form **tiru'* is bisyllabic and consonant-initial but nevertheless this form is out because it violates the constraint MorphDis. The optimal candidate is then *təniru'*, even if it is trisyllabic:

Past Tense affix	Stem *tiru'* 'teach'	ONSET	PPN	MORPHDIS	PRWD=FT	UNIFORMITY
ni-	nitiru'		*!		*	
☞-ən-	təniru'				*	
-i-	tiiru'	*!	*		*	
-i-	tiru'			*!		*

Table 4 Past Tense of verbs starting with Ci

The selection of the Dependent allomorph of consonant-initial verbs with a high penultimate vowel, as in Table 5 can be explained in the same way: prefixation with *m-* results in an illicit, clustered onset while infixation with *-u-* in combination with vowel coalescence is impossible here. Therefore the optimal candidate is infixed with *-əm-*, even though it renders the form trisyllabic.

Dependent affix	Stem *sukot* 'ask'	ONSET	PPN	COMPL ONSET	MORPH DIS	PRW D=FT	UNIFORMITY
m-	msukot			*!			
☞-əm-	səmukot					*	
-u-	suukot	*!	*			*	
-u-	sukot				*!		*

Table 5 Dependent of verbs starting with Cu

3.2 *Vowel-initial stems*

All vowel-initial stems are prefixed with *ni-* in the Past Tense and with *m-* in the Dependent, regardless of the quality of their initial vowel. Examples of all possible types are given below.

(15)
stem	gloss	Past Tense	Dependent
əgkot	'work'	*nigkot*	*məgkot*
abput	'bite'	*nebput*	*mabput*
issa'	'put'	*nissa'*	*missa'*
usur	'tell'	*nisur*	*musur*

Consider Table 6. The first, winning, candidate *nesso* only violates UNIFORMITY because of its vowel coalescence. The two other candidates **ənasso* and **esso* lack an onset and are therefore out.

Affix	Stems: əgkot 'work', abput 'bite', issa' 'put', usur 'tell'	ONSET	PRWD =FT	UNIFORMITY
☞ ni-	nigkot /nebput / nissa' / nisur			*
-ən-	ənəgkot / ənabput / ənissa' / ənusur	*!	*	
-i-	igkot / ebput / issa' / isur	*!		*

Table 6 Past Tense of vowel-initial verbs

The Dependent of the verb *asso* 'read' is shown in Table 7. The prefixed candidates *mabput, məgkot* and *missa'* do not violate any constraint whereas the other candidates **əmasso, *əməgkot, *əmissa'* and **osso, *ugkot, *uissa'* lack an onset.

Affix	Stem əgkot 'work', abput 'bite', issa' 'put'	ONSET	PRWD =FT	UNIFORMITY
☞ m-	masso / məgkot / missa'			
-əm-	əmasso / əməgkot / əmissa'	!*	*	
-u-	osso / ugkot / uissa'	!*		*

Table 7 Dependent vowel-initial verbs

Prefixation then, provides vowel-initial stems with an onset and vowel coalescence, in the case of the Past Tense, keeps them bisyllabic.

3.3 *The Past Tense and Dependent affix after stems starting with a liquid*
Verbs starting with a liquid followed by a high vowel often display metathesis of the consonants of the first syllable when they are infixed with *-ən-* or *-əm-*. This metathesis is obligatory for some forms, and optional for other forms. The variation is lexically conditioned, but there may be sociolinguistic influence. Examples are given in (16).

(16)	**stem**	**gloss**	**Past Tense**	**Dependent**
	liug	'swap'	*nəliug*	*məliug*
	riu'	'bathe'	-	*məriu'*
	lauy	'flee'	-	*məlauy*
	riksa'	'examine'	*nəriksa'/rəniksa'*	*məriksa'/rəmiksa'*
	longgo	'lay child to sleep on the floor'	-	*məlonggo*
	lera'	'look after'	*nəlera'/lənera'*	*məlera'/ləmera'*
	runi	'talk'	*rini*	*rəmuni*

Begak is not the only Austronesian language that forbids infixation of an infix that contains a sonorant after a stem-initial sonorant.[8] Chamorro (Klein 2004) and Inonhan (Goudswaard 1998) also have metathesis in this context, whereas

Toba Batak shows assimilation and Tagalog does not parse the infix in this context (Klein 2004).[9]

Metathesis in this context is not an OCP effect of the type "no two sonorant consonants", because it is not just any Sonorant-Vowel-Sonorant sequence that is forbidden here. The reverse order, where the infix is infixed after a non-sonorant-initial consonant but before a sonorant of the stem, C-əm-SonVC, does not allow metathesis, as (17) shows. If metathesis were an OCP effect, we would expect infixation without metathesis in these stems to be ungrammatical as well.[10]

(17)

stem	gloss	infixation	gloss
tulud	'fly'	*təmulud*	'DEP-fly'
tula'	'blame'	*tərula'*	'REC-blame' 'blame each other'
tiru'	'teach'	*təmiru'*	'DEP-teach'
tumis	'stirfry'	*təmumis*	'DEP-stirfry'

In other words, we need to invoke a constraint that refers only to the initial-consonant of the stem:

(18)   *AFF (SON): Sonorant affix avoidance
        "An affix containing a sonorant is prohibited after a morpheme-initial sonorant" (Klein 2002)

This constraint interacts with the faithfulness constraint LINEARITY:

(19)   LINEARITY: S_1 reflects the precedence structure of S_2 and vica versa.
        If x, y $\in S_1$; x', y' $\in S_2$; x R x' and xR y'; then x < y iff x' < y'.(McCarthy and Prince 1995)

The ranking of the constraints is then:

(20)   ONSET, *COMPLONSET, PPN >> MORPHDIS >> PRWD=FT >>
        UNIFORMITY >> *AFF(SON), LINEARITY >>
        ALIGN-BY-SEG (AFFIX, WD, L)

The constraints *AFF(SON) and LINEARITY are unordered with respect to each other, in order to explain the speaker variation between metathesized and non metathesized forms.

Table 8 shows how the metathesized form can win in sonorant-initial stems. The forms prefixed with *ni-* or *m-* or infixed with *-i-* or *-u-* are ruled out for reasons described in the sections above. The form infixed with *-ən-* or *-əm-* violates *AFF(SON) but respects LINEARITY and is therefore optimal in some peoples speech. The metathesized form violates LINEARITY but respects *AFF(SON) and is thus the winning candidate in most people's speech. Table 8 shows the Past Tense of verbs starting with a liquid followed by /i/. The candidate prefixed with *ni-* has a full vowel in the prepenultimate syllable and is therefore out. The last candidate infixed with *-i-* is ruled out because it does not only have a full vowel in the prepenultimate syllable but also violates ONSET (and not shown here, the antihomophony constraint). The candidate

infixed with -ən- and the candidate prefixed with nə- are both longer than a foot and consequently they violate PRWD=FT. The winning candidate is the one prefixed with nə- as it does not violate *AFF(SON), which is ranked higher than LINEARITY.

Past Tense affix	stem *liug* 'swap'	ONSET	PPN	PRWD =FT	*AFF (SON)	LINEARITY	ALIGN-BY-SEG (AFFIX,WD, L)
ni-	niliug		*!	*			
-ən-	ləniug			*	*		*!
☞-ən-	nəliug			*		*	
-i-	liiug	*!	*				*

Table 8 Past Tense of verbs starting with Li

Table 9 shows basically the same picture as Table 8: the first and last candidates are ruled out because they violate PPN and a few other constraints. But this time the winning candidate is not the prefixed form *nətiru'* but the infixed form *təniru'*, as neither candidate violates *AFF(SON). The next constraint is LINEARITY, which is violated by the prefixed, metathesized form *nətiru'*. The winning candidate, then, is infixed form *təniru'*.

Past Tense affix	stem tiru' 'teach'	ONSET	PPN	PRWD =FT	*AFF (SON)	LINEARITY	ALIGN-BY-SEG (AFFIX, WD, L)
ni-	nitiru'		*!	*			
☞-ən-	təniru'			*			*
-ən-	nətiru'			*		*!	
-i-	tiiru'	*!	*	*			*

Table 9 Past Tense of verbs starting with Ci

Table 10 shows how the Dependent form is derived from liquid-initial stems and is almost identical to Table 8:

Dependent affix	Stem *liug* 'swap'	*COMPL ONSET	ONSET	PRWD =FT	*AFF (SON)	LINEARITY	ALIGN-BY-SEG (AFFIX, WD, L)
m-	mliug	*!		*			
-əm-	ləmiug			*	*		*!
☞-əm-	məliug			*		*	
-u-	luiug		*!				*

Table 10 Dependent of verbs starting with Li

The first candidate *məliug* has a complex onset and is therefore out. The last candidate lacks an onset in its second syllable and is therefore out. The

98

candidate infixed with -əm- violates *AFF (SON) and is therefore ruled out. The candidate prefixed with a metathesized form of -əm- wins.

4. Discussion: listed allomorphs or one abstract input IN and UM?

Until now, I have assumed that the Past Tense and Dependent affixes are listed, suppletive allomorphs that cannot be derived from an underlying abstract morpheme, for example *IN or *UM. I will now provide some evidence for this claim. Assuming that the underlying form of the affixes is indeed *IN or *UM, it is easy to derive -ən- and -əm-: the vowels of IN and UM are reduced to schwa because of PPN. The shape of -i- and -u- is also predictable because of bisyllabicity as a maximum. However, the shape of *ni-* and *m-* is unpredictable: there is an asymmetry between the form of the Past Tense prefix *ni-* and the Dependent prefix *m-*. We would expect them to be either both of the shape CV- (*ni-* and *mu-* respectively) or both of the shape C- (*n-* and *m-* respectively), but this is not the case. The shape of the allomorphs is determined by historical processes, because in Mukah Melanau (Blust 1997; Ussishkin 2000) the corresponding morphemes are *nə-* versus *mə-* versus *mu-*: the reverse of Begak.[11] Therefore, a listed allomorph analysis is the best way to explain the data.

5. Conclusion

We have seen that the Begak Past Tense and Dependent allomorphy can be best explained as suppletive allomorphy where the phonology chooses the optimal form. The choice of the optimal allomorphs is directed towards creating consonant-initial bisyllabic words, i.e. words that consist of exactly one foot. Prefixation provides vowel-initial words with an onset and infixation in combination with vowel coalescence keeps consonant-initial words bisyllabic.

Not only prosodic but also segmental constraints play a role in infixation-prefixation alternations: metathesis prevents sonorant affixes to be infixed after a sonorant.

Endnotes

1. The Ida'an Sungai or Subpan have largely intermarried with the people living along the Segama river, who are popularly called the Dusun Segama, and no longer consitute a distinct group.

2. The Past Tense and Dependent verb forms treated in this paper are Undergoer Voice forms by default, as they lack Actor Voice prefixes. The Actor Voice in Begak is marked with Actor Voice prefixes while the Undergoer Voice is characterized by the absence of these prefixes. The Dependent is used for imperatives, verbs of motion, successive actions in stories and after auxiliaries.

3. The optimal prosodic word is always bisyllabic in Begak, not just bimoraic.

4. There is only one exception: the Undergoer Voice Stative prefix *a-* has a full vowel. I have no explanation for this exception; perhaps this vowel receives secundary stress and is therefore footed in some way or another

5. An alternative approach is that infixes can be subcategorized for infixation and the prefixes for prefixation (Yu 2003).

6. There is evidence for anti-homophony in Begak. No transitive active (non-stative) verbal root can start with /b/, /p/, /m/, /n/ or /ŋ/ to avoid homophony with the prefixes *b-*, *p-*, *m-*, *ni-*, and *(mə)ŋ-* respectively. Stative verbal roots can start with nasals or labials but very few stative verbal roots can be affixed with Past Tense or Dependent morphology. If a loan word starts with one of these consonants, it is deleted in the Begak adaptation: *pikir* 'think' from Malay becomes not **məmikir (məng-pikir)* in the AV but *məng-ikir*.

7. This constraint can be ranked anywhere as long as it is higher than PRWD=FT.

8. All examples shown in (16) start with a liquid because Begak does not have any verbal stems starting with a nasal. Another context where infixation is forbidden after a stem initial sonorant or liquid is in the Reciprocal. The Reciprocal is marked by the infix *-ər-* if the stem starts with a consonant. Stems starting with a vowel or with a liquid mark the Reciprocal with a suppletive allomorphic process,:CV reduplication.

9. In the Bisayan language Inonhan the plural actor infix *-Vr-* causes metathesis after an initial liquid. The examples in (i) show the ordinary infixation pattern if the stem does not start with a liquid, although it may contain one:

(i) | *súlat* | 'read' | *nag-surúlat* | 'pres.perf-read-pl' |
     | *pilá* | 'spit' | *nag-piríla* | 'pres.perf-spit-pl' |
     | *sáqot* | 'dance' | *nag-saráqot* | 'pres.perf-dance-pl' |
     | *kánta* | 'sing' | *nag-karánta* | 'pres.perf-sing-pl' |
     | *abót* | 'arrive' | *nag-qarábot* | 'pres.perf-arrive-pl' |

The following (elicited) examples illustrate metathesis after an initial liquid:

(ii) | *líbot* | 'surround' | *nag-rilíbot* | 'pres.perf-surround-pl' |
     | *lohúd* | 'kneel' | *nag-rolóhud* | 'pres.perf-kneel-pl' |
     | *lítson* | 'roast a pig' | *nag-rilítson* | 'pres.perf-roast a pig-pl' |

10. Another Begak infix that cannot occur after a stem-initial liquid is the Reciprocal infix *-ər-.*:

(iii) | *kədtut* | 'pinch' | *k-ər-ədtut* | 'pinch each other' |
      | *kati* | 'tease' | *k-ər-ati* | 'tease each other' |
      | *tadtas* | 'chase' | *t-ər-adtas* | 'chase each other' |

Reciprocals of liquid-initial stems are formed with the suppletive morphological process of prefixation with the AV-prefix *gə-* combined with CV-reduplication:

(iv) | *rakop* | 'wrestle' | *gəgərakop* | 'wrestle with each other' |
     | *lapas* | 'pass' | *gəgəlapas* | 'pass each other by' |

100

langu' 'relative' *gəgəlangu'* 'be relatives of each other'

There is an OCP effect to some extent: the infix -*ər*- cannot occur in stems that contain /r/.

(v)   *danggar*   'bump'   *gəgədanggar*   'bump on each other (on purpose)'
      tiru'   'teach'   *tətiru'*   'teach each other'
      taru'   'put'   *tətaru'*   'put on each other'

The OCP effect of /r/ must be analysed as a distinct phenomenon that is different from the constraint against affixation after sonorants, as the OCP effect applies specifically to /r/ and not to any other non-initial sonorant or liquid: *t-ər-ula'* 'blame each other' from *tula'* 'blame' is perfectly grammatical although the root contains a liquid /l/.

11. In Mukah Melanau, the allomorphs -*i*- and -*u*- replace schwa in the stem, *n*- and *m*- occur before vowel-initial stems, the allomorphs *mə*- and the *mə*- before consonant-initial stems, but *mu*- before labial-initial stems. There is no synchronic explanation for this asymmetry.

References

Alderete, John. 1995. Faithfulness to Prosodic Heads. ms., University of Massachusetts, Amherst. ROA-94-0000

Banker, John E. 1984. The Ida'an language. Languages of Sabah: a survey report, ed. by Julie K. King & John Wayne King. Pacific Linguistics C-78, Canberra: Australian National University.

Beckman, Jill. 1997. Positional Faithfulness. Ph.D. dissertation, University of Massachussetts, Amherst.

Blust, Robert. 1997. Ablaut in Northwest Borneo. Diachronica 14.1.1-30.

Goudswaard, Nelleke. 1998. Inonhan, een Bisaya taal, MA thesis. Dept. of linguistics, Vrije Universiteit Amsterdam

Klein, Thomas B. in press. Infixation and segmental constraints effects: UM and IN in Tagalog, Chamorro and Toba Batak. Lingua 2004, ROA535-0802

McCarthy, John. 2002. Against gradience. ms. University of Massachusetts. Amherst: ROA #510

McCarthy, John & Alan Prince. 1995. Faithfulness and reduplicative identity. University of Massachusetts Occasional Papers in Linguistics 18,

Papers on Optimality Theory GLSA ed. by Jill Beckman, Laura Walsh Dickey and Suzanna Urbanczsyk, 249-384. University of Massachusetts, Amherst.

Moody, David C. 1993. Ida'an phonemics. Sabah Museum Monograph, volume 4, Kota Kinabalu.

Prince, Alan and Paul Smolensky. 1993. Optimality Theory: Constraint Interaction in Generative Grammar. ms. Rutger University and University of Colorado, Boulder, ROA#537-0802 version 8/2002.

Steriade, Donca. 1994a. Licensing by Cue. ms., UCLA.

Steriade, Donca. 1994b. Positional Neutralization and the Expression of Contrast. ms., UCLA.

Ussishkin, Adam. 2000. Fixed effects in Austronesian: an Optimality-Theoretic account. Proceedings of AFLA7, ed. by Marian Klamer. Vrije Universiteit Amsterdam

Alan Yu. 2003. The Morphology and Phonology of Infixation. Ph.D. dissertation. University of California, Berkeley

Zoll, Cheryl. 1998. Positional Asymmetries and Licensing. ms., MIT ROA-282-0998

Nelleke Goudswaard
Vrije Universiteit Amsterdam
Letteren, Taalwetenschap
De Boelelaan 1105, 1081 HV Amsterdam

On Statives and Potentives in Western Austronesian
(Mostly Tagalog)

Nikolaus P. Himmelmann
Ruhr-Universität Bochum

1. Introduction

This contribution is concerned with prefixed forms in western Austronesian languages[1] which have been called a wide variety of names including 'stative', 'accidental', 'involuntary', 'potential', 'coincidence', 'momentary', and so on. Although widely neglected in the literature,[2] these formations are of major import to the grammar of many western Austronesian languages, where for all event expressions there is an obligatory choice between a neutral form and a form marked for 'involuntariness', 'potentiality', 'coincidence', or the like. Furthermore, this distinction has implications for a wide range of theoretical issues, including the nature of unaccusativity and causativity, split-intransitivity, and the grammar of control and complementation.

The main goal of this contribution is to bring some basic order to the fairly broad and, on first sight at least, somewhat heterogeneous range of uses and meanings associated with these forms. I will argue that the different uses can be grouped into two semantically and morphosyntactically quite different construction types, which I will call STATIVE (proper) and POTENTIVE, respectively.

Section 2 presents the major uses of the 'stative' prefix *ma-* in Tagalog. In section 3, it is shown that despite superficial similarities the various examples with *ma*-marked predicates presented in section 2 involve two different constructions and that the prefix *ma-* belongs to two different morphological paradigms. Section 4, finally, provides a systematization of stative and potentive uses and discusses similarities and differences between the Tagalog system and superficially similar systems in so-called split-S languages.

2. Typical uses of *ma-* in Tagalog

In Tagalog, the prefix marking stative and related types of predicates is *ma-* in non-realis formations and *na-* in realis formations. There is also a variant with a long vowel which is orthographically represented as *mà-* (realis *nà-*). It is not unlikely that historically these two variants represent two different formations. But in current Tagalog there is no longer a systematic grammatical distinction between them (cp. Schachter & Otanes 1972:330, among others). As usual in the literature, *ma-* is used here as the citation form of the prefix.

The major uses of *ma-* can be roughly grouped into the following seven semantic classes.

First, *ma-* regularly occurs on property-denoting ('adjectival') predicates, regardless of whether these are used attributively (as in (1)) or predicatively (as in (2)). This usage differs formally from all the remaining uses in that *ma-* here is invariable, i.e. there is no mood alternation (*ma-* vs. *na-*).[3]

PROPERTY

(1)   ang   **ma-liít**   na   hayop
       SPEC   ST-smallness   LK   animal
       'the small animal'

(2)   **ma**-saráp   ang   pag-kain
       ST-satisfaction   SPEC   GER-eating
       'the food was good'

A closely related use is the occurrence of *ma-* on predicates denoting states or changes of state (more precisely: entering into the state denoted by the base). This includes positionals (as in (5)) or locationals (as in (6)). Locationals typically occur only in realis mood (i.e. with *na-*), while all other state expressions occur in both moods.

STATE/ENTERING STATE

(3)   **na**-tà-takot   silá
       RLS.ST-RDP1-fear   3.PL
       'they were afraid (of the snake)'

(4)   **na**-pipe   sya.
       RLS.ST-dumb   3.SG
       'He got dumb.' (Bloomfield 1917:285)

(5)   isá-ng   araw   **na**-upó   syá   sa   taburete
       one-LK   day   RLS.ST-sitting   3.SG   LOC   stool
       'One day he sat down on the chair (between the four pits) ...'
       (Bloomfield 1917:24)

(6)   semantala-ng   sya   'y   **na**-sa tabí   ng   ilog
       meanwhile-LK   3.SG   PM   RLS.ST-LOC side   GEN   river
       'When he was close to the riverside, ...'

A third major usage of *ma-* is with predicates denoting involuntary actions, i.e. eventualities which in principle involve a controlling agent but in the instance at hand this agent lacks full control. Lack of control may pertain to lack of physical control (as in (7) and (8)) but also to the lack of intention. In the latter case the agent performs a controlled action without intending its outcome (as in (9)).

ACCIDENTAL

(7)   **na**-hulog   siyá   sa   kabayo.
       RLS.ST-fall   3.SG   LOC   horse
       'S/he fell from a horse.' (English 1986:664)

(8)   **nà**-ihí'   akó   sa   kà-ta-tawa.
       RLS.ST-urine   1.SG   LOC   ??-RDP-laugh
       'I laughed so hard I wet my pants.' (Wolff et al. 1991:1135)

104

(9)   **na**-dalá   ko   ang   libró
         RLS.POT.PV-carried   1.SG.POSS   SPEC   book
         'I took the book by accident.' (cp. Wolff et al. 1991:285)

Inanimate effectors by definition lack intentions as well as the ability for physical control. Events involving them are thus often also marked with *ma-*, as in (10).

INANIMATE EFFECTOR

(10)   ang   dahun   ay   **na**-dà-dalá   ng   tubig
         SPEC   leaf   PM   RLS.POT.PV-RDP1-carried   GEN   water
         'The leaf was being carried along by the water, ...'

Lack of control also plays a role for non-volitional or spontaneous perception predicates. In Tagalog, these are regularly marked with *ma-*, as in (11).

SPONTANEOUS PERCEPTION

(11)   **nà**-kíta   niyá   ang   dugó'
         RLS.POT.PV-seen   3.SG.POSS   SPEC   blood
         'She saw (happened to see) the blood.'

A sixth use of *ma-* pertains to expressions conveying the ability of an agent to do or achieve something. This may refer purely to the (mental or physical) capabilities inherent in the agent, as in (12) or to the fact that the agent was successful in overcoming difficult circumstances in performing an action (as in (13), cp. English 'manage to' or 'succeed in').

ABILITY

(12)   **nà**-kì-kíta   ba   ninyó   yung   iskinita?
         RLS.POT.PV-RDP1-seen   Q   2.PL.POSS   DIST.LK   street corner
         'Can you (are you able to) see that corner? (Wolff et al 1991:286)

(13)   ay   **na**-kuha   niyá   ang   dahon
         PM   RLS.POT.PV-getting   3.SG.POSS   SPEC   leaf
         'he was able to (managed to) get the leaf'

Finally, *ma-* is also used when asserting (or denying) that the possibility or opportunity to do something exists, regardless of the capabilities of the agent involved in the eventuality. In addition to the following example from a narrative, compare also more or less fixed expressions such as *ma-basa* (POT-reading) 'can be read, legible'.

POSSIBILITY/OPPORTUNITY

(14)   kung   **mà**-bi-bilí   iyán
         if   POT.PV-RDP-sale   MED
         'if that can be sold/if this is sellable'

3. Two different constructions and morphological paradigms: stative and potentive

The examples of *ma-* presented above involve two clearly different constructions and morphological paradigms. This is not immediately obvious when presenting the examples in the way it was done in the previous section (and in much of the literature). But the syntactic differences become obvious when comparing two semantically similar expression types, namely predicates for spontaneous perceptions and for emotions. Both expression types denote mental states of animate experiencers, usually directed at or caused by some entity outside of, or different from, the experiencer. It would thus not be very surprising if these two expression types were to be constructed morphosyntactically in the same or at least a similar way.

Tagalog, however, works differently. Let us first take a closer look at a perception predicate. In the following constructed example (which essentially repeats example (11) above), the perception predicate *kita* 'seen' is prefixed with *na-* and followed by an experiencer expression in possessive (or genitive) case, which in turn is followed by the stimulus (the thing seen) functioning as the subject of the overall construction (marked by the proclitic specific article *ang*).

(15) **nà**-kita   niyá   ang   aso
      RLS.POT.PV-seen 3.SG.POSS SPEC   dog
      'She saw a/the dog.'

The emotion predicate *galit* 'angry' (see also *takot* in (3) above) is constructed quite differently. Here the experiencer is the subject (being a pronoun in (16), it appears in *ang*-form), while the stimulus occurs in locative case marked by the general locative preposition *sa*.

(16) **na**-galit   siyá   sa   aso
      RLS.ST-anger   3.SG   LOC   dog
      'She was angry with the dog.'

A correlated difference pertains to the fact that with perception predicates the stimulus is obligatory in the sense that a stimulus is always understood to be present even if not overtly expressed. With emotion predicates, the stimulus is optional. Altogether, these differences suggest that despite the identical marking on the predicate, we are dealing with two different constructions in which semantic roles are differently aligned with syntactic functions.[4] For reasons that will become obvious shortly, the perception predicate construction is called here a *potentive* construction, while the emotion predicate construction is called a *stative* construction.

Stative and potentive predicates also differ in that they allow for different voice alternations. An alternative way to express the state of affairs in (15) is to use a (potentive) actor voice construction. Here, the experiencer appears as the subject (in *ang*-form) and the stimulus is marked as genitive. Concomitantly, the predicate is prefixed with an actor voice prefix (*naka-*).[5]

(17)   **nakà**-kita   siyá   ng   aso
        RLS.POT.AV-seen   3.SG   GEN   dog
        'She saw a dog.'

An alternative for the emotion construction in (16) is a (stative) locative voice construction, where the predicate is marked with the circumfix *ka—an*. Here the experiencer appears in the possessive form while the stimulus functions as subject (marked by *ang*):

(18)   **k<in>a-galit-an**   niyá   ang   aso
        ST<RLS(UG)>-anger-LV   3.SG.POSS   SPEC   dog
        'She was angry with the dog.'

Table 1 summarizes the alignment differences between the two constructions. The form of the stative predicate which is marked with *ma-* (as in (16) above) is called *basic form*, because this is the most unmarked, frequent and widespread form of stative marking. The role repertoire indicated for each construction includes other possible roles, some of which are exemplified in the examples in the preceding section: Examples (9), (10), (13) and (14) illustrate potentives with agent and theme rather than experiencer and stimulus arguments. Example (7) involves a stative predicate with a theme and a source argument, and (8) one with a theme and a cause argument.

Table 1: Alignment of semantic role and syntactic function in (semantically) transitive potentives and statives

	POTENTIVE	STATIVE	
AV	SUBJ = AGENT/EXPERIENCER GEN = PATIENT/THEME/ STIMULUS	SUBJ = THEME/EXPERIENCER (LOC for SOURCE/GOAL/CAUSE/ STIMULUS)	Basic Form
PV	SUBJ = PATIENT/THEME/ STIMULUS GEN = AGENT/EXPERIENCER	SUBJ = SOURCE/GOAL/CAUSE/ STIMULUS GEN = THEME/EXPERIENCER	LV/CV

The voice alternations illustrated in (17) and (18) are part of the two more extensive morphological paradigms for statives and potentives shown in Table 2. This table also includes the well-known basic voice affixes *-um-*, *-in*, etc., which correlate directly with the potentive forms (see further section 4.2 below). Evidence for the correlations underlying these paradigms is provided in Himmelmann (forthcoming). Here it will be sufficient to take note of the following points.

Table 2: Dynamic and stative paradigms in Tagalog

| | NON-STATIVE (DYNAMIC) | | STATIVE | |
	NON-POTENTIVE	POTENTIVE		
AV	-um-, mag-	maka-	(maka-)	ST.AV
PV	-in	**ma-**	**ma-**	ST
LV	-an	ma--an	ka--an	ST.LV
CV	i-	ma-i-	i-ka-	ST.CV

First, note that potentive and stative formations are identically marked in the second row (both are simply prefixed with *ma-*), but they differ clearly in locative and conveyance voice. Potentives and statives also receive identical marking in actor voice (first row). Here, however, a further difference exists in that stative actor voice is not fully productive, as indicated by the fact that *maka-* appears in parentheses (see Wolff et al. 1991:419f and Himmelmann (forthcoming) for details).

Second, on first sight the occurrence of a *ka-* prefix and the lack of *ma-* in some forms of the stative paradigm may appear to be somewhat unusual. However, it is a well established fact that in many western Austronesian languages *ma-* regularly alternates with *ka-* in a number of grammatically defined environments which are not related to voice. Compare, for example, the following pair of clauses from Mantauran Rukai. In the first clause, the predicate denoting the property of being thick occurs as the single main predicate in clause-initial position and is marked with (stative) *ma-*. In the second clause, the same predicate occurs as the second predicate in a coordinate construction marked by *la*. In this environment, stative predicates in Mantauran Rukai are generally marked by *ka-*.

(19) Example 25 from Zeitoun (2000:429)

a. **ma-**|iʔəməθə   ðona   koapə
   STAT-thick   this   sock
   'This/these sock(s) is/are thick.'

b. ðonaʔi   koapə   ʔa   ma-poli   **la**   *ka-*|iʔəməθə
   that   sock   TOP   STAT-white   and   STAT-thick
   'Those socks, (they) were white and thick.'

Based on the evidence provided by Mantauran Rukai and other languages it has been suggested that, historically at least, the prefix *ma-* is a clipped version of **kuma-*, i.e. *ka-* infixed with *-um-* (cp. Ross 1995:740, Blust 2003:440 passim).

Note, however, that there are also western Austronesian languages where the *ka-* prefix occurs both in the potentive and in the stative paradigm, as illustrated by the paradigms from Ratahan given in Table 3. In this instance, the two paradigms are almost identical (for Ratahan examples and more discussion, see Himmelmann & Wolff 1999:52-63).

Table 3: Dynamic and stative paradigms in Ratahan

	NON-STATIVE (DYNAMIC)			STATIVE	
	NON-POTENTIVE		POTENTIVE		
	PAST	NON-PAST			
AV	-im-	-um-	maka-	??	ST.AV
PV	-in-	-an	-	ma-	ST
LV	-in--an	-an	**ka--an**	**ka--an**	ST.LV
CV	-in-	0	**ka-**	**ka-**	ST.CV

As an aside, it may be noted that the Ratahan data suggest the view that the potentive vs. stative distinction is an innovation which happened after the major branches of Western Malayo-Polynesian (WMP) split. See Zobel (in prep.) for further data and discussion.

So far only constructed example pairs have been used in this section for reasons of clarity and ease of processing. The following examples are added to show that the voice alternations adduced above actually occur in natural discourse.

The pair of attested examples below illustrates the actor voice vs. patient voice alternation for potentive (perception) predicates

(20)   nang   **mà-rinĩg**   itò   ng   Kastila'
         when   POT.PV-audible PRX GEN Spaniard
         'When the Spaniard heard this, ...' (Bloomfield 1917:28/19)

(21)   at   **nakà-rinĩg**   siyá   ng   mga huni   ng   ibon
         and RLS.POT.AV-audible 3.SG   GEN PL   chirping GEN bird
         '... and then he heard some birds chirping.'

The following two pairs of examples illustrate the alternation of basic stative voice and locative stative voice for stative (emotion) predicates:

(22)   **na-tà-tákot**   ako   sa   ahas
         RLS.ST-RDP1-fear 1.SG   LOC snake
         'I am afraid of snakes ...'

(23)   **k<in>a-tà-takut-an**   siyá   ng   mga tao   dito.
         ST<RLS(UG)>-RDP1-fear-LV 3.SG   GEN PL   people   PRX.LOC
         'People here are afraid of him.' (cp. Wolff et al. 1991:699)

(24)   **na-mù-muhî'**   akó   sa   kanyá.
         RLS.ST-RDP1-detestation 1.SG LOC   3.SG.DAT
         'I am disgusted with him/loathe him/despise him.' (English 1986:917)

(25) k<in>a-mù-muhi-án   ko   siyá   sa
     ST<RLS(UG)>-RDP1-detestation-LV   1.SG.POSS 3.SG   LOC
     g<in>awá'   niyá-ng   iyón
     <RLS(UG)>act   3.SG.POSS-LK   DIST
     'I hate him for doing that.' (English 1977:439)

4. Subclassifying stative and potentive uses

While statives and potentives allow a number of different affixations as seen in Table 2 above, forms prefixed with *ma-* are by far the most frequent formations in Tagalog and other Philippine-type languages. Since *ma-*prefixed forms can be either stative or potentive, they are easily confused. But the preceding discussion provides for two independent diagnostics to keep stative and potentive formations apart:

1. Potentive formations with *ma-* allow for genitive-marked arguments (since they are in patient voice they usually will be (semantically) transitive). Statives with *ma-* never allow for genitive-marked arguments.
2. Stative *ma-* alternates with *ka—an* or *ika-*, while potentive *ma-* alternates with *maka-*, *ma—an* or *ma-i-*.

The result of applying these diagnostics to the different uses of Tagalog *ma-* reviewed in section 2 above is the proposal for a systematization of the different subclasses of statives and potentives given in (26).

(26) Subclasses of STATIVES and POTENTIVES in Tagalog

STATIVE (-ACT)

property: have a certain property/quality (*be red/good/small*)
state: be in/get into a state (*be/become broken/flooded/ angry/ afraid/dumb/surprised* etc.); includes *alive/live, dead/die, sleep* (?) and **positionals** (*sit/lie/stand* etc.) and **locationals**; a few (semantically) transitive expressions for feelings: *hate, fed up with, disgust, fear, like* (?)
'unaccusatives': *fall, drown, slide, collapse*, directed motion (?)

POTENTIVE (+ACT)

potential
    abilitative: be able to do sth; succeed in doing
    possibility/opportunity: can be done/possible to do
involuntary/non-intentional
    accidental: do sth by accident, the action is done intentionally but the outcome is not intended (*I took the wrong pills, hit accidentally*, etc.)
    coincidental: happen to do sth without having any prior intentions to do so (*to bump into someone*); includes:
        spontaneous perceptions (*to notice/see/hear/sense, find*)
        spontaneous cognitive acts (*discover, understand, remember, forget*, but see also section 4.1 below)
    inanimate effector (cp. example (10))

110

This classification implies that the distinction between potentives and statives in Tagalog involves two semantic parameters, i.e. the (aspectual) type of eventuality (state vs. event) and control/intentionality (or agency). As amply illustrated by Mithun (1991), these two parameters also play an important role in many languages with Split-S (active/stative) case or person marking systems. Tagalog and other western Austronesian languages with stative and potentive paradigms, however, are crosslinguistically somewhat unusual for the following interrelated reasons, all of which will be discussed in more detail in the ensuing subsections. First, unlike typical Split-S languages which have binary distinctions (agentive/eventive vs. stative), in Tagalog there is a three-way distinction between neutral (i.e. non-potentive, non-stative), potentive and stative forms (see Table 2 above). Second, the distinction is not restricted to intransitive predicates but applies to both (semantically) transitive and intransitive predicates. Third, both semantic parameters (type of eventuality and control/intentionality) play an equally important role but are distributed asymmetrically across the three categories.

Regarding the last point, note that states by definition lack actors (-ACT) and hence are by default non-controlled and non-intentional. However, the category STATIVE in Tagalog is not restricted to states proper but also includes a few events which lack an actor-like core argument (for lack of a better term called *unaccusative* in (26), which is further discussed in the next section). Potentive marking occurs on event-denoting predicates which include an actor (+ACT) but the potentive morphology indicates that this actor lacks full control or intentionality. Tagalog morphology thus suggests that one can downgrade agentivity/intentionality for event predicates without changing the event type or the semantic role structure. This is unusual from a crosslinguistic point of view in that in many split-S languages, lack of control or intentionality usually involves stative marking and hence a change in event type.

In this regard, it should be noted that examples (7)-(9), which from a purely semantic point of view can be characterized as *accidental* eventualities, actually belong to two distinct morphosyntactic classes. Examples (7) and (8) are statives (of the unaccusative variety), which means that they lack an actor-like core argument. Example (9), on the other hand, is potentive (of the accidental variety), i.e. the predicate includes an actor-like core argument but this actor is depicted as lacking intention. Note also that non-controlled (spontaneous) perceptions and other cognitive acts are potentive rather than stative, i.e. the morphology implies that these acts involve an actor (= experiencer) who lacks control rather than completely lacking an actor role.

The classification given in (26) raises a number of further issues, some of which are dealt with in the following subsections.

4.1. Further notes on statives

Stative is a marked category in two regards. On the one hand, it is marked in the superficial morphological sense of involving more morphological marking than non-statives, at least in locative and conveyance voice. On the other hand, it is marked in the functional sense of being the marked member in a binary opposition. Evidence for this status is provided by the fact that not all semantically stative eventualities require stative morphological marking. For example,

all 'copula-like' expressions for 'equal', 'weigh' and 'cost' usually are not marked as statives, as seen in the following example:

(27) t<um>ì-timbáng   akó   ng   110   libra.
       RDP1<AV>-weight   1.SG GEN   pound
       I weigh 110 pounds. (English 1977:1180)

Similarly, it is not the case that all expressions for involuntary bodily actions are marked with *ma-* (as in (8)) above). Instead, lexical bases denoting such actions generally do not occur with stative morphology. Examples include *sinìsinók/nagsìsinók* 'hiccup' and *umubó/ubuhín* 'cough'.

Furthermore, a set of predicates including *tulog* 'sleep' and the positionals allows for stative and non-stative affixation without an obvious difference in meaning (e.g. both *na-tùtulog siyá* and *t<um>ùtulog siyá* mean 's/he is sleeping'). Some positionals actually occur primarily with non-stative affixation (e.g. *l<um>awít* 'be hanging down'). Note that the stative and non-stative forms usually differ in some of their senses (for example, only stative *ma-tulog* can be used in the expression for oversleeping). The point here is that both forms can be used to refer to the state denoted by the base form.[6]

Finally, stative marking is also only one of several options for expressions referring to feelings ('like', 'hate', 'love', etc.). Impressionistically speaking, it seems that negative feelings in particular are referred to with stative predicates while positive feelings are often expressed by predicates with non-stative affixation. But this needs further research and is perhaps only of marginal interest because there is a strong tendency to use both positive and negative feeling predicates *without* any affixation. This holds true in particular for the two high-frequency items *gustó* 'want, like' and *ayaw* 'dislike'. Roughly the same comments apply to predicates for cognitive states, in particular *alam* 'know'.

Of course, rather than saying that the 'exceptions' just mentioned show that STATIVE is a marked category, one could also hold that reference to a state is not a major common semantic denominator of the predicates thus marked. This view would be supported by the fact that the category STATIVE includes a few event-type predicates such as 'fall', etc. (see (7) and (8) above).

Reasoning along these lines, one could entertain the idea of calling the whole class *unaccusative* instead of *stative*. But note that here again one would have to take note of the fact that predicates such as 'come' and 'arrive' which are unaccusative in a great many languages where this distinction is clearly attested are generally not marked with *ma-* in Tagalog (a possible exception are some predicates of directed motion, including *puntá* 'go', which in some uses allow affixation with *ma-* as an alternative to infixation with –*um-*). In this regard, it may also be noted that intransitive achievements are generally non-stative, e.g. *pumutók* 'burst', *sumabog* 'burst/explode', *magkalamat* 'crack' (< *lamat* 'crack').

The use of the term *unaccusative* for 'fall', etc. in the above systematization is somewhat ad-hoc and mainly motivated by the lack of a better alternative. The basic point here is that these predicates are clearly stative according to the two diagnostics given at the beginning of section 4, but they do not

denote states in terms of their aspectual characteristics. A more descriptive name for these predicates would be 'actor-less event predicates'.

4.2. Further notes on potentives

The term *potentive* is an innovation, first used in Rubino 1997. The innovation is motivated by the fact that this category comprises two subclasses, i.e. potential and involuntary, which, on first sight at least, do not appear to have much in common. There is, however, little doubt that this is not a case of accidental homonymy but rather a perhaps unusual, but still coherent category. The two meanings are widely rendered by the same formative in western Austronesian languages, regardless of whether they have two paradigms as in Tagalog or render potentive and a smallish subset of stative meanings with the help of some other morphological marker (e.g. Malay *ter-* and its cognates in many languages of Indonesia). Furthermore, the combination of potential and involuntary meanings is also attested in languages outside the Austronesian family, including the *involitive* in Sinhala (Inman 1993) and *out of control* morphology in Salishan languages (cp. Davis & Demirdache 2000).

A major feature that distinguishes potentives from statives is the fact that there is an absolutely regular and fully productive relation between potentive and neutral (non-potentive) forms: for every potentive form there is a corresponding non-potentive one (with a few exceptions, the reverse also holds) and in terms of their formal make-up, potentive forms are completely predictable on the basis of neutral forms and vice versa (see Schachter & Otanes 1972:331). This (almost) perfect match is the reason why in Table 2 potentive and neutral forms are represented as belonging to a single higher-level category (non-stative/dynamic) despite the fact that potentives and statives are more similar in terms of their formal make-up.

While use of the potentive form is usually obligatory when the actor of a given predicate is not in full control of the event denoted by the predicate, there are a few exceptions to this rule (far fewer though than in the case of statives). Most importantly, perhaps, for some predicates denoting cognitive acts such as *isip* 'think' neutral affixation or no affixation at all is the rule rather than exception. Similarly, not all clauses with inanimate effectors necessarily involve potentive morphology, as seen in the following example:

(28)   ni-lunod   **ng**   **sirena ng**   **bapór** ang
       RLS(UG)-drown   GEN   siren   GEN   boat   SPEC

       s<in>à-sabi   sa   amin
       RDP1<RLS(UG)>-statement   LOC   1.PL.EX.DAT

       ng   dalaga.
       GEN   young.woman
       'The boat's whistle drowned what the young woman was telling us.'
       (English 1986:856)

4.3. Stative and potentive are (largely) orthogonal to lexical categories

It is a matter of debate whether and to what extent lexical bases in Tagalog belong to different lexical categories. There is no need to review this debate here (see Himmelmann (in print) for a recent assessment) because the main

point can be made without having to specify the relevant categories, if any: Stative as well as potentive affixations are not restricted to a specific class of lexical bases but both may, in principle, occur on any lexical base, provided the resulting form makes semantic and pragmatic sense. Thus, for example, lexical bases such as *takot* 'fear' are not restricted to stative formations but also occur in neutral (non-stative) forms as in:

(29)   Huwág   mo-ng   takut-**in**   ang   bata'.
       NEG.IMP   2.SG.POSS-LK   fear-PV   SPEC   child
       'Don't frighten/scare the child!' (English 1986)

(30)   Sino   ang   t<**um**>akot   sa   iyó?
       who   SPEC   <AV>fear   LOC   2.SG.DAT
       'Who frightened you?' (English 1986)

Consequently, one would also expect to find 'minimal pairs' of potentive and stative formations derived from the same lexical base. And this is indeed what we find. In the following two pairs of examples, the first example is potentive, the second stative:

(31)   Hindí   **ma-hulug-an**   ng   karayom   ang   lugar
       NEG   POT-fall-LV   GEN   needle   SPEC   place

       sa   dami   ng   tao.
       LOC   amount   GEN   people
       'One could not drop a needle in the place because of the amount of people.' (google)

(32)   iyón   ang   patibóng   na   **k<in>a-hulug-an**   ni   Gideon
       DIST   SPEC   trap   LK   ST<RLS(UG)>-fall-LV   PN.POSS
       'that's the trap into which Gideon fell' (google)

(33)   **Na-upu-án**   ko   ang   eyeglasses   ko.
       RLS.POT-sitting-LV   1.SG.POSS   SPEC   1.SG.POSS
       'I happened to sit on my eyeglasses.' (google)

(34)   bundók   na   **k<in>a-ù-upu-án**   ng   babae.
       mountain   LK   ST<RLS(UG)>-RDP1-sitting-LV   GEN   woman
       '... the mountains where the women were sitting.' (google)

With regard to 'unaccusatives' such as *hulog* 'fall' it may be of interest to note that here again Tagalog differs from typical split-S languages in that the non-stative form(s) do not mean '(let oneself) fall intentionally' but rather 'make fall, drop'. Self-induced falling is expressed by the prefix combinations *magpati-* or *magpaka-*, both of which contain the causative prefix *pa-*.

4.4. Stative and potentive forms do not constitute a split-S system

In a number of recent publications the proposal has been advanced that some western Austronesian languages are split-S or split-intransitive (or "active") languages in terms of the familiar division of languages into nominative-accu-

sative, ergative, and split-intransitive languages. In typical split-intransitive languages, the main parameter determining the grammatical properties of core arguments in basic transitive and intransitive clauses pertains to the question of whether the participant-role is more ACTOR- or more UNDERGOER-like. That is, the single core argument of some intransitive predicates (S_a) aligns with the more ACTOR-like core argument of transitive predicates (A) while the single core argument of the remaining intransitive predicates (S_o) aligns with the more UNDERGOER-like core argument of transitive predicates (O). Consequently, there is no grammatical relation or function which encompasses one of the core arguments of a transitive predicate and *all* core arguments of intransitive predicates.

It is a matter of debate whether any western Austronesian languages are split-intransitive languages in this sense. The most explicit and convincing case for a split-intransitive analysis has been made by Durie (1987) for Acehnese. In the previous section, it was already pointed out that the stative/potentive system in Tagalog differs from a typical split-intransitive system in a number of important ways. Perhaps the most important point is the fact that stative and potentive marking applies to (semantically) transitive and intransitive predicates alike. Hence, strictly speaking, the stative vs. potentive vs. neutral distinction cannot instantiate a split-intransitive system.

Nevertheless, (semantically) intransitive predicates in Tagalog show, of course, different kinds of morphological marking as seen in the following example pair:

(35) **Natàtakot silá.**   They were frightened.
     Tumàtakbo silá.   They are running (away).

But does this difference in morphological marking indicate a difference in basic clause structure? In the brief definition of split-S systems given above, the crucial characteristic for a split-S system is that it provides for two (or more) different intransitive *clause* structures, i.e. different constituent structures or, in LFG terms, different f- or c-structures. A difference in clause structure would mean that there are different positional regularities for the two clauses (e.g. in one clause the subject may be preposed, in the other it may not) or that the core arguments have clearly differing morphosyntactic properties (different case marking, different control properties, different constraints on zero anaphora, etc.). None of these differences applies to the two clauses in (35). They have exactly the same properties with regard to basic clause structure, the only major difference being that the subject arguments carry different semantic roles. Hence, they do not exemplify a split-S system as defined above.

Put in more general terms, the point here is that split intransitivity involves basic and pervasive differences in clause structure and not just minor, superficial differences in morphological marking. Otherwise the concept of split-intransitivity is in danger of loosing most of its empirical content and all typological relevance. In many, if not all languages stative and eventive intransitive clauses will differ with regard to one or two minor morphosyntactic features, most often in the morphological marking of the predicate. If such differences were to be taken as evidence for split-intransitivity then (almost) all lan-

guages are split-intransitive, which renders it a typologically uninteresting category.

Pursuing this point a bit further, it is in fact also not sufficient to make just any syntactic difference a pre-requisite for split-intransitive status. For example, inasmuch as the difference between unergatives and unaccusatives is considered to be universal and to be essentially a syntactic distinction (predicates with and without internal subject arguments), then again all languages would have to be regarded as split-intransitive languages. But obviously, the idea behind the three-way typological distinction between nominative-accusative, ergative[7] and split-intransitive languages is that these languages differ in fundamental aspects of clause structural organization which cross-cut the distinction between transitive and intransitive clauses.

Endnotes

1. The term *western Austronesian languages* here refers to all non-Oceanic Austronesian languages. See Himmelmann (2004:111f) for more extensive definitions of this and other widely used terms for geographical groupings (e.g. *Philippine languages, languages of western Indonesia*, etc.).
2. Until quite recently, the only work specifically dealing with these formations appears to be Dell (1983), on which Kroeger (1993:80-85, passim) heavily draws. Despite its title, Gerdts (1978) hardly deals with statives or potentives at all but rather with the advancement analysis of Ilokano "passives" (in relational grammar terms). Recently, statives have attracted more attention as seen, for example, in work by Zeitoun (2000), Goddard (2003) and Blust (2003).
3. Unless indicated otherwise, all examples in this paper are taken from natural discourse. Sources are the author's own corpus of spontaneous spoken narratives which includes stories from Wolff et al.'s (1991) textbook, Tagalog websites (coded as *google*) and the texts in Bloomfield (1917). Note that most example sentences in the dictionaries by English (1977, 1986) are from written literary sources. The examples from spoken narratives retain features of the spoken language (in particular common reductions).
4. One way in which this difference could be characterized for the examples given so far is to say that the perception construction is a transitive construction (with *ng* marking a non-subject core argument) and the emotion construction is an intransitive one (with *sa* marking a peripheral argument). However, it is far from clear to what extent (semantic) transitivity is actually grammaticalized in Tagalog and whether the difference between *ng-* and *sa*-marking of non-subject arguments actually correlates with core vs. peripheral status (see Himmelmann 1999:259-261, 2004:147f passim). As we will see shortly, both predicate types allow voice alternations promoting non-subject arguments to subject position

and also behave very much alike with regard to other morphosyntactic characteristics. Furthermore, as also noted later on in the text, all kinds of predicates, semantically transitive as well as intransitive ones, allow both stative and potentive marking. All of this makes it highly doubtful whether (syntactic) transitivity is of major relevance here.

5. Unless other factors interfere, the undergoer in an actor voice construction is interpreted to be indefinite, as indicated in the translation of this example. This is of no import to the point at hand. Note also that the claim frequently made in the literature that undergoers in actor voice constructions are always indefinite or even non-specific is not correct (cp. Himmelmann 2004:148, 172f).

6. The grammar of positionals and locationals is much more complex than indicated here (see Himmelmann (forthcoming) for some additional comments). It is not at all impossible that a more thorough investigation will reveal that it is preferable not to analyze them as statives.

7. As seen in the following quote from Dixon, the term *ergative* has been applied to such a heterogeneous variety of phenomena that it has lost its typological significance in the sense that the feature *ergative* does not correlate with any other morphosyntactic features:

> "What then does it mean for a language to be ergative? Exactly what we said in the first paragraph of Chapter 1- that S is treated in the same way as O and differently from A in some part or parts of the grammar. **Nothing else necessarily accompanies this.**" (Dixon 1994:219, emphasis added)

Using *split-intransitive* in a similarly liberal way for all kinds of marking differences between intransitive predicates will have the same effect of rendering it typologically vacuous.

Abbreviations

ACT	ACTOR	PN	PROPER NOUN
AV	ACTOR VOICE	POSS	POSSESSIVE
CV	CONVEYANCE VOICE	POT	POTENTIVE
DAT	DATIVE	PRX	PROXIMAL
DIST	DISTAL	PV	PATIENT VOICE
EX	EXCLUSIVE	Q	QUESTION MARKER
GEN	GENITIVE	RLS	REALIS
GER	GERUND	RDP	REDUPLICATION (NUMBERS INDICATE DIFFERENT FORMAL TYPES OF REDUPLICATION)
IMP	IMPERATIVE		
LK	LINKER		
LOC	LOCATIVE		
LV	LOCATIVE VOICE	SG	SINGULAR
MED	MEDIAL	SPEC	SPECIFIC ARTICLE
NEG	NEGATION	ST	STATIVE
PL	PLURAL	UG	UNDERGOER
PM	PREDICATE MARKER		

References

Bloomfield, Leonard, 1917, *Tagalog texts with grammatical analysis*, Urbana: The University of Illinois

Blust, Robert A., 2003, 'Three notes on early Austronesian morphology', *Oceanic Linguistics* 42: 438-478

Davis, Henry & Demirdache, Hamida, 2000, 'On lexical verb meanings: Evidence form Salish', in: Tenny & Pustejovsky (eds), 97-142

Dell, Francois, 1983, 'An aspectual distinction in Tagalog', *Oceanic Linguistics* 22-23:175-206

Dixon, R.M.W., 1994, *Ergativity,* Cambridge: Cambridge University Press

Durie, Mark, 1987, 'Grammatical relations in Acehnese', *Studies in Language* 11:365-399

English, Leo J., 1977, *English-Tagalog Dictionary*, Manila: National Book Store

English, Leo J., 1986, *Tagalog-English Dictionary*, Manila: National Book Store

Gerdts, Donna B., 1978, 'Out of control in Ilokano', *Berkeley Linguistic Society* 5:81-93

Goddard, Cliff, 2003, 'Dynamic *ter-* in Malay (Bahasa Melayu). A study in grammatical polysemy', *Studies in Language* 27: 287–322

Himmelmann, Nikolaus P., 1999, 'The lack of zero anaphora and incipient person marking in Tagalog', *Oceanic Linguistics* 38:231-269

Himmelmann, Nikolaus P., 2004, 'The Austronesian languages of Asia and Madagascar: Typological overview', in: Alexander Adelaar & Nikolaus P. Himmelmann (eds), *The Austronesian languages of Asia and Madagascar*, London: Routledge

Himmelmann, Nikolaus P., in print, 'Lexical categories and voice in Tagalog', in: Peter Austin & Simon Musgrave (eds), *Voice and Grammatical Functions in Austronesian Languages*, Stanford: CSLI

Himmelmann, Nikolaus P., forthcoming, 'How to miss a paradigm or two: Multifunctional *ma-* in Tagalog', in: Felix Ameka, Alan Dench and Nicholas Evans (eds) *Catching Language*, Berlin: Mouton de Gruyter [http://www.linguistics.ruhr-uni-bochum.de/~himmelmann/publications.html]

Himmelmann, Nikolaus P. & Wolff, John U., 1999, *Toratán (Ratahan)*, München: Lincom (Languages of the World, Materials 130)

Inman, Michael V., 1993, *Semantics and pragmatics of Colloquial Sinhala involitive verbs*, PhD thesis, Stanford University

Kroeger, Paul R., 1993, *Phrase Structure and Grammatical Relations in Tagalog*, Stanford: Stanford University Press (CSLI publications)

Mithun, Marianne, 1991, 'Active/agentive case marking and its motivations', *Language* 67:510-546

Ross, Malcolm D., 1995, 'Reconstructing Proto-Austronesian verbal morphology: evidence from Taiwan', in: Li, Paul J., Ho, Dah-an, Huang, Ying-kuei and Tsang, Cheng-hwa (eds), *Austronesian studies relating to Taiwan*, 727-791 [Symposium Series of the Institute of History and Philology, Academia Sinica 4. Taipei: Institute of History and Philology, Academia Sinica]

Rubino, Carl, 1997, *A Reference Grammar of Ilocano*, PhD thesis University of California, Santa Barbara

Schachter, Paul & Otanes, Fay, 1972, *Tagalog Reference Grammar*, Berkeley/Los Angeles: University of California Press

Tenny, Carol & Pustejovsky, James (eds), 2000, *Events as grammatical objects*, Stanford: CSLI

Wolff, John U. with Maria Theresa C. Centeno & Der-Hwa V. Rau, 1991, *Pilipino through Self-Instruction*, 4 vols., Ithaca: Cornell Southeast Asia Program

Zeitoun, Elisabeth, 2000, 'Dynamic vs. stative verbs in Mantauran (Rukai)', *Oceanic Linguistics* 39:415-427

Zobel, Erik, in prep., *A diachronic study of Austronesian verb morphosyntax*, PhD Thesis, Universität Frankfurt

Intraposition and Formosan Adverbial Heads

Arthur Holmer
University of Lund

1. Background

From its inception, the adoption of Kayne's (1994) Antisymmetry hypothesis has posed important challenges for the analysis of VOS languages. The initial problems (how to derive the VOS word order itself) were successfully dealt with by among others Pearson (1998), who proposed that VOS order be derived by raising the predicate past subject position. In Massam (2000) predicate raising is analysed as VP raising to SpecTP as the result of a [+pred] EPP feature in (some) verb-initial languages (1a), contrasting with a [+D] EPP feature in SVO languages (1b).

(1) a. $[_{TP}[_{VP}$ **V OBJ**$]_i$ $[_{T'}$ T $[_{vP}$ **SUBJ** $t_i]]]$
     b. $[_{TP}$ **SUBJ**$_i$ $[_{T'}$ T $[_{vP}$ t_i $[_{v'}$ v $[_{VP}$ **V OBJ**$]]]]]$

One problem with this approach is that it does not account for the extremely head-initial nature of VOS languages in general. VOS languages do not only have VOS word order: at some descriptive level, they seem to require a structure which is consistently Head-Complement-Specifier, an option disallowed by the Antisymmetry hypothesis. In fact, as Aldridge (2002) and Holmer (2004) have shown, recursive instances of predicate raising are required simply to allow for the order of the arguments, verbal heads and post-subject particles. The situation is further exacerbated if we include the ordering of adverbs in VOS languages, this issue being the topic of the present paper.

In Cinque (1997) a universal hierarchical ordering of adverbs is suggested, based on a wide range of languages. A subset of this hierarchy, quoted from Rackowski & Travis (2000:121) is shown in (2).

(2)
```
speech act 1
    generally 2
        Neg 3
        already 4
            still 5
               at.all 6
                  anymore 7
                     always 8
                     completely 9
                         well 10    VP
```
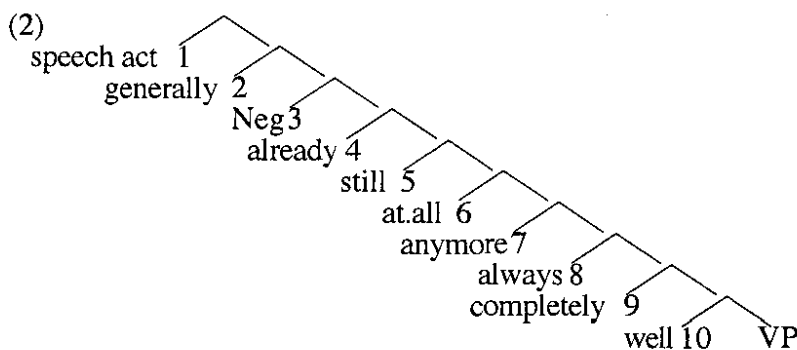
Given Antisymmetry, the hierarchical ordering suggested by Cinque should, *ceteris paribus*, be mirrored by a universal linear ordering. The facts in the VOS language Malagasy, the topic of Rackowski & Travis (2000), are unexpected in this light. Given the numbering in (2), the linear order found in Malagasy is that in (3).

(3) 2 - 3 - 4 - 5 - **VERB** - 10 - 9 - 8 - 7 - 6 - 1 - **SUBJ**

Basically, the Malagasy data conforms linearly to the Cinque hierarchy for preverbal adverbs, whereas the reverse order holds for all postverbal adverbs. This fact leads Rackowski & Travis (2000) to posit a generalized system of predicate raising termed intraposition, where a large portion of the structure is rolled up, as it were, and surfaces in reversed linear order to the right of VP (4).

(4)
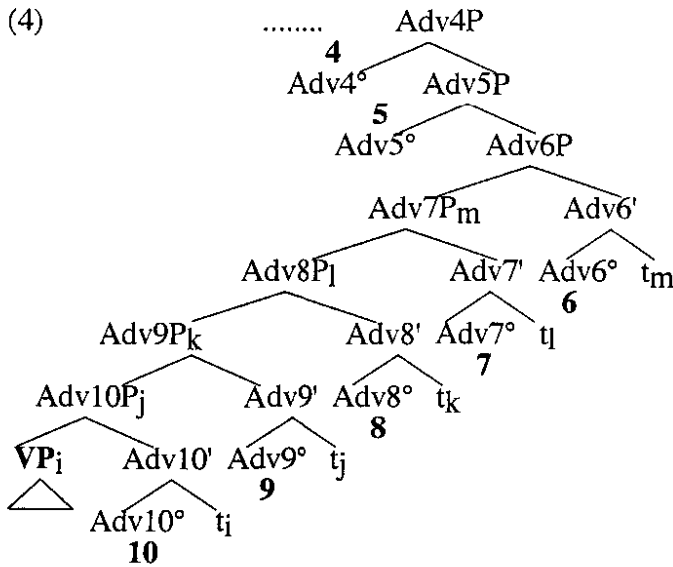
An obvious question at this stage would concern the motivation for movement – Rackowski & Travis' suggestion is that predicate raising is not a feature of a single head (such as T) but rather of every level which is an expansion of VP (e.g. AspP, TP, various AdvP's), possibly extending to the clause as a whole. This would be the defining feature of a predicate raising language as opposed to an argument raising language (cf. Pearson 1998). In a predicate raising language, a head X will trigger raising of its complement to SpecXP (5).

(5) $[_{XP} YP_i [_{X'} X° t_i]]$
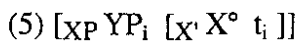
Viewed in this light, the predicate raising mechanism as such is not a problem. Rather, the more serious question would instead seem to be what prevents the process from continuing throughout the whole structure: why are not all Malagasy adverbs postverbal with reverse Cinque order?

The predicate raising mechanism illustrated in (4) and (5) operates around heads, and this leads Rackowski & Travis (2000:122) to suggest that preverbal adverbs are not heads, but are phrasal, and are located in the Specifier positions themselves. The crucial consequence of this is that the specifier position is blocked, thus effectively preventing further predicate raising. Given that the entire analysis crucially rests on the assumption that certain elements are heads and others are phrases, it would be an advantage if some independent evidence for the X / XP status of the elements could be unearthed. Unfortunately, such evidence is hard to come by in Malagasy. However, other Austronesian languages with similar word order patterns do display rather robust evidence for the head status of certain elements. One such language in the Formosan language Seediq.

121

2. Seediq

2.1 Headhood in Seediq

Seediq is an Atayalic language spoken in the area between Puli in central Taiwan and Hualien on the Pacific coast. It is a classic example of a VOS language (6a), although this is sometimes obscured by various other facts, including emphatic subject fronting (6b), the typical Austronesian voice (or "focus") system (6c) and pronominal cliticization (6d).

(6) a. M-n-ekan   bunga   ka   qolic.
      ACTF-PST-eat   sweet.potato   NOM   rat
      'The rat ate sweet potatoes.'

     b. Pawan   (ka)   m-n-imah   sino.
      Pawan   NOM   ACTF-PST-drink   wine
      '*Pawan* (was the one who) drank wine.'

     c. P-n-uq-an   qolic   ka   bunga.
      -PST-eat-LOCF   rat   NOM   sweet.potato
      'A rat ate the sweet potato.'

     d. M-n-imah=ku   sino   kiya.
      ACTF-PST-drink=1SG.NOM   wine   that
      'I drank that wine.'

Most crucially for the present discussion, Seediq also has a similar patterning when it comes to adverb ordering. Relevant examples, with numbering of adverb types as in (2) or (3), are given in (7a-e).

(7) a. ye=su   ini   ekan   hlama?
      INTERR=2SG   NEG   eat.ACTF.CONNEG   *hlama*
      1   **3**   **V**
      'Don't you eat *hlama* [steamed rice with honey]?'

     b. ini   ba   mhmet-i
      NEG   indeed   at.random-PATF.CONNEG
      3   **6**   **10**

        m-angal   pala   q-n-ada   seedaq   sa
        ACTF-take   cloth   -PATF.PST-throw   person   QUOT
        V   **S**   **1**
      '(They) don't just take clothes which (other) people have thrown.'

     c. M-uuyas   ruru   kiya   klaali   heya.
      ACTF-sing   stream   there   always   3SG.NOM
      V   **8**   **S**
      'It (the frog) always sings by the stream.'

     d. Ini=ku   k-qeni   **na.**
      NEG=1SG.NOM   CONNEG-thirsty   NA
      3   **V**   **5**
      'I am not thirsty (yet).'

     e. m-usa   m-ekan   seedaq   kiya   gaga   cghuun   kiya   **di**   **si**[1]
      ACTF-go   ACTF-eat   person   that   be   hang   there   DI   SA
      V   **4**   **1**
      'they (the crows) go and eat the hanged person, so it is said'

122

The pattern for adverbs in Seediq can be summarized as in (8). The basic picture is similar to that in Malagasy: preverbal adverbs appear in Cinque order, and postverbal adverbs appear in reverse Cinque order. One noticeable difference between Seediq and Malagasy is that all postverbal adverbs are presubject in Malagasy, whereas some of them are postsubject in Seediq, the reverse Cinque ordering still being the same as in Malagasy, however. The difference between Seediq and Malagasy could be expressed in intraposition terms such that the raised predicate in Malagasy never contains the subject, whereas in Seediq the subject may be included in the raised structure.

(8) 1 - 3 - 6 - 10 - VERB - 8 - SUBJ - 5 / 4 - 1

From the above, it is clear that there is no a priori reason to assume that adverb ordering is derived differently in Seediq and Malagasy – rather, we expect Rackowski & Travis' analysis to carry over directly to Seediq. For this reason, the head-status tests which Seediq syntax can offer us are directly relevant to the Rackowski & Travis analysis.

In Seediq, pronominal clitics attach to the highest head in the clause. This can be a subordinator (9a) or an interrogative particle (9b), but it can also be a T/A-marker (9c), a negator (9d) or the main verb (9e).

(9) a. **Netun**=su   m-imah   sino,   bsukan=su   dhenu.
        if-2SG.NOM   ACTF-drink   wine   drunk-2SG.NOM   consequently
        'If you drink wine you will get drunk.'
    b. **Ye**=su   m-n-imah   sino   ciga?
        INTERR-2SG.NOM   ACTF-PST-drink   wine   yesterday
        'Did you drink wine yesterday?'
    c. Wada=mu   qta-un   ka   huling=su.
        PST=1SG.GEN   see-PATF   NOM   dog=2SG.GEN
        'I saw your dog.'
    d. Ini=ku   kela   r-m-engo   kari   seediq.
        NEG=1SG.NOM   know.ACTF.CONNEG   -ACTF-talk language   person
        'I can't speak Seediq.'
    e. M-n-ekan=ku   ido   ciga.
        ACTF-PST-eat=1SG.NOM   rice   yesterday
        'I ate rice yesterday.'

It is important to note that, in contrast to Tagalog, the cliticization process does not involve 2nd position Wackernagel clitics. Rather, cliticization is a syntactic process which is sensitive to the status of the host. Cliticization may not take place to conjunctions (10a), nor to wh-phrases (10b).

(10) a Anisa-(*ku)   **ini**=ku   ekan   tmaku.
        but=(*1SG.NOM)   NEG=1SG.NOM   eat.ACTF.CONNEG   tobacco
        'But I don't smoke.'
     b. Ima=(*su)   **q-n-ta-an**=su   ciga?
        who=(*2SG.GEN)   -PST-see-LOCF=2SG.GEN   yesterday
        'Who did you see yesterday?'

123

On the other hand, a wh-word may attract clitics if and only if it is a syntactic head, e.g. the wh-verb *hmuwa / hwaun* 'to do what' (11a). Further, no phrasal adverbials of any kind may attract clitics (11b). It is crucial to note that, in contemporary Seediq, the phrasal adverb *ini huwa* 'it is OK' is not clause-external, as evidenced from the fact that the clitic can climb past it to an interrogative particle (11c)[2].

(11) a. **Hwa-un**=ta   seediq   so   nii   me-eguy?
       do.what-PATF=1PL.INCL   person   like   this   ACTF-steal
       'What shall we do with a thief like this?'

     b. Ini-(*su)   huwa-(*su)   **m-ekan**=su   tmaku   hini.
       NEG   how.ACTF.CONNEG   ACTF-eat=2SG.NOM   tobacco   here
       'It's OK if you smoke here.'

     c. **Ye**=ku   ini   huwa   m-ekan   tmaku?
       INTERR=1SG.NOM   NEG   how.ACTF.CONNEG   ACTF-eat   tobacco
       'Is it OK if I smoke?'

Seediq also displays clear morphological evidence for head-status, e.g. the behaviour of connegatives. The rules governing the use of the connegative are as follows: if there are two verbs present, such as in a control construction, both are realized in normal affirmative morphology (12a). If the clause is negated, the superordinate verb must be realized with connegative morphology (12b, c), which is formally identical, for all voices, with the imperative. However, any verb subordinate to the negated verb may not be realized in connegative form, but must be realized in default morphology (12d). These facts can be summarized graphically as in (12e).

(12) a. m-kela=ku   r-m-engo   kari   seediq
       ACTF-know=1SG.NOM   -ACTF-talk   language   person
       'I can speak Seediq.'

     b. ini=ku   **kela**   r-m-engo   kari   seediq
       NEG=1SG.NOM   ACTF.CONNEG.know   -ACTF-talk   language   person
       'I can't speak Seediq.'

     c. *ini=ku   **m-kela**   r-m-engo   kari   seediq
       NEG=1SG.NOM   ACTF-know   -ACTF-talk   language   person

     d. *ini=ku   kela   rengo   kari seediq.
       NEG=1SG.NOM   ACTF.CONNEG.know   ACTF.CONNEG.talk Seediq

     e. **ini**   **+**   **CONNEG**   **+**   **not** CONNEG

The fact that a verb can be assigned connegative morphology by the negation, and that this morphology can be blocked by an intervening verb, is most easily captured under the assumption that all elements involved in this interaction are heads.

Another piece of evidence concerns the T/A markers, which occur in complementary distribution with the corresponding morphology on the verb. For example, past tense in perfective aspect can be expressed with the particle *wada* (13a) or with overt morphology on the verb (13b), but not with both (13c). If the particle *wada* is used, the main verb must be realized in a form unmarked for T/A (cf. 13a).

(13) a. wada=mu   qta-un   ka   sapah=su
PST=1SG.GEN   see-PATF   NOM   house=2SG.GEN
'I saw your house.'

b. q-n-ta-an=mu   ka   sapah=su
-PST-see-LOCF=1SG.GEN NOM   house=2SG.GEN
'I saw your house.'

c. *wada=mu   q-n-ta-an   ka   sapah=su
PST=1SG.GEN   **-PST**-see-LOCF NOM   house=2SG.GEN

The same occurs with various directional preverbs. If the clause occurs in a certain distinctive voice / "focus" form, this can be realized on the directional preverb, if one is present (14a), or on the main verb, as long as no directional preverb is present (14b), but never on both (14c). If the directional preverb is used, the main verb must be realized in morphology which is unmarked for voice / "focus", i.e. actor focus (ActF), cf (14a).

(14) a. yah-un   m-ekan   qolic   ka   bunga
come-PATF ACTF-eat rat   NOM   sweet.potato
'Rats will come and eat the sweet potatoes.'

b. puq-un   qolic   ka   bunga
eat-PATF   rat   NOM   sweet.potato
'Rats will eat the sweet potatoes.'

c. *yah-un   puq-un   qolic   ka   bunga
come-**PATF** eat-**PATF** rat   NOM   sweet.potato

These features are all typical of the assumed behaviour of heads in syntactic structure. They also hold for a substantial set of elements which have meanings typical of manner adverbials. This class of adverbials display the same kind of behaviour as directional preverbs, in that they can realize the voice morphology which semantically corresponds to the main verb of the clause (15a, b), at the same time preventing the same morphology from being realized on the main verb (15c).

(15) a. **tte-un**=daha   <u>t-m-ekan</u>   ka   macu
to.pieces-PATF=3PL.GEN -ACTF-pound NOM   millet
'They pound the millet to pieces.'

b. <u>tkan-un</u>=daha   ka   macu
pound-PATF=3PL.GEN NOM   millet
'They pound the millet.'

c. *tte-un=daha   <u>tkan-un</u>   ka   macu
to.pieces-PATF=3PL.GEN pound-PATF NOM   millet

Some representative examples of the distribution of voice morphology are given in (16a-d). Example (16d) is particularly illustrative, in that it shows the same verb *tmuting / ttingun* 'beat' occurring in two constructions, providing a near-minimal pair.

(16) a. **skret-an**=daha   <u>m-ekuy</u>   quwaq   salo
      tight-LOCF=3PL.GEN   ACTF-tie   mouth   pot
      'They tie the mouth of the pot tightly.'

   b. **bleq-un**=daha   <u>g-m-emuk</u>
      well-PATF=3PL.GEN   -ACTF-cover
      'They cover it well...'

   c. **gguy-un**=misu   <u>s-m-neru</u>
      secretly-PATF=1SG->2SG   -ACTF-tell
      'I'll startle you.'

   d. **tting-un**=daha   qhuni   ka   qmegi,
      beat-PATF=3PL.GEN   tree   NOM   soapwort,
         nme-un=daha   <u>t-m-uting</u>...
         to.powder-PATF=3PL.GEN   -ACTF-beat
      'They beat the soapwort berries off a tree and pound them to powder.'

Examples (17a-c) show that it is the adverbial head, not the main verb, which receives connegative morphology when combined with the negator *ini*. Likewise, it is the adverbial head, not the main verb, which is realized in imperative if the entire utterance is a command (17d).

(17) a. ini=daha   **mhmet-i**   <u>s-m-ipaq</u>   sa
      NEG=3PL.GEN   needlessly-PATF.CONNEG   -ACTF-kill   QUOT
      'It is said that they don't kill them (100-pacer snakes) needlessly.'

   b. ini   **burux**   <u>m-ekan</u>   ka   seediq   cbeyo
      NEG   alone.CONNEG   ACTF-eat   NOM   person   long.ago
      'The people of old didn't eat alone.'

   c. ini=daha   **trmex-i**   <u>m-ekan</u>...
      NEG=3PL.GEN   on.its.own-PATF.CONNEG   ACTF-eat
      'They don't eat it (the chili) on its own... (because it's so hot).'

   d. **k-tengi** hari   <u>m-ekan</u>!
      IMP-full a.bit   ACTF-eat
      'Eat some more! / Eat until you are more full.'

Most elements which occupy this position convey meanings typically corresponding to adverbs of manner. However, this group also contains a handful of adverbs referring to frequency and duration (18a, b).

(18) a. Ini=daha   **kntte-i**   <u>m-ekan</u> beras   baso,
      NEG=3PL.GEN   often-PATF.CONNEG   ACTF-eat grain   baso
         pcnga-un=daha   <u>m-ekan</u>
         sometimes-PATF=3PL.GEN   ACTF-eat
      'They don't eat *baso* grain often, they eat it occasionally.'

   b. ini=daha   **qbsyaq-i**   <u>pure</u>   heya
      NEG=3PL.GEN   long.time-PATF.CONNEG   (ACTF).cook   3SG.NOM
      'they don't cook it (taro) too long.'

At least for the above sets of manner adverbs, it is difficult to conclude that they can be anything other than heads. Other preverbal elements with adverbial meanings also seem to display head characteristics, although less clearly, due to

the fact that they are located structurally higher than the most relevant types of morphology. One clear instance is the negator *ini*, which triggers connegative morphology and attracts clitics (19a). Another element which attracts clitics is the particle *tena* 'already' which is located preceding the tensed verb, and which for this reason is not involved in any morphological processes typical to verbs (19b). The status of *tena* as a head can only be surmised from the fact that it attracts clitics, which may be subject to debate, although it has been argued here that clitic attraction in Seediq is an exclusive property of heads. At least one preverbal element is clearly not a head according to the above criteria, namely *ba* 'indeed', given that it can intervene linearly between the negator *ini* and the main verb without blocking the assignation of connegative morphology (19c).

(19) a. Ini=mu   qta-i   ka   quyu   kiya
        NEG=3SG.GEN   see-PATF.CONNEG   NOM   snake   that
        'I didn't see that snake.'

     b. Tena=ku   m-n-ekan   ido.
        already=1SG.NOM   ACTF-PST-eat   rice
        'I've already eaten.'

     c. Kiya ini=daha   ba   tleng-i
        thus   NEG=3PL.GEN   indeed   touch-PATF.CONNEG
           ka   seedaq   m-n-cghuun.
           NOM   person   ACTF-PST-hang
        'they certainly don't touch people who have hanged themselves.'

2.2 The analysis

The conclusion we can draw from the above is that preverbal adverbial meanings in Seediq can be expressed by both heads and non-heads. It follows that the Rackowski & Travis' (2000) account of what prevents predicate raising past preverbal adverbs does not carry over to Seediq. In fact, given that head status evidence is visible in Seediq but not in Malagasy, it could be argued that the Seediq evidence weakens Rackowski & Travis' arguments for Malagasy as well. Therefore, at least for Seediq, an alternative analysis accounting for the presence vs. absence of predicate raising must be found.

One possibility might be to posit that preverbal adverbs are heads and postverbal adverbs are phrases, in analogy with Shlonsky's (2003) analysis of Hebrew NP's. However, one problem with this view is that postverbal adverbs (in particular final particles) can not be expanded (a behaviour otherwise typical of phrasal elements, cf. the discussion in Holmer 2004). An illustrative example of the expansion argument concerns negation in English and German. In English, *not* is a head, since it triggers do-insertion (in certain versions of the theory, cf. Ouhalla 1991, by virtue of blocking verb movement), whereas German *nicht* is phrasal (presumably in the Specifier of NegP) and does not block verb movement. English *not* cannot be expanded (in its canonical position and with its canonical behaviour) into complex negations such as 'not at all' or 'never', whereas German *nicht* can easily be replaced by *überhaupt nicht* 'not at all', *nie im Leben* 'never ever (lit. never in one's life)' and many other types of complex negations. In cases of doubt, I assume here that the impossibility of phrasal expansion can serve as evidence for head status.

Under the assumption that postverbal adverbs (including final particles) are

also heads, another problem arises, namely how it can be that these heads, if they are heads, do not block head movement according to the Head Movement Constraint. There is plausible evidence for head movement (V->C) in Seediq, given that clitics, which are crucially not 2nd position clitics, can attach to either subordinators (20a) and verbs (20b), depending on which is first. Here the default assumption would be that they attach to a single discrete head (i.e. C°) and that head movement ensures that a C° which is not filled by a subordinator is lexically filled by movement instead (cf. Holmberg & Platzack's (1995) analysis of V2 languages).

(20) a. Netun=su   ini   ekan   ido...
        if=2SG.NOM   NEG   eat.ACTF.CONNEG   rice
        'If you don't eat...'
     b. M-n-ekan=ku   ido   ciga.
        ACTF-PST-eat=1SG.NOM   rice   yesterday
        'I ate rice yesterday.'

If we discount cliticization, other evidence of head movement is hard to come by, given the extremely head-initial nature of the language. One possible case might be the behaviour of *ba* 'indeed'. In (21a), *ba* is located to the left of the manner adverb, whereas in (21b) is is located to the right of the manner adverb. These two examples can be captured under a head movement analysis assuming that *ba* is not a head and that the adverb in (21b) moves past it because the position to the left of *ba* is not occupied by any other element.

(21) a. ini   ba   mhmet-i   m-angal   pala
        NEG   indeed   at.random-PATF.CONNEG   ACTF-take   cloth
        'they certainly don't just take clothes ...'
     b. blequn=daha   ba   s-m-netun   ma   rees-un=daha
        well=3PL.GEN   indeed   -ACTF-follow and   bury-PATF=3PL.GEN
        'they observe (the law) meticulously and bury them'

Even if overt verb movement could be shown not to occur, the relation between the T/A markers and the overt morphology of the verb, namely that they cannot cooccur (22a-c), suggest that they are checked in the same syntactic position, if not overtly, then at least covertly. The intervening postverbal adverbs do not block this covert checking relation either.[3]

(22) a. wada=mu   qta-un   ka   huling
        PST=1S.GEN   see-PATF   NOM   dog
        'I saw the dog.'
     b. q-n-ta-an=mu   ka   huling
        -PST-see-LOCF=1S.GEN   NOM   dog
        'I saw the dog.'
     c. *wada=mu   q-n-ta-an   ka   huling
        PST=1S.GEN   -PST-see-LOCF   NOM   dog

For the reasons outlined above, the relation between head-raising and possible intervening heads seems to be a serious problem which must be addressed.

Apparent violations of the Head Movement Constraint are not rare among the world's languages. Possibly the most publicized phenomenon of this type is Long Head Movement (cf. Borsley et al. 1996). Carnie, Harley & Pyatt (2000) discuss similar facts in Old Irish and suggest, following Borsley et al. (1996), that heads, in analogy with phrasal positions, can also be classified in terms of the A / A' distinction. Just as a filled A position does not block A'-movement of a phrase, neither does a filled A head position block movement to an A'-head position.

In analogy with this distinction, we suggest that heads in Seediq can be of two classes, one class which is involved in head movement and which is typically associated with left-marginal or postverbal position, and one class which triggers predicate raising, which is in no way involved in head raising, and which is typically associated with postverbal position. Whether or this distinction is analogous to the A / A' distinction as proposed by Carnie, Harley & Pratt (2000) is an open question. Certainly, neither can be derived straightforwardly from the other. For instance, neither of the two types of head in Seediq is particularly or exclusively connected with C° – in fact, conditional subordination itself can be expressed either by the clause-initial head *netun* 'if' which blocks head movement (23a, b) or by the final particle *do* 'if' (23c).

(23) a. m-imah=su   sino   klaali
         ACTF-drink=2SG.NOM   wine   always
         'You drink wine all the time.'
     b. netun=su   m-imah(*=su)   sino   klaali...
         if=2SG.NOM   ACTF-drink(*=2SG.GEN)   wine   always
         'If you drink wine all the time...'
     c. m-imah=su   sino   klaali   do...
         ACTF-drink=2SG.NOM   wine   always   COND
         'If you drink wine all the time...'

At this stage, we shall refer to the two types of head as X-heads and Y-heads respectively, deferring to future research the issue of whether this distinction has any common denominator with the A / A'-distinction. The properties of the two types of heads are given in (24).

(24) X heads: undergo / block head-raising
     Y heads: irrelevant for head-raising; trigger PRED-raising

Given that X heads are involved in head raising, it is naturally only X-heads which will be able to host typical verbal morphology. Thus a further property of X-heads crystallizes, namely that they can be part of the verbal system of morphology (behaving like prototypical heads in the sense we are used to from European languages), something which never occurs for Y-heads.[4]

In Holmer (2004) the distinction between X-heads and Y-heads was harnessed to account for the fact that final particles (Y-heads) do not block verb movement. This analysis can be extended to account for the ordering of adverbs in Seediq as well. If we follow Rackowski & Travis (2000) in analysing postverbal adverbs as heads, the distinction between preverbal and postverbal heads can easily be reduced to a distinction between X-heads (preverbal) and Y-

heads (postverbal). Final particles are simply a special instance of Y-heads, hierarchically so high in the structure that the raised predicate includes the position containing the grammatical subject, whereas postverbal (but pre-subject) adverbs are located lower in the structure (as we indeed would expect, given their relation to the Cinque hierarchy), so that the raised predicate does not include the subject.

The proposed distinction thus allows us to apply Rackowski & Travis' (2000) predicate raising analysis to Seediq postverbal adverbs, and to reconcile this analysis with morphological evidence from preverbal adverbs. In the narrowest sense, it only claims to account for data in Seediq (and other languages where manner adverbs are realized as verbs). However, the same analysis can be carried over to Malagasy and other VOS languages with similar adverb ordering, under the assumption that the distinction between X-heads and Y-heads can be maintained without morphological desinences.

3. Geographical and genetic distribution

One question to be asked at this stage is how widely spread this phenomenon really is: where else, outside Seediq do we find morphological head evidence for preverbal adverbs? In the closest relatives of Seediq, namely the other Atayalic languages in Taiwan, Squliq Atayal and C'uli Atayal, we find much the same phenomenon, behaving exactly as in Seediq. Examples (25a-c) illustrate this for Squliq Atayal, and examples (25d-f) illustrate it for C'uli Atayal.

(25) a. **m-in-glu**=ta?   <u>m-aniq</u>   hira?
     ACTF-PST-together=1PL.INCL   ACTF-eat   yesterday
     'We ate together yesterday.' (Huang 1993: 90)

    b. *****m-in-glu**=ta?   <u>m-in-aniq</u>   hira?
     ACTF-PST-together=1PL.INCL   ACTF-PST-eat   yesterday

    c. **leq-un**=maku?   <u>m-ita?</u>
     careful-PATF=1SG.GEN   ACTF-see
     'It examined (it) carefully.' (Huang 1993: 90)

    d. **naqaru-un**=mi'   <u>ma-bahuq</u>   ku'   situing   la
     finish-PATF=1SG.GEN   ACTF-wash   NOM   clothes   PRT
     'I have finished washing the clothes.' (Huang 1995: 193)

    e. **lihka**=ci'   <u>ma-ktaliyun</u>
     fast=1SG.NOM   ACTF-run
     'I run fast.' (Huang 1995: 195)

    f. **si-iwan**'i'   <u>ma-quwas</u> ni'   yumin   'i'   limuy
     INSF-replace   LINK   ACTF-sing   GEN   Yumin   NOM   Limuy
     'Yumin sang instead of Limuy.' (Huang 1995: 194)

A similar pattern is found in the Paiwanic language Bunun, also spoken in Taiwan (Jeng 1977), as shown in examples (26a, b).

(26) a. **qasmav-un**   ?ista   <u>ma-tas?i?</u>   palangan
     diligent-PATF   3SG.GEN   ACTF-make   rattan-basket
     'He is diligent making rattan baskets.' (Jeng 1977:210)

130

b. **ma-qasmav**   ?aipa?   <u>ma-tas?i?</u>   palangan
   ACTF-diligent   3SG.NOM   ACTF-make   rattan-basket
   'He is diligent making rattan baskets.' (Jeng 1977:205)

In the third Formosan group, Tsouic, similar facts obtain, albeit with an important difference. We have noted that in the Atayalic languages and (at least some) Paiwanic languages verbal morphology on the adverb precludes morphological distinctions on the main verb, which is instead realized in default ActF. In Tsou, on the other hand, both the adverb and the verb display the focus morphology corresponding to the whole clause. Thus, the verbal morphology on the adverb and the main verb agrees, although the full distinction is only realized on the main verb, the adverb only realizing a defective distinction between ActF and non-ActF (glossed here as UNDF 'Undergoer Focus', following the source of the examples, Szakos 1994). Relevant examples are given in (27a, b).

(27) a. Ø-o-si-cu   **aha'-va**   <u>eh-tothom-neni</u>
       UNDF-PR-3-PERF sudden-UNDF   against-fight-BENF
            <u>le-tothom-neni</u>   na   'e   eatatiskova
            hit-fight-BENF   ART   DEM   person
       'She suddenly attacked the man and fought him.' (op.cit.2)
  b. m-oh-cu   **aha'o**   <u>mi-hcihci</u>   ho
       ACTF-PST-PERF sudden-ACTF   ACTF-teethbare   &
            <u>mi-se'u</u>   to   t'nι
            ACTF-grimace   LOC   cliff (op.cit.6)
       'Suddenly she bared her teeth and grimaced towards the cliff.'

Thus, all Formosan branches of Austronesian display this phenomenon to a certain extent. The facts in Formosan languages can be summarized as in (28).

(28) - Atayalic: full voice on ADV, default on V
     - Paiwanic: full voice on ADV, default on V
     - Tsouic: defective voice agreement on ADV, full voice on V

Outside Formosa, the phenomenon is rarer, although not non-existent. In Tagalog, only the linear order and the occurrence of the ligature *na / nang* between the manner adverb and the remainder of the clause points to what may be predicative status for the manner adverb.

(29) a. **mabilis**na <u>naglakad</u> si   Pedro
      quick   LIG   walk NOM Pedro
      'Pedro walked quickly' (Schachter & Otanes 1972: op.cit. 436)
  b. <u>naglakad</u>   si   Pedro   nang   **mabilis**
      walk   NOM   Pedro   LIG quick
      'Pedro walked quickly' (Schachter & Otanes 1972: 436)

However, in Tukang Besi (Mark Donohue, p.c.) the perfective morpheme *mo* 'PRF' which is prototypically attached to verbs (30a), intstead attaches to an adverb of a certain class if one is present (30b, c). Attaching it to the main verb instead of to the adverb varies from ungrammatical to marginal (30d, e), whereas

attaching it to both is ungrammatical (30f, g). This is irrespective of the possible verbal status of the adverb, since it holds even for adverbs which do not allow predication in their own right (30h).

(30) a. no-tinti=mo
3SG-run-PRF
'S/he ran.'

    b. no-menti'i=mo   no-tinti
3SG-quickly=PRF 3SG-run
'S/he ran quickly.'

    c. po'oli=mo   no-tinti
already=PRF 3SG-run
'S/he has already run.'

    d. ?*no-menti'i   no-tinti=mo
3SG-quickly   3SG-run-PRF

    e. ?po'oli   no-tinti=mo
already   3SG-run=PRF

    f. ?*no-menti'i=mo   no-tinti=mo
3SG-quickly=PRF 3SG-run=PRF

    g. *po'oli=mo   no-tinti=mo
already=PRF 3SG-run=PRF

    h. *no-po'oli
3SG-already

Admittedly, the categories which are realized on the adverb are not the same in Tukang Besi (perfective) and Formosan languages (voice), but the existence of the phenomenon in a relatively wide genetic space among Austronesian languages may be an indication that it is part of the linguistic inheritance of Austronesian languages, despite the fact that most Austronesian languages have lost the morphological desinences relating to it. Under such an assumption, the ordering of adverbs in Malgasy falls out naturally as the consequence of a distinction between two kinds of heads which in itself serves to permit the cooccurrence of head-raising and predicate raising in the same language.

4. Conclusion

To summarize, we have argued that the problems of adverb ordering in Malagasy and other typologically similar VOS languages can be solved by means of a classification of heads into two major categories for which we have morphosyntactic evidence in at least some Austronesian languages, and that it is not necesary to make use of a morphosyntactically unmotivated distinction between XP vs. X° status for adverbs to block intraposition in either Seediq or Malagasy in order to derive the correct linear order. The exact nature of X-heads and Y-heads and their possible connection with an A / A'-distinction as outlined by Carnie, Harley & Pratt (2000) is an issue which we defer to future research.

Endnotes

1. [si] and [s] are idiolectal variants of *sa*.
2. Although it is likely that it derives etymologically from a matrix clause taking the remainder of the utterance as its complement.
3. The question might be posed whether a checking relation must necessarily be local. However, in a language like English, it appears to be the presence of a Neg head between V and T which forces do-insertion to realize tense, whereas non-local checking could lead to clauses like '*He not drinks beer.'
4. Given this fact, Sandra Chung (p.c.) has suggested an alternative analysis, following Chang (2004), which dispenses with the X/Y distinction, instead reducing similar facts in the Formosan language Kavalan to a generalization that adverbial meanings are grammatically encoded as verbs in languages of this type, and that lexical verbs behave in a different way from other types of head. While it is clear that manner adverbs in languages such as Seediq and Kavalan *are* verbs in the very relevant sense of being able to bear verbal morphology, as well as possibly expressing the primary semantic predication in the clause (as discussed in Holmer 2002), this does not allow us to do away with X-heads altogether: over and above manner adverbs / verbs, X-heads include tense markers, negators and some (but not all) subordinators: these can hardly be considered to be verbs under any analysis, yet they share the typical X-head trait of being involved in head movement and thereby being realized preverbally. Further, while the verbal status of manner adverbs is clear in a language like Seediq where these elements can bear verbal morphology, it is less clear in a language like Malagasy, where preverbal adverbs do not display any traits typical of lexical verbs. At the same time, the present analysis allows us to account for the largely identical adverb ordering in both Seediq and Malagasy at the same time, making use of exactly the same mechanism. This generalization would be lost under a lexical approach. Instead, I favour the idea that X-heads which have lexical content are indistinguishable from verbs by virtue of the very fact that they have both lexical content and verbal morphology, i.e. that they are verbs in a very real sense, but as a consequence of other properties rather than as a primitive.

References

Aldridge, Edith. 2002. Nominalization and WH-movement in Seediq and Tagalog. *Language and Linguistics* 3.2: 393-426. Taipei: Academia Sinica

Borsley, Robert, Maria-Luisa Rivero & Janig Stephens. 1996. Long Head Movement in Breton. In Ian Roberts & Robert Borsley (eds). *The syntax of the Celtic languages*. Cambrudge: CUP. 53–74

Carnie, Andrew, Harley, Heidi & Pyatt, Elizabeth. 2000. VSO order as raising out of IP? Some evidence from Old Irish. In Andrew Carnie & Eithne Guilfoyle (eds). *The syntax of verb-initial languages*. Oxford: OUP. 39–59

Chang Yung-Li. 2004. The guest playing host: Modifiers as matrix verbs in Kavalan. Paper presented at AFLA XI, Berlin, April 2004.

Cinque, Guglielmo. 1997. *Adverbs and functional heads. A cross-linguistic perspective*. Oxford: Oxford University Press

Holmer, Arthur. forthcoming 2004. Seediq – antisymmetry and final particles in a Formosan VOS language. Andrew Carnie, Sheila Dooley Collberg & Heidi Harley (eds). *Verb First: Papers on the Syntax of Verb Initial Languages.* John Benjamins.

Huang. Lillian. 1993. *A study of Atayal syntax.* Taipei: Crane Publishing Co.

Huang, Lillian. 1995. *A study of Mayrinax syntax.* Taipei: Crane Publishing Co.

Jeng, Heng-hsiung. 1977. Topic and focus in Bunun. (Inst of History and Philology, Academia Sinica Special Publication no. 72). Taipei: Academia Sinica

Massam, Diane. 2000. VSO and VOS. Aspects of Niuean word order. In Andrew Carnie & Eithne Guilfoyle (eds). *The syntax of verb-initial languages.* Oxford: OUP. 97-116

Pearson, Matt. 1998. Predicate raising and VOS order in Malagasy. *UCLA Occasional Papers in Linguistics* 20: 94-110

Rackowski, Andrea & Travis, Lisa. 2000. V-initial languages: X or XP movement and adverbial placement. In Andrew Carnie & Eithne Guilfoyle (eds). *The syntax of verb-initial languages.* Oxford: OUP. 117–141.

Shlonsky, Ur. 2003. The form of Semitic noun phrases. to appear in *Lingua.*

Szakos, Jószef. 1994. *Die Sprache der Cou. Untersuchungen zur Synchronie einer austronesischen Sprache auf Taiwan.* (2. Teil: Texte und lexikalische Analyse). PhD thesis. Friedrich-Wilhelms-Universität, Bonn.

Schachter, Paul & Fe Otanes. 1972. *Tagalog reference grammar.* Berkeley: University of California Press.

Arthur Holmer
Lund University
Dept of Linguistics and Phonetics
Centre for Languages and Literature
Box 201, SE-221 00 Lund
Sweden
arthur.holmer@ling.lu.se

Deriving the Order of Heads and Adjuncts: The Case of Niuean DPs

Arsalan Kahnemuyipour and Diane Massam
Syracuse University/University of Toronto

0. Introduction

This paper examines substantive noun phrases in Niuean, a Polynesian language of the Tongic subgroup with VSO word order, isolating morphology, and an ergative case system. We describe the allowable orderings of elements in the Niuean noun phrase, which include certain variations in the placement of numerals and the genitive possessor, then we provide a phrasal movement analysis for these variations, treating first the possessor variation, then the numeral variation. Parallels will be drawn between the derivation of nominal and sentential word order.

There has been a large quantity of work, both recent and traditional, attempting to understand why certain orders of elements in clauses seem to be universally ruled out. To account for this, some linguists have posited that there is a universal order of elements and that allowable variations on this order are derived by various movement patterns (e.g. Cinque 1996, 1999, 2000, 2002, Belleti 2004, Rizzi 1997, 2003). This position allows for the theory to rule out ungrammatical orders by universal constraints on movement, rather than by typological stipulations. Our exploration of the Niuean DP takes place in light of this type of work, as well as similar work specifically on DPs such as den Dikken 2003, Pearce 2002, 2003. and Shlonsky 2004. Given that Niuean DPs have a N-initial order, similarly to the V-initial order of clauses, our analysis also addresses the relation between nominal structure and sentential structure, finding striking parallels between the two clause-types. We confirm a movement constraint observed by Rackowski and Travis (2000), which states that purely relational functional projections such as Agreement, which have no semantic content, are invisible to certain types of movement.

1. Description of the Niuean DP

The Niuean DP is described in Seiter (1980), and in Massam and Sperlich (2000). In a DP without a possessor or numeral, the order of elements is as shown in (1) below. First, there is a portmanteau morpheme, which indicates the case of the DP as well as whether it is common or proper (where proper includes pronominal). In (1a,b) this particle is *e* (absolutive common), whereas in (1e) it is *a* (absolutive proper). (Ergative common DPs begin with the particle *he*, and ergative proper DPs begin with *e*.) This is followed by an optional marker for number, which also has classifier-like properties, as can be seen in (1a) and (1b) where a different plural marker appears depending on the nature of the noun or group. Other plural classifiers include *lafu* for a family group, *atu* for a row, and *na:* for a pair. (For a discussion of the relation between classifiers and number markers, see Fassi Fehri

and Vinet, 2004). This marker can also have article-like properties, in that an indefinite plural NP can appear with *falu (a)*, as in (1c) and in the case of a singular indefinite NP, the marker *taha* or *ha* can appear in this position, as in (1d). We will refer to this complex morpheme simply as number (#) in this paper. This morpheme is followed by the head noun, which is in turn optionally followed by one or more adjectives as in (1a) and a demonstrative as in (1a). (1e) shows a proper DP. The order of elements in the DP is fixed.

(1) Order of Elements without Possessors and Numerals
     (C=common P=proper/pronoun, Colon indicates a long vowel)

Case+P/C	#	Noun	Adjs	Dem

a. e   tau   manu kula fulufuluola   e:
   AbsC Pl   bird   red   beautiful   Dem
   'those beautiful red birds' (Field Notes.01)

b. e   kau   kaiha
   AbsC group thieves
   'a group of thieves' (Seiter 1980.100a)

c. e   falu (a)   tagata
   AbsC some   person
   'some people' (Sperlich 1997.67)

d. e   taha   tagata
   AbsC one   person
   'a person' (Field Notes.01)

e. a   Moka
   AbsP Moka
   'Moka' (Field Notes.97)

In DPs with possessors, the situation is a little more complex. (Examples with numerals will be discussed further below.) There are two possible orders, as shown in (2). The first order finds the genitive case marked possessor in pre-nominal position. In this order, there is a ligature item *a* appearing between the possessor and the noun, as in (2a). The second order finds the genitive marked possessor at the end of the entire DP (after the demonstrative if there is one), as in (2b).

(2) Orders of Elements with Possessors

		#	Noun	Adjs	Dem	
a. Case+P/C	**Poss *a***	#	Noun	Adjs	Dem	
b. Case+P/C		#	Noun	Adjs	Dem	**Poss**

a'. e   ha   Sione a   leo
    AbsC GenP Sione *a* voice
    'Sione's voice' (Seiter 1980.92b)
b'. e   leo   ha   Sione
    AbsC voice GenP Sione
    'Sione's voice/voice of Sione' (Field Notes.97)

The pre-nominal possessive construction has two particular properties distinct from the properties of the construction with the possessor at the end of the clause. First, the pre-nominal possessor gives a definiteness reading to the DP as a whole, similarly to the situation in Hebrew and Arabic (see, e.g. Ritter 1988, Shlonsky 1988, Borer 1999), as shown in (3). (3a) has a definite reading, whereas (3b), like non-possessed Niuean DPs, can be definite or indefinite.

(3) a. ko e   haana a   fale
     Pred   his   a   house
     'It's his house.' [definite] (Sperlich 1997.103)
   b. ko e   fale   haana
     Pred   house   his
     'It's his house/a house of his' (Sperlich 1997.103)

The second property of the pre-nominal possessor construction is that the pre-nominal possessor must be proper as in (2a), or pronominal as in (3a). It is ungrammatical to have a common pre-nominal possessor, although such a possessor is fine in final position, as shown in (4a,b).

(4) a. Ko   e   pepa he   faiaoga
     Pred   AbsC   book GenC teacher
     'the book of the teacher' (Field Notes.01)
   b. *Ko   e   he   faiaoga   a   pepa
     Pred   AbsC   GenC teacher   *a*   book
     ('the teacher's book') (Field Notes.01)

Given the facts described above, we will address the following two questions: How do we derive the order of elements? How do we account for the two positions (and corresponding properties) of the possessor?

2. Setting the stage

Let us first address the question of the order of elements. One logical possibility is to assume that N is base-generated in the same place it surfaces in (2a), i.e. between the # and the Adjectives. Given the impossibility of N taking Adjectives and Demonstratives as complements, the only way to have this option is to assume a combination of right and left branching, contra Kayne's (1994) antisymmetric system, which disallows left branching universally. Note that this would violate even a weaker version of an antisymmetric system, which would

137

allow cross-linguistic variation in branching direction, but not different directions of branching within a single language or within a single phrasal category. We thus take N in (2a) to be base-generated at the end of the phrase, as shown in (5) (where all other elements appear in their surface order), and we derive its surface position, and those of the other elements, in a manner to be elaborated below.

(5) Case+P/C   #   Adjs   Dem   N

One way of deriving the order in (2a) is to allow N to move over Dem and Adjs to the medial position. We do not adopt this option for two reasons. First, if this is an instance of head-movement, it violates the Head Movement Constraint. If taken to be XP-movement, some constraint on the movement would be required. Otherwise, it would have to be stipulated that this element moves between # and Adjs, and not, for instance, between Case and # or Adjs and Dem, etc. More importantly, linguists who assume a basic universal order and have constraints on movement to account for the order of elements, account for some typological generalizations that would be hard to capture if we allowed this type of movement.

Let us look at some of these typological facts and see how they are accounted for by assuming a universal order of elements. The usual order of elements in the Noun Phrase was perhaps first observed by Greenberg (1966). His observation is given in (6).

(6) Universal 20 (Greenberg 1966:111, see also Hawkins 1983)
"When any or all of the items – demonstrative, numeral, and descriptive adjective – precede the noun, they are always found in that order. If they follow, the order is either the same or its exact opposite."

The generalization in (6) is partially summarized in (7). We leave aside numerals for the moment.

(7) Cinque (1996) and subsequent work, see also Kayne (1994)
   a. Dem – A – N   =Base Ordering
   b. *A – Dem – N   =Impossible
   c. N – Dem – A   =Noun Movement
   d. N – A – Dem   =Successive XP raising   ⟶ Niuean

Cinque (1996) accounts for the ordering restrictions in (7) in the following manner. (7a) is the basic order. The order in (7c) is the result of N-movement (later revised to remnant movement, which we leave aside here). Finally, the mirror-image order in (7d) is the result of successive XP-movement, which we suggest is what happens in Niuean. Crucially, if the XP-movement is successive and local, (7b) is impossible. Our approach is different in technical details from that of Cinque, but the idea is the same. One technical difference, for instance, is that Cinque places the pre-nominal modifiers in Specifier positions, whereas we consider them to be Heads. Our analysis is in line with Rackowski and Travis

(2000) in these respects. See also Shlonsky (2004), who considers that some elements are heads, while others are in specifier position.

Another related fact is the order of descriptive adjectives. It has been suggested that there is a universal order of descriptive adjectives (Laenzlinger 2000, Scott 1998, Sproat & Shih 1991), given in (8).

8. Proposed Universal Order of Adjectives

Quantification > Quality > Size > Shape > Color > Nationality

If in a language like Niuean there is successive XP-movement to derive the mirror-image order, one would expect the adjectives to appear in reverse order as well. This prediction is borne out in all the examples we found in texts. We saw an example of this in (1a) with 'color' and 'quality' in the opposite order. The same phenomenon is observed in (9a) for 'color' and 'size' and in (9b) for 'size' and *oti* 'all', which we take to be a quantificational adjective.

(9) a. e   letio   kula   tote
        AbsC   radio   red   little
        'the little red radio' (Nelisi 1995.6)
     b. e   tau koloa   ikiiki   oti ia   haaku...
        AbsC   Pl   store   small(Pl) all Dem   my
        'all those small stores of mine' (de Sousa 2001.50)

In the next section, we discuss the details of how the inverse order is derived in Niuean.

3. Deriving inverse order

The base order we assume for functional heads is given in (10). This order is based on a body of work on functional categories within the noun phrase. (For examples, see, Ghomeshi and Ritter 1996, Megerdoomian 2002, Pearce 2002, Phan 2001, Ritter 1991, 1995, Schoorlemmer 1998, Travis 1992, among others).

(10) Basic DP Order

K	D	Poss	Dem	A	#	N

The order of K and D in (10) follows standard assumptions in generative theory (see; for K, Bittner and Hale 1996 and for D, Abney 1987). Note that D is null in Niuean.

With respect to the Poss head, it has been suggested in the literature that there are two positions across languages, one lower position much closer to the noun, which is utilized, for instance, by Semitic languages and one higher one, which we suggest is the one used in Niuean. (English possibly uses both positions, e.g. 'John's damaged car door.') Schoorlemmer (1998) discusses these

two possibilities and the properties she attributes to the languages that use the high position coincide with properties of Niuean. We will return to this point below.

For Dem and A, we are following Cinque as in (7). For Number (#), we are following Ritter (1991, 1992).

The order (2a) is derived in a manner illustrated in (11) which involves successive 'intraposition' or roll-up movement of the complements to their empty specifiers. As shown in (11), the #P moves to the spec of AP, then the whole AP moves to the spec of DemP, DemP cannot move to the already filled spec of PossP and finally PossP moves to the spec of DP. DP, with an empty head (see below), does not move, hence K (not represented in (11)), appears at the far left.

(11) Pre-nominal possessor derivation (2a)

$$[_{DP} \underline{\quad\quad} D \ [_{PossP} \underline{DP_{Poss}} \ Poss \ [_{DemP} \underline{\quad\quad} Dem \ [_{AP} \underline{\quad\quad} A \ [_{\#P} \# \ [_{NP} N]]$$
$$a$$
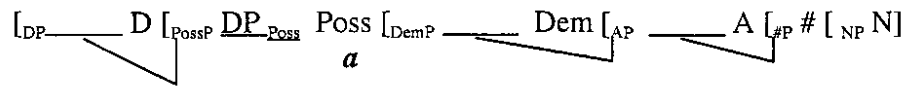
The question arises as to why NP does not move to the specifier of #P. This can be answered easily if # is in specifier position of #P, in which case it would block the movement of NP. This claim is supported by the fact *falu* usually appears with an optional particle *a*, which is arguably the head of #P, and by the fact that *taha* has an alternative form *ha*, suggesting that this number marker is formed of two parts, with *ta* in specifier and *ha* in head position.

The pre-nominal order of the possessor has two properties given below.

Property 1: Pre-nominal possessor gives a definiteness reading to the DP as a whole (3a).

According to Schoorlemmer (1998), in languages with the high PossP, Poss is a potential carrier of a value for definiteness. We posit that *a* in Niuean is one such element. The Poss head, which is home to *a*, an element with semantic content, gives the whole DP the definite reading.

Property 2: The pre-nominal possessor must be proper or pronominal (2a, 3a).

We posit that the Poss morpheme *a* has a [proper] feature which must be shared with its specifier. This is supported by the fact that *a* has three other uses in Niuean that bear the feature proper (Absolutive proper case, proper article in goal DPs and Genitive proper case). Thus, *a* has two roles, giving the definite reading to the whole DP and the [proper] feature to the possessor.

Let us now turn to the order (2b). This is shown in (12).

140

(12) Post-nominal possessor derivation (2b)

$$[_{DP}\_\_\_ D\ [\ _{PossP}\ \underline{DP}_{\underline{Poss}}\ Poss\ [_{DemP}\_\_\_\_\ Dem\ [_{AP}\_\_\_\_\ A[_{\#P}\ \#\ [\ _{NP}\ N]$$
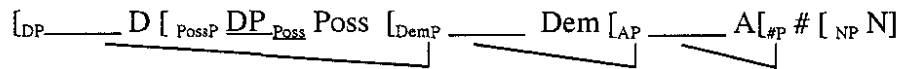
The first two movements are exactly the same as (11), #P to spec of AP and AP to spec of DemP. The only difference here is that there is no *a* in Poss. To get the right order, we need the whole DemP to move over PossP to the spec of DP, as shown in (12). The question remains, however, as to why in (12) DemP, rather than PossP, moves to spec of DP. Recall that in (11), it was PossP that moved, which is expected under some version of relativized minimality or shortest move. To explain the phenomenon in (12), we make use of an idea in Rackowski and Travis (2000), where they derive the order of adverbs in Malagasy and Niuean from Cinque's universal order of adverbs. Let us look at their analysis briefly.

Their derivation for the Niuean verb phrase is given in (13).

(13) Niuean clausal derivation (adapted from Rackowski & Travis 2000)

$$[_{QP}\_Q[_{AgrSP}\_AgrS[_{AgrOP}\_AgrO[_{Asp\text{-}AdvP}\_\_Asp\text{-}Adv\ [_{aiP}\_ai[_{ManP}\_\ Man[_{DirP}\_Dir\ [_{VP}V...$$
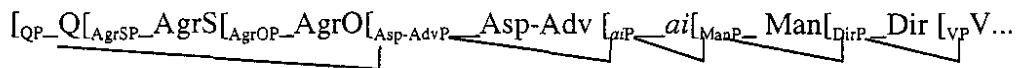
Note the striking parallel between (13) and (12), which essentially involve the same series of movements. Thus, in (13), VP moves to spec of DirP, DirP to spec of ManP, and so forth. Crucially, when the movement sequence gets to the AgrO and AgrS phrases, they are skipped and they cannot themselves move. To account for this fact, Rackowski and Travis suggest a restriction on movement given in (14).

(14) Rackowski & Travis (2000:127)
> "To avoid this ungrammatical derivation, there must be a restriction in the grammar such that non-contentful phrases like AgrP are invisible to movement and cannot themselves move. In contrast to this, contentful phrases like AdvPs can and, in this case must, move."

The restriction is that non-contentful phrases like AgrP are invisible to movement and cannot themselves move. We suggest that the same restriction is in place for PossP in (13). Note the plausibility of this suggestion, given the parallel between AgrP and PossP. In fact, we seem to have come across a striking example to support their proposal. Here, we have a head, which is contentful in one case and non-contentful in the other. When it is non-contentful as in (12), it is skipped and cannot itself move. In (11), on the other hand, the Poss head is contentful; it contains the feature definite realized by *a*. In this case, as predicted by Rackowski and Travis, the PossP moves which results in the pre-nominal possessor order. (See den Dikken (to appear) for a different approach to word order and to linking items.)

4. Numerals

As well as variation in word order of possessors, Niuean exhibits variation in word order of numerals. Given what we have proposed for Niuean, we expect an inverse order for numerals compared with adjectives and demonstratives: [Adjective - Numeral – Demonstrative], as outlined in (6) from Greenberg (1966) and also discussed in Cinque (1996). Let us consider how numerals actually do behave in Niuean.

In fact, in Niuean, numerals are found both preceding and following the noun, as shown in (15). When numerals precede the noun, a ligature item *e* appears between the numeral and the noun, similarly to pre-nominal possessors. It is preferred that the number marker be null in such cases. In case of post-nominal numerals, this ligature item does not appear. (Note when counting humans, the prefix *toko* appears on the numeral, which is sometimes written as part of the numeral, or with a hyphen, and sometimes written as a separate word as in the examples.)

(15) a. toko-lima   e   tagata loloa
     Pers-five   Lig person tall
     'five tall people' (de Sousa 2001.33)
    b. Maori toko ua
     Maori Pers three
     'three Maoris' (Blanc and Togakilo 1965)

Pre-nominal numerals can co-occur with possessors. The most commonly found construction with both a numeral and a possessor is one where the numeral precedes the noun and the possessor follows it, as in (16).

(16) Ko e   toko fa: e tama a   Matakuhifi
     Pred AbsC Pers four Lig boy   Gen   Matakuhifi
     'Matakuhifi's four sons' (Blanc and Togakilo 1965)

Let us consider first the pre-nominal numeral order in (15a) and (16). If we take the base order to be that assumed by Cinque (1996), namely [Dem Numeral Adj N], the order in (15a) and (16) can be derived in a straightforward manner. (The examples here, however, do not include a demonstrative or # marker.)

(17) [$_{DP}$ D [$_{PossP}$ DP$_{Poss}$ Poss [$_{DemP}$ Dem [$_{NumP}$ Num *e* [$_{AP}$ A [$_{\#P}$ # [$_{NP}$ N]

First, the #P moves to the specifier of AP, just as in all the derivations so far. At this point, it is not possible to move the AP into the specifier of NumP, because this position is filled by the Numeral, while the head of NumP is filled by the particle *e*. Instead, the NumP is moved to the specifier of DemP. This

movement exactly parallels the movement in (11) of PossP to DP, in place of movement of DemP to PossP, the latter of which is similarly blocked by a full specifier position. Then, the derivation proceeds as does that in (12), to yield the order in (15a) and (16) where the numeral is pre-nominal. This derivation yields a post-nominal possessor and a pre-nominal numeral as in (16).

In Niuean, each of the possessor and the numeral can appear pre-nominally (2a & 15a). It is unclear whether both the possessor and the numeral can be pre-nominal in the same phrase, because there is a strong tendency on the part of speakers to avoid having more than one such element on one or other side of the noun. We leave this question open, pending future fieldwork.

Let us now address the post-nominal numerals, as in (15b). Note that there is no ligature item *e* in these cases. Our analysis of nominal movements laid out in this paper might predict that in such a situation, NumP will resist movement, as PossP did in (12) to yield a word order [Poss # N A Dem Num], which is not what is found, since numerals do not appear after demonstratives in Niuean.

(18) An Incorrect Prediction

[$_{DP}$ D [$_{PossP}$ DP $_{Poss}$ Poss [$_{DemP}$ Dem [$_{NumP}$ Num *null* [$_{AP}$ _ A [$_{\#P}$ # [$_{NP}$ N]
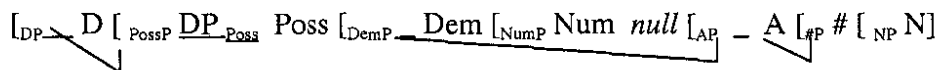
A solution lies in the analysis of numerals as heads in such cases (see Shlonsky 2003). If this position is tenable, the correct derivation results.

(19) [$_{DP}$ D [$_{PossP}$ DP $_{Poss}$ Poss [$_{DemP}$ _ Dem [$_{NumP}$ Num [$_{AP}$ _ A [$_{nP}$ # [$_{NP}$ N]
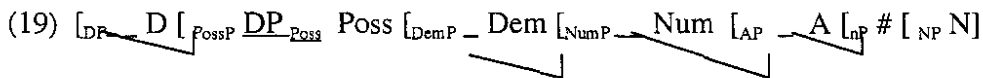
This derivation yields the word order [Poss # N A Num Dem]. The important question is, does Niuean conform to the Greenberg generalization that when post-nominal, numerals follow adjectives? It is difficult to determine this, because, as just discussed, there is a strong preference to avoid having strings of multiple 'modifiers' (in the loose sense, including numerals and possessors). Thus, when a noun is modified by both a possessor or adjective and a numeral, one or other of them is usually found pre-nominally, while the remaining one is found post-nominally. This is true of all natural examples we have found in texts, as (16). When the data was elicited, in fact, varying orders were accepted when presented, as shown in (20).

(20) a. e   tau manu ua   kula   fulufuluola   e:
          AbsC Pl bird   three   red   beautiful   Dem
          'those three beautiful red birds' (Field Notes 2001)
      b. e   tau manu kula   ua   fulufuluola e:
          AbsC Pl bird   red   three beautiful Dem
          'those three beautiful red birds' (Field Notes 2001)

The status of such orders remains unclear, however, given the lack of such examples in Niuean texts. We note here simply that it is possible to derive these varying orders by choices regarding piedpiping, but since the status of these examples seems marginal, we do not discuss them further here.

In addition to the uses above, Niuean numerals can also function as predicates as in (21). In this function, they often modify noun phrases within a relative clause, as in (22). This serves as another strategy to avoid pile-up of modifiers in the noun phrase, and is common when there are multiple modifiers, as in (22), and (23) from de Sousa (2002).

(21)   Valu   e   hui   he   feke
          eight AbsC leg   Gen   octopus
          'An octopus has eight legs.' (Sperlich 1997)

(22)   motu ikiiki mo e tokolalo   ne   fa꞉
          island small and   sandy   Comp four
          'four small and sandy islets' ((lit. 'small and sandy islets that are four')
          (Blanc and Togakilo 1965)

(23)   fiamanako   au ke   sela e   tau koloa ikiki oti ia   haaku
          want   I   Comp sell AbsC Pl   store   small all Dem my
          ne   lima   i Niue   ne   mai   e   Sione
          Comp five   Loc Niue Comp give Erg:P Sione
          'I want to sell all those five small shops of mine in Niue which I have
          inherited from Sione.' (de Sousa 2001.50)

Examples such as (22) and (23), with a relative clause, and examples with a PP argument, as in (23), along with deverbal nominal clauses raise further questions about word order in Niuean nominal phrases, however these questions remain for future research.

5. Conclusion

Following Cinque (2000), we have presented a roll-up analysis of Niuean DPs that derives the correct word orders and accounts for the position and properties of possessors, and for the variation in pre-nominal and post-nominal numerals. Having a filled specifier can affect the pattern of movement (Shlonsky 2004), and in addition, we have claimed, following Rackowski and Travis (2000), that the stoppage of movement that is seen in some cases in Niuean is tied to the content of functional heads. If the head is filled, movement of the phrase is possible, but if the head is phonologically and semantically empty, the phrase does not move. In the case of numerals we posited that when the ligature morpheme is absent, the numeral itself is in the head position.

References

Abney, Steven. 1987. *The English Noun Phrase in its Sentential Aspect*. Doctoral dissertation, MIT.

Belleti, Adriana. (ed.) 2004. *Structures and Beyond – The Cartography of Syntactic Structures*. Vol 3. Oxford: Oxford University Press.

Bittner, Maria and Ken Hale. 1996. The Structural Determination of Case and Agreement. *Linguistic Inquiry, 27*, 1-68.

Blanc, R. R. V. and Togakilo (translator). 1965. *Ne Toka Hifo e Kuki e Higoa Haana he Tau Aelani*. Islands Education Division, Dept. of Education, Wellington, New Zealand.

Borer, Hagit. 1999. Deconstructing the Construct. in Kyle Johnson and Ian Roberts (eds.) *Beyond Principles and Parameters*. Kluwer Publishing:.Dordrecht. pp.43-89.

Cinque, Guglielmo. 1996. The Antisymmetric Programme: Theoretical and Typological Implications. *Journal of Linguistics, 32*, 447-464.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-linguistic Perspective*. Oxford: Oxford University Press.

Cinque, Guglielmo. 2000. On Greenberg's Universal 20 and the Semitic DP. *The University of Venice Working Papers in Linguistics*, 10.2, 45-61.

Cinque, Guglielmo, (ed.) 2002. *The functional structure of DP and IP – The Cartography of Syntactic Structures*. Vol.1. Oxford: Oxford University Press.

den Dikken, Marcel. 2003. *The Structure of the Noun Phrase in Rotuman*. LINCOM Studies in Austronesian Linguistics 05, München, Germany.

den Dikken Marcel. To appear. *Relators and Linkers: The Syntax of Predication, Predicate Inversion and Copulas*. Cambridge: MIT Press.

Fassi Fehri, Abdelkader and Marie-Thérèse Vinet. To appear. Distribution of Number and Classifier in Arabic and Chinese and Parameterization. *Linguistic Research 9.1*. IERA Publications. Rabat.

Ghomeshi, Jila, and Elizabeth Ritter. 1992. Binding, Possessives, and the Structure of DP. *Proceedings of NELS 26*, K. Kusumoto (ed), GLSA, University of Massachusetts.

Greenberg, Joseph. 1966. *Language Universals: With Special Reference to Feature Hierarchies*. The Hague: Mouton.

Hawkins, John A. 1983. *Word Order Universals*. New York: Academic Press.

Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge: MIT Press.

Laenzlinger, Christopher. 2000. French Adjective Ordering: Perspectives on DP-internal Movement Types. *Generative Grammar in Geneva* 1, 55-104.

Massam, Diane and Wolfgang Sperlich. 2000. Possession in Niuean. in Steven Fischer (ed.) *Possessive Markers in Central Pacific Languages*. Thematic volume of *Language Typology and Universals*, 53.3/4, 281-292.

Megerdoomian, Karine. 2002. *Beyond Words and Phrases: A Unified Theory of Predicate Composition*. Doctoral dissertation. University of Southern California.

Nelisi, Lino. 1995. *Ko e Letio Kula*. Wellington: Ministry of Education.

Pearce, Elizabeth. 2002. DP Structure and DP Movement in Maori. Paper presented at COOL5, University of Canberra, Australia.

Pearce, Elizabeth. 2003. Phrasal movement within the Maori DP. *Digests of Selected Papers Presented at AFLA X*. University of Hawai'i at Manoa Working Papers in Linguistics 34.2. Pp. 41-42.

Phan, Kristin. 2001. Not so Bare after all: Observations on the Cantonese Nominal Phrase. ms. University of Toronto.

Rackowski, Andrea and Lisa Travis. 2000. V-initial Languages: X or XP Movement and Adverb Placement. in Andrew Carnie and Eithne Guilfoyle (eds.) *The Syntax of Verb Initial Languages*. Oxford: Oxford University Press.

Ritter, Elizabeth. 1988. A Head Movement Approach to Construct-State Noun Phrases. *Linguistics*, 26, 909-929.

Ritter, Elizabeth. 1991. Two Functional Categories in Noun Phrases: Evidence from Modern Hebrew. in Susan. D. Rothstein (ed.) *Perspectives on Phrase Structure: Heads and Licensing*. Syntax and Semantics 25. New York: Academic Press. 37-62.

Ritter, Elizabeth. 1993. Where's Gender? *Linguistic Inquiry*: 24.4: 795-803.

Rizzi, Luigi. 1997. The Fine Structure of the Left Periphery. in Liliane Haegeman (ed.) *Elements of Grammar*. Dordrecht: Kluwer Publishing.

Rizzi, Luigi. (ed.) 2003. *The Structure of CP and IP – The Cartography of Syntactic Structures*. Vol. 2. Oxford: Oxford University Press.

Schoorlemmer, Maaike. 1998. Possessors, Articles and Definiteness. in Artemis Alexiadou and Chris Wilder (eds.) *Possessors, Predicates and Movement in the Determiner Phrase*. Amsterdam: Benjamins.

Scott, G.-J. 1998. Stacked Adjectival Modification and the Structure of Nominal Phrases. *SOAS Working Papers in Linguistics and Phonetics*, Vol. 8, 59-89.

Seiter, William. 1980. *Studies in Niuean Syntax*. New York: Garland Press.

Shlonsky, Ur. 1988. Government and Binding in Hebrew Nominals. *Linguistics*, 26, 951-976.

Shlonsky, Ur. 2004. The Form of Semitic Noun Phrases. *Lingua* 114.12: 1465-1526.

de Sousa, Hilario. 2001. Noun Phrase Structure and the Case Marking System in Niuean. ms. University of Auckland.

Sperlich, Wolfgang. 1997. *Tohi Vagahau Niue/Niue Language Dictionary*. Honolulu and Alofi: University of Hawai'i Press and the Government of Niue.

Sproat, Richard & Chilin Shih. 1991. The Cross-linguistic Distribution of Adjective Ordering Restrictions. in Carol Georgopolous & R. Ishihara (eds.) *Interdisciplinary Approaches to Language. Essays in Honor of S. Y. Kuroda.* Dordrecht: Kluwer Publishing.

Travis, Lisa. 1992. Inner Tense with NP: the Position of Number. *CLA Proceedings.* University of Toronto Working Papers in Linguistics. pp. 329-346.

Arsalan Kahnemuyipour
Syracuse University
Department of Languages, Literatures, and Linguistics
340 H. B. Crouse Hall
Syracuse
New York 13244 1160
akahnemu@syr.edu

Diane Massam
University of Toronto
Department of Linguistics
130 St. George Street
Toronto, Ontario
M5S 3H1
diane.massam@utoronto.ca

Rigidity versus Relativity in Adverbial Syntax

Daniel Kaufman
Cornell University

Two diametrically opposed stances have emerged from recent theoretical debates on adverbial syntax. One approach, represented by Alexiadou (1997) and Cinque (1999), espouses a rigid hierarchy of functional projections hosting individual adverbs. The other, represented broadly by Jackendoff (1972), McConnell-Ginet (1982) and most recently Ernst (2002), takes adverb placement to be determined by the semantics of the adverbs themselves as opposed to the functional architecture of the clause. Under the latter view, adverbs may be divided into several categories based on their meaning with each category being licensed in a certain range within the sentence.

Here, I undertake a detailed examination of Tagalog adverbs and compare the predictions of the two best articulated recent theories of adverbs, that of Cinque (1999, 2004) and Ernst (2002). The results offer support for some of the basic predictions of the semantically based approach of Ernst. Particularly important are scopal facts which do not obtain a clear explanation under a functional projection-based theory such as Cinque's.

1.0 Two theories of adverbs compared

Because proponents of both theories considered here employ certain ad-hoc mechanisms for handling exceptions – and are thus able to achieve similar empirical coverage – *naturalness* must figure prominently as an evaluation metric.[1] Since each theory is suited to naturally handle a well circumscribed set of phenomena the theories under consideration must be evaluated in terms of how well the *core predictions* are corroborated by the data. For the purposes of the paper I will refer to the Alexiadou-Cinque theory as the *Rigid Approach* and Ernst's theory as the *Relative Approach*.[2]

1.1 Fundamentals and predictions of the Rigid Approach

Cinque (1997) bases his proposal on the notion that adverb order is much stricter than can be predicted by scopal properties and semantic domains. He offers evidence, primarily from Romance languages, to show that ordering relations between adverbs are fixed and transitive. That is to say, if the linear relations between adverbs A, B and C are such that A>B and B>C then it is possible to conclude that A>C.

Two central tenets of Cinque's proposal are: (i) Adverbs are not adjuncts but rather specifiers within an array of strictly and universally ordered functional projections; (ii) Each adverb corresponds to an inflectional head which carries an analogous meaning. Thus, the same order is reflected by verb morphology.

Strong evidence for the specifier hypothesis is the relation between single adverbs and the verb in several Romance languages. In these languages, the verb may appear on differing sides of an adverb based on whether the verb is finite, infinitival, participial, etc. This suggests that the verb crosses certain adverbs when raising to an inflectional head. This phenomenon clearly requires a syntactic explanation (in the narrow sense) since scopally, a verb cannot be

148

said to interact with a single adverb in the same way multiple adverbs interact with each other. In any case, this phenomenon is largely irrelevant in dealing with Austronesian languages since verb movement is not "gradated" as in Romance languages, i.e., we do not find overt evidence that the verb is in different positions based on finiteness or other inflectional features.

In support of a Spec-Head relationship between adverbs and verbal morphology, Cinque offers data from a wide range of languages to show that the order posited for adverbs is reflected (in reverse) in the domain of the word. Thus, an aspectual adverb, for instance, would be found in the specifier of a functional projection such as PerfectiveP while its corresponding verbal morpheme would constitute the head of this phrase.

Unexpected surface orders (i.e. those not corresponding to the underlying hierarchy of FPs) result from one of the following (Cinque 1999:3-4): (i) "When an AdvP directly modifies (is the specifier of) another AdvP."; (ii) "When a lower portion of the clause (containing an AdvP) is raised across a higher AdvP (for focus-presupposition requirements)."; (iii) "When one AdvP is wh-moved across another."; (iv) "When one and the same AdvP can be 'base generated' in two different positions in the clause (with one of the two positions to the left, and the other to the right of another AdvP)."; (v) "When a noninherently 'focusing' AdvP (e.g. probably) is used as a 'focusing' adverb (like only and simply)."; (vi) "When an adverb is used 'parenthetically'."

As Cinque notes, (iii) and (vi) are uncontroversial because they are apparent from surface form and tied to more general phenomenon. The other exception-creating circumstances, however, are more contentious in that they are less detectable and more specific to adverbs. Because of space restrictions, we focus our scrutiny on (iv) as this allowance risks proliferating functional projections to capture different interpretations of single adverbs.

The core predictions of the *Rigid Approach* may be summarized as the following: (i) Ordering of adverbs should be highly restricted both cross-linguistically and within single languages; (ii) Deviant orders should show the hallmarks of XP movement to a higher Spec of a functional projection (esp. FocP, TopP); (iii) Scope may often not correspond to linear order since there need not be a direct link between the scope of an adverb and its position in relation to other elements.

Cinque (2004) cites two arguments from Rizzi (2002) in support of prediction (ii). Movement of adverbs to the pragmatically unmarked ModifierP (Rizzi 2002) appears to be subject to Relativized Minimality effects (1-2).[3]

(1)   Rapidamente, qualcuno farà sparire i documenti
        Quickly, someone will make the documents vanish

(2)   *Rapidamente, qualcuno probabilmente farà sparire i documenti
        Quickly, someone will probably make the documents vanish

In addition, this displacement appears to be clause-bounded (3) (adapted from Cinque 2004:703 ex. 39-41), unlike focalization and topicalization (4-5).[4]

(3)   *Rapidamente, credo che qualcuno farà sparire i documenti
        Quickly, I think that someone will make the documents vanish

149

(4)   Rapidamente, credo che nessuno farà sparire i documenti
       Quickly (topic), I think that nobody will make the documents vanish

(5)   RAPIDAMENTE, credo che qualcuno farà sparire i documenti
       Quickly (focus), I think that someone will make the documents vanish

In regard to prediction (iii), Cinque is not precise about how the relative scope of adverbs is to be derived. This is a major point of divergence between the two theories since, for the *Relative Approach*, scope is the primary determinant in the positioning of adverbs of the same class. For Cinque, adverbs that are generally understood to have variable scope in relation to other scope sensitive material (e.g. frequency adverbs) are generated in unique positions in the hierarchy. Support for the *Rigid Approach*, therefore, may come in the form of adverbs that may obtain several scopal interpretations from one "canonical" surface position.

1.2 Fundamentals and predictions of the Relative Approach
Three principles within Ernst's system which contrast with the Rigid approach are the following: (i) Adverbs are adjoined freely, constrained only by general restrictions on adjunction; (ii) Sentences are composed incrementally from several semantic layers and there exists a *minimal* correspondence between semantic layers and syntactic phrases in the clausal architecture; (iii) An adverb is uninterpretable if it cannot access the semantic layer it requires from its position in the sentence. This happens, for instance, when a lower level adverb (e.g. Manner) is composed after a higher level adverb (e.g. Speech Act).
       Ernst (2002) develops a theory of adverbs in the former tradition of treating them as adjunctions. Applying a more articulated theory of propositional semantics (based largely on Parsons 1990), Ernst is able to restrict the positions of adverbs according to what type of semantic layer, or Fact-Event-Object (FEO), they may modify. FEOs are thus treated as the semantic arguments of adverbs. Based on their semantic class, adverbs are specified to take particular FEO types and are capable of shifting that type as a result of composition. Composition is a step-wise process and layers are added under syntactic sisterhood. The basic FEO types/layers employed are the following:

(6)   Speech-Act > Fact > Proposition > Event > Specified Event

The constraints on how these types combine, together with ostensibly universal syntactic principles, are understood to be responsible for the distribution of adverbs within the clause. A bare verb is associated only with the lowest FEO type (the Event) which must eventually be built into a Speech Act by the time the sentence is complete. This, however, is not a strictly unidirectional process which corresponds directly to phrase structure. Rather, FEO types can raise independently on either side of the predicate and thus adverbs should show similar linear and scopal effects on either side of the verb. One important mitigating factor is the special connection between clausal Predicational adverbs and the syntax. According to Ernst, clausal predicational adverbs are subject to the same syntactic restrictions as predicational heads and thus follow the similar ordering principles. This is posited to account for the fact that these

150

adverbs regularly precede the verb cross-linguistically. Non-predicational adverbs are not bound by such syntactic restrictions and therefore display a wider range of potential positions.

Adverb movement falls under the scope of more generally accepted cases of movement such as topicalization, *wh*-movement, clefting, etc. Under this view, then, there is no independent motivation for adverb movement and thus it should not be considered a separate "type" of movement on par with those mentioned above. Nor should there exist special phrases to host adverbs such as Rizzi's (2002) ModifierP.[5] For Ernst, the only strict correspondence between FEO types and specific syntactic projections is that only event-internal modification is possible within the "L-syntax" (i.e., corresponding to VP *pace* Hale & Keyser 1993). This entails that adverbs adjoined within VP can only receive limited interpretations (e.g. manner, measure).

Ernst proposes a semantically motivated taxonomy of adverbs which have their own FEO selection requirements and which may yield different FEO types upon composition. For example, epistemic adverbs such as *probably* may be represented as $[_{FACT}ADV[_{PROP}]]$ indicating that they take a Proposition as their complement and yield a Fact (cf. Parsons 1990 for discussion of these terms).

The building up of propositions and the composition of adverbs follows the FEO calculus. The primary principle of this is that any FEO type may be freely converted to any higher FEO type; but not lowered (Ernst 2004:761, 2002:50). In principle, the FEO type of a sentence fragment may be lowered *through composition* but this is very rare (and may also probably be subject to a non-lowering analysis). In general, adverbs and operators either raise the FEO type or maintain it. Because the FEO type may be freely raised without overt operators, an adverb that selects for a higher FEO type such as Fact or Speech-Act need not occur at a syntactically designated peripheral position but can rather be accommodated in more internal positions through FEO type raising (modulo VP-internal positions, the domain of event-internal modification).

The manner/clausal ambiguity with adverbs such as *clearly* as in (7) results from the fact that manner adverbials are composed via the Manner Rule. This is a rule which obligatorily applies to predicational adverbs within L-syntax and which may optionally apply to constituents of the Event type (although a clausal reading is preferred outside of VP).

(7)   John clearly saw everything written on the chart.
        (He saw it clearly ~ Clearly, he saw it.)

The relevant predictions of the *Relative Approach* are summarized as the following: (i) An adverb that selects for a lower FEO type should not appear higher than one which selects for a higher FEO type; (ii) The positions in which an adverb is licensed should constitute a contiguous range w.r.t. non-FEO-type changing material; (iii) Adverb scope should be reflected by surface order in a *concentric* fashion with the predicate as the center due to the symmetric manner in which semantic layers are built in the FEO calculus.

2.0 Tagalog adverbs

Tagalog offers a good testing ground for the predictions of the two theories as Tagalog adverbs are introduced into the clause by means of overt markers that elucidate their syntactic position.

The lowest attachment of adverbs is marked with *nang*, the same particle employed to introduce internal arguments. Manner adverbials are typically introduced in this way as can be seen in (8). Here we see that the position of the manner adverb is free within the post-verbal domain and that the various positions are interpreted uniformly.[6]

(8) P-in-ag-aral-an (nang maigi) ng piloto (✓) ang mapa (✓)
     PRF-TR-study-LV *nang* well GEN pilot   SUB map
     'The pilot studied the map thoroughly.'

(9) P-um-asok (nang madalas) sa opisina (✓) si Ben (✓)
     AV.PRF-enter *nang* often OBL office   P.SUB B.
     'Ben went to the office often.'

Adverbs may also be introduced in the same way as topics, that is, in clause-initial position preceded by the marker *ay*. Compare the topicalization construction in (10) with the adverbials in (11) and (12).

(10) Ang tubero ay nag-trabaho sa kusina.
     SUB plumber TOP AV.PRF-work OBL kitchen
     'The plumber, worked in the kitchen.'

(11) Kadalasan ay hindi siya p-um-apasok sa klase
     usually   TOP NEG 3sg.SUB <AF>IMPF.enter OBL class
     'Usually, he doesn't come to class.'

(12) Malamang ay nan-daya   sila
     probably TOP AF.PRF-cheat 3p.SUB
     'They probably cheated.'

Whereas we see from the above that adverbs may be marked as arguments, it is also possible to introduce adverbs in the same way as adjectives, that is, directly adjacent to the modified head with the linker *na/-ŋ*. Compare (13) and (14).

(13) Madalas na   pag-ulan.
     frequent LNK NOMZ-rain
     'Frequent rain.'

(14) Madalas na umuulan.
     frequent LNK AV.IMPF-rain
     'It rains frequently.'

For this type of modification, adjacency is required between the adverbial and the verb as shown by the ungrammatical (15).[7]

(15) *P-um-apasok si   Juang madalas
      AV.IMPF-enter P.SUB J.-LNK often
      'Juan enters often.'

Adverbial notions are also often introduced syntactically as clefted adjectives. Strictly speaking, these should not be considered as adverbials at all since syntactically and semantically such constructions are indistinguishable from adjectival predicate constructions. However, the fact that this is the only way to express certain adverbial notions in Tagalog is instructive and thus merits inclusion here.

(16) Malinaw na mag-sasalita   mamaya sa miting si   Bobong.
      clear   COMP AV-IRR.speak later   OBL meeting P.SUB Bobong
      'Clearly, Bobong will speak later at the meeting.'
      (or, 'It's clear that Bobong will speak later.')

Other adverbs may be introduced into the clause without any marker at all. The adverbs which allow this are typically mono-morphemic (not containing the *ma-* adjectival prefix) and may also occasionally be treated as clitics, being positioned within the second position clitic cluster.

(17) B-in-atíkos   (uli) siya (✓) ng mga guro   (✓).
      PV.PRF-criticize again 3s.SUB   GEN PL teacher
      'He was criticized again by the teachers.'

(18) K-um-agat   sa kanya (kahapon) ang aso (✓)
      AV.PRF-bite OBL 3s.OBL yesterday SUB dog
      'The dog bit him yesterday.'

Finally, there is the class of clitic adverbs that is restricted to appearing in second position (19-21). These consist of aspectual, mood, evidential and functional adverbs. A small number of adverbs that fall in this class may also appear in topic position as the first occurrence of *sana* 'OPTM' in (21) does.

(19) B-um-alik   na   nga po pala sila!
      AV.PRF-return already EMPH POL SURP 3p.SUB
      'They really already returned!'

(20) Mag-tatago pa rin   ba kaya kayo?
      AV-IRR.hide still also Q RHET 2p.SUB
      'I wonder, will you still hide?'

(21) (Sana ay) g-um-radweyt (sana) siya   ngayon.
      OPTM TOP AV.INF-graduate OPTM 3s.SUB now
      'Hopefully, he'll graduate.'

We are now in a position to look at the relation between adverb types and positions in Tagalog. Sentences (22-25) show the positional possibilities of an ambiguous clausal/manner adverb.

153

Malinaw 'clear(ly)': (i) Manner reading – [SPECEVENT ADV [SPECEVENT]]
(ii) Evidential reading – [STATE ADV [FACT]]

(22)   [TP K-um-akantaᵢ kaⱼ [VP tᵢ tⱼ nang malinaw ngayon]]
AV.IMPF-sing 2s.SUB   *nang* clear   now
'You're singing clearly now.' (≠ 'Clearly, you're singing now.')

(23)   [TP Malinaw kangⱼ   [TP k-um-akantaᵢ [VP tᵢ tⱼ ngayon]]]
clear   2s.SUB-LNK   AV.IMPF-sing   now
'You're singing clearly now.' (≠ 'Clearly, you're singing now.')

(24)   [TopP Malinaw ay [TP k-um-akantaᵢ kaⱼ [VP tᵢ tⱼ ngayon]]]
clear   TOP   AV.IMPF-sing 2s.SUB   now
'Clearly, you're singing now.' (≠ 'You're singing clearly now.')

(25)   [TP Malinaw [CP na [TP k-um-akantaᵢ kaⱼ [VP tᵢ tⱼ ngayon]]]
clear   COMP   AV.IMPF-sing 2s.SUB   now
'Clearly, you're singing now.' (≠ 'You're singing clearly now.')

The adverb in (22) represents the *nang* (internal argument-like) attachment showing that only a manner reading is possible from this position; (23) shows the adverb adjoined like an adjective in the preverbal position; (24) exemplifies the adverb in topic position; (25) an adverbial meaning is obtained through an adjectival predicate with a clausal complement. Note that there is a difference in the readings between (22-23) on the one hand and (24-25) on the other hand. In the former group the adverb can only obtain a manner interpretation while in the latter group the adverb can only obtain a clausal interpretation. Generally, the argument-like *nang* attachment of adverbs necessitates a manner reading. The positional possibilities for a modal adverb are shown in (26-29).

Malamang 'probably': modal adverb – [FACT ADV [PROP]]

(26)   *[TP K-um-akantaᵢ siyaⱼ [VP tᵢ tⱼ nang malamang ngayon]]
AV.IMPF-sing 3s.SUB   *nang* probably   now

(27)   *[FocP Malamang siyangⱼ   [TP k-um-akantaᵢ [VP tᵢ tⱼ ngayon]]
probably 3s.SUB-LNK AV.IMPF-sing   now

(28)   [TopP Malamang ay [TP k-um-akantaᵢ siyaⱼ [VP tᵢ tⱼ ngayon]]
probably TOP   AV.IMPF-sing 3s.SUB   now
'Probably, he's singing now.'

(29)   [TP Malamang [CP na [TP k-um-akantaᵢ siyaⱼ [VP tᵢ tⱼ ngayon]]
probably COMP   AV.IMPF-sing 3s.SUB   now
'Probably, he's singing now.'

A summary of the possible positions (excluding predicate position) for several types of adverbs is given in Table 1 (m = manner interpretation; c = clausal, *epi* = epistemic, *deo* = deontic).

Table 1. Summary of adverb type-position relations

	[TopP ADV *ay*]	[TP ADV (clitic)	[VP *nang* ADV]]]
a. Pure manner i. *mahigpit* 'tight(ly)'	*	✓ *m*	✓ *m*
b. Manner/clausal i. *malinaw* 'clear(ly)'	✓ *c*	✓ *m*	✓ *m*
c. Exocomparative i. *ganito* 'like this'	✓ *c*	✓ *m*	✓ *m*
d. Subject oriented i. *kusang-loob* 'willingly'	✓	✓	✓
e. Functional–quant. i. *madalas* 'often' ii. *kadalasan* 'most often'	✓ ✓	✓ *	✓ *
f. Modal i. *dapat* 'should' ii. *malamang* 'probably'	✓ *epi.* ✓	✓ *deo.* *	* *
g. Speech act i. *sa madaling salita* 'briefly'	✓	*	*

What emerges clearly from Table 1 is that the clause appears to be bifurcated at the edge of TP. Differences in interpretation and grammaticality are found to relate to the boundary marked by the topic position. As seen in (a-c), manner readings are only found within TP. Modals present a divergence from the pattern as they are the only category that may appear at the edge of TP but not within VP. The TP boundary is still relevant however since the deontic reading is highly favored within TP while the epistemic reading is favored outside of TP. Although it is difficult to make a direct connection between event-internal modification and deontic modality, the restricted readings for modals should be tied to the fact that the lower range does not allow access to higher FEO types. Finally, a number of adverbs cannot appear within TP at all.

The facts in Table 1 support two of Ernst's claims: that manner readings are restricted to (but free within) a lower domain and that clausal predicative adverbs are linearized in relation to their FEO complements in accordance with the general head-complement pattern of head-initial languages. This latter claim is supported by the fact that clausal adverbs such as (f.ii) and (g) in Table 1 can only appear as topics, preceding the entire TP complement.

Two deviations from Ernst's theory should be noted here. The edge of event-internal modification appears to be higher than VP and subject oriented adverbs such as 'willingly' are not expected to surface in the domain of event-internal modification. I will not offer an explanation for these facts here as they do not affect the more far-reaching debate between the two theories.

4.0 Evidence for Relativity

While the facts shown in the previous section suggest a good deal of freedom, they do not provide an iron-clad argument for relativity in and of themselves since there always exists the possibility of movement. In this section I will present evidence which bears directly on the issue of relativity.

4.1 Syntactic locality versus the FEO calculus

As seen above, a frequency adverb such as *madalas* 'often' may be introduced in several different positions in the clause as seen in (30-32).

(30) Um-uulan   nang madalas dito
AV .IMPF-rain *nang* often   here
'It rains often here.'

(31) Madalas ay   um-uulan   dito
often   TOP AV-IMPF.rain here
'It rains often here.'

(32) Madalas um-uulan   dito
often   AV-IMP.rain here
'It rains often here.'

Crucially however, these positions are limited with the presence of another adverb. The modal clitic adverbs *sana* 'hopefully' and *yata* 'perhaps' appear in the second position of the clause as in (33) and (34).

(33) B-um-isita sana   si Juan.
AF.PRF-visit OPTM P.SUB Juan
'Hopefully, Juan visited'

(34) B-um-isita yata si   Juan.
AF.PRF-visit EPST P.SUB Juan
'Perhaps Juan visited.'

Whereas a frequency adverb such as 'often' may be found in topic position when it occurs alone as in (32), the presence of a modal adverb seems to block the possibility of *madalas* in topic position as seen in (35-37).

(35) Madalas ay b-um-isita   si Juan
often   TOP AF.PRF-visit P.SUB Juan
'Juan visited often.'

(36) *Madalas ay b-um-isita   sana   si   Juan
often   TOP AF.PRF-visit OPTM P.SUB Juan

(37) *Madalas ay b-um-isita   yata si   Juan.
often   TOP AF. PRF-visit EPST P.SUB Juan

That this is not a categorical constraint on the co-occurrence of these adverbs is seen from (38) and (39) where the frequency adverb is attached within TP.

(38) B-um-isita   sana si Juan nang madalas
AF. PRF-visit OPTM P.SUB Juan *nang* frequent
'Hopefully, Juan visited often.'

156

(39)   B-um-isita   yata si   Juan nang madalas
       AF. PRF-visit EPST P.SUB Juan *nang* frequent
       'Perhaps Juan visited often.'

Unlike Cinque's similar examples in (1-5) above, there can be no Relativized Minimality account here for two reasons. First, however these adverbs are characterized, it is doubtful that *yata* can be the same kind of element as *madalas* since *madalas* can appear in Spec positions and may be modified while *yata* cannot. Second, as a prosodically positioned Wackernagel clitic, *yata* is ordered at PF and therefore cannot interfere with syntactic movement.

These facts are expected under Ernst's analysis which does not rely on movement but rather on the FEO calculus to rule out sentences such as (36) and (37). Specifically, the presence of modal adverbs (*yata, sana*) convert the FEO type to Fact while a frequency adverb (*madalas*) requires a lower type (e.g., Event/Proposition). The prosodic domain within which the clitics are positioned is the TP, *excluding* the Topic. It is therefore unambiguous that the adverb in topic position must be composed after clitics within TP.

4.2   Concentric scope phenomena

Key evidence for Ernst's theory comes from scope facts and, in particular, data which shows that outer adverbs scope over inner adverbs on both sides of the predicate. The predictions of symmetric semantic composition are borne out by the Tagalog data. Observe the relative scope of scope of *bigla* 'suddenly' and *lagi* 'always' within the TP (40-41). When both adverbs precede the verb, the leftmost adverb must scope over the adverb to its right.

(40)   Biglang   laging   na-lulungkot si   Juan
       suddenly-LNK always-LNK   STA-IMPF.sad P.SUB Juan
       'Suddenly, Juan is always sad.'   (*suddenly*>*always* only)

(41)   Laging   biglang   na-lulungkot si   Juan
       always-LNK   suddenly-LNK STA-IMPF.sad P.SUB Juan
       'Juan is always suddenly sad.'   (*always*>*suddenly* only)

Crucially, when the adverbs "sandwich" the predicate as in (42), either reading is available. This shows that adverb scope does not necessarily abide by an asymmetric left-to-right hierarchy.

(42)   Biglang   na-lulungkot   si   Juan lagi
       suddenly-LNK STA-IMPF.sad P.SUB Juan always (ambiguous)

On the right edge of the clause we find the same effects. When two (non-clitic) adverbs appear following the verb, the outer one must scope over the inner. Compare the relative scope of the adverbs *dalawang beses* 'twice' and *di-sinadya* 'unintentionally' in (43) and (44).

(43)   S-in-ingil   nila   ako   nang   di-sinadya   nang dalawang beses
       PV.PRF-charge 3p.GEN 1s.SUB *nang* NEG-intentional *nang* two-LNK times
       'They charged me unintentionally twice.'   (*twice*>*unintentionally* only)

157

(44)   S-in-ingil   nila   ako   nang dalawang beses nang di-sinadya
       PV.PRF-charge 3p.GEN 1s.SUB *nang* two-LNK times *nang* NEG-intentional
       'They charged me twice unintentionally.'   (*unintentionally>twice* only)

When the adverbs flank the verb both readings are possible as seen in (45).

(45)   Dalawang beses nila   akong   s-in-ingil   nang di-sinadya
       two-LNK times 3p.GEN 1SUB-LNK PV.PRF-charge *nang* NEG-intentional
       'They twice charged me unintentionally.' (ambiguous)

Under the *Relative* theory, the ambiguity of (45) is predicted since such a configuration may correspond to two different bracketings in the FEO structure:

(46)   [ᴇ·Dalawang beses [ᴇnila akong siningil nang di-sinadya]] = (43)

(47)   [ᴇ·[ᴇDalawang beses nila akong siningil] nang di-sinadya] = (44)

Cinque (1999:25) discusses similar English data from Andrews 1983, reanalyzing the original adjunction derivation along the lines of the specifier theory. To capture the scope facts of English (48) he posits the existence of two separate base positions for *twice*, a higher position with an "iterative" meaning and a lower position with a "repetitive" meaning. In (48a), the entire clause has raised to the left of the iterative *twice*.

(48) a. John knocked on the door intentionally twice. = *twice>intentionally*
     b. John knocked on the door twice intentionally. = *intentionally>twice*

By positing two projections for *twice*, Cinque seems to mistake a simple scope alternation for a more fundamental difference in meaning. His distinction between iterative and repetitive only emerges clearly when there is a second adverbial to scope under or over. Furthermore, his approach requires positing multiple phrases for every adverb that displays similar concentric effects (e.g. *almost, again, always* etc.) leading to an almost endless array of projections with essentially identical content. From a technical standpoint, there is also the problem of explaining why the higher *twice* is marked in its base position even for the iterative meaning ("John twice knocked on the door"). Unmotivated movement of the VP plus PP constituent is necessary to derive the unmarked order.

Another point for *Relativity* comes from the fact that two readings of *often*, which Cinque attributes to base generation in two different functional projections, are both present in the right edge of the clause in the reverse order which Cinque posits. [8] Because of the *nang* marker, lower and higher attachment of adverbs is overtly distinguished on the right periphery. In the higher attachment in (49), the adverb takes wide scope over negation while in (50) it takes narrow scope.

(49)   Hindi siya   s-in-ingil(,)   dalawang beses
       NEG   3s.SUB PV.PRF-charge two-LNK times
       'He wasn't charged twice.'   (*twice>NEG*)

(50) Hindi siya   s-in-ingil   nang dalawang beses
     NEG   3s.SUB PV.PRF-charge *nang* two-LNK times
     'He wasn't charged twice.'   (NEG>*twice*)

There seems to be no principled explanation why "repetitive" *twice*
could not raise over negation to FocP (51a) or TopP (51b) with the adverbial
being interpreted in its reconstructed position (as suggested by Cinque) beneath
negation. However, the data shows that this is impossible since *twice* can only
be interpreted in (51) with wide scope in regard to negation.

(51) a. Dalawang beses siya   hindi s-in-ingil
        two-LNK   times 3s.SUB NEG PV.PRF-charge   (*twice*>NEG)
     b. Dalawang beses ay hindi siya   s-in-ingil
        two-LNK   times TOP NEG 3s.SUB PV.PRF-charge   (*twice*>NEG)

Again, this is predicted by the linear order of the adverbial and negation in
regard to the verb. Because negation must be composed with the predicate
before *twice*, the former must take narrow scope.

4.3 Epistemic versus deontic readings of modals

Similar to the repetitive/iterative division, Cinque proposes that epistemic and
deontic readings of modals are the result of base generation in two different
functional projections. He takes the fact that two instances of the same modal
can occur simultaneously as evidence for the existence of two separate
functional projections. Thus, sentences such as (52) and (53) could be taken to
show that the epistemic phrase is to the left of negation while the deontic
(alternatively, "alethic") is to the right.

(52) Dapat hindi sila   dapat mag-aral
     must NEG   3p.SUB must AV.INF-study
     'It should be that they do not have to study.'

(53) Maaaring hindi sila   maaaring mag-aral
     can-LNK   NEG 3p.SUB can-LNK AV.INF-study
     'It is possible that they are unable to study.'

As Ernst notes, this necessarily treats the formal similarity between
epistemic and deontic modals across languages as accidental homophony.[9] It
appears more felicitous to treat modals as yielding an epistemic reading when
taking a higher FEO complement (i.e. Proposition) and a deontic reading when
taking a lower FEO complement (i.e. Event). Therefore, when modals are
stacked on one side of the predicate, the outer modal may only receive an
epistemic reading while the inner modal, a deontic one.

It must also be noted that many deontic/alethic modals in Tagalog (as in
many other languages, cf. de Haan 1997) can appear on either side of negation
with transparent scope relations, as found in (54-55).

(54) Puwede akong   hindi ma-tulog.
     can   1s.SUB-LNK NEG STA.INF-sleep
     'I'm able to not sleep.'   (deontic *can*>NEG)

159

(55) Hindi ako   puwedeng ma-tulog.
      NEG 1s.SUB can-LNK   STA.INF-sleep
      'I'm unable to sleep.'   (NEG>deontic *can*)

Thus, a multiple base generation approach must still rely on movement to capture all the facts while a free adjunction approach accounts for the freedom and scope transparency simultaneously.

5.0 Conclusion and prospects

The evidence presented here argues strongly for a *Relative* approach to adverbs along the lines of Ernst 2002. Because Tagalog has a wide variety of means for introducing adverbs into the clause and a large number of clitic adverbs, we were able to differentiate between different predictions made by Cinque and Ernst in regard to multiple adverb facts. It was seen that a syntactic (Relativized Minimality) account of certain ungrammaticalities fails in Tagalog while a more semantically based account explains the data parsimoniously. Evidence was also shown in support of treating concentric effects as the norm in adverbial interpretation. Under this view, right adjunction is necessary to account for the different interpretive possibilities of adjuncts on the right edge of the clause.

Taking concentric layering to be a natural outcome of the FEO calculus, the "inverse" ordering of adverbs in Malagasy (Rackowski 1998) and Seediq (Holmer this volume) in the postverbal domain is less surprising then previously considered. The following order for Malagasy adverbs in relation to the verb is given by Rackowski (1998) as (56).

(56) *Na*(*dia*) > *Matetika* > *Tsy* > *Efa* / *Mbola* > *Tsy* > VERB > *Tsara* >
      'Even'   generally   NEG   already still   NEG   well

      Tanteraka > *Foana* > *Intsony* > *Mihitsy* > *Aza* > *Ve*
      completely   always anymore   at-all   though SPEECH ACT (Q)

In the preverbal domain the adverbs are ordered as predicted by Cinque's hierarchy but in the postverbal domain the order appears to be reversed.

Concentric effects are clearer in Malagasy than in Tagalog since Malagasy has few or no mitigating prosodic factors such as (prosodically determined) clitic movement. Furthermore, Malagasy seems to treat all adverbs similarly in terms of how they are introduced into the clause; topicalization and other peripheral positions are less common than in Tagalog. Taking this into account, the FEO predicts this order naturally in a verb initial language. Additionally, there is evidence that adverbs within a single class are in fact permutable in Malagasy with corresponding differences in scope (cf. Rackowski 1998:10). Adverb order should therefore not be taken as a strong argument for a "roll-up"/intraposition derivation (Pearson 2001) in-lieu of evidence that it is possible for a verb-initial language to display the Cinque order of adverbs in the postverbal domain.

160

Endnotes

1. Cinque allows for unmotivated movement through an extensive array of functional projections and multiple loci for base generation of certain adverbs. Ernst on the other hand introduces stipulatory syntactic features such as [+Right] which is satisfied by linearization to the right of a certain head and PF requirements on "heavy/light" items which have no basis in the phonological content of the items themselves.

2. I choose these terms over the more common Tight-fit versus Loose-fit since these terms underemphasize a crucial distinction between the two approaches: the predictions of relative scope between adverbs.

3. Relativized Minimality as envisioned in Rizzi 2002 takes the more specific categories quantificational, modificational, topic and argument to be relevant in calculating minimality.

4. It must be noted that the judgments above can also be predicted by Ernst's theory since a frequency adverb must take an Event as its FEO complement while a modal adverb such as probably raises the FEO type to Proposition, which cannot be subsequently lowered to satisfy the outer adverb. Although the two theories cannot be differentiated on the basis of the above data, there are, in principle, cases which force different predictions from these two approaches as will be shown later.

5. Of course, these "types" of movements have all been taken to be instantiations of a single rule since the advent of GB. The empirical prediction underlying the rejection of "adverb movement" as an independent phenomenon is that adverbs which are not interpreted in their surface position are displaced because of positional requirements on topicalized, focused or wh- constituents.

6. Abbreviations used: AV – actor voice; COMP – complementizer; EMPH – emphatic; EPST – epistemic; GEN – genitive; IMPF – imperfective aspect; INF – infinitive; IRR – irrealis; LNK – linker; LV - locative voice; NEG – negation; NOMZ – nominalizer; OBL – oblique case; OPT – optative; P – personal/[+human] (case marker); PL – plural; POL – politeness marker; PV – patient voice; Q – question marker; RHET – rhetorical question marker; STA – stative; SUB – 'subject' (absolutive/nominative case); SURP – surprise; TOP – topic marker; TR – transitive;

7. Clitics can intervene since they are ordered post-syntactically.

8. There is no evidence from binding or word order in Tagalog for a roll-up/intraposition derivation as Pearson (2001) posits for Malagasy and other "inverse" predicate initial-languages.

9. More recently, Cinque (2004) addresses the homophony problem by suggesting that certain adverbs are 'underspecified' for certain features (e.g. epistemic/deontic) and can thus be compatible with two different positions. However, the fact that these 'underspecified' adverbs are still base-generated separately leaves open the question of what, precisely, their connection is.

References

Alexiadou, Artemis. 1997. Adverb placement. Amsterdam: John Benjamins.

Andrews, Avery. 1983. A note on the constituent structure of modifiers. Linguistic Inquiry 14. 695-697.

Cinque, Guglielmo. 1999. Adverbs and functional heads: a cross-linguistic perspective. Oxford: Oxford University Press.

Cinque, Guglielmo. 2004. Issues in adverbial syntax. Lingua 114. 683-710.

Ernst, Thomas. 1984. Towards in integrated theory of adverb position in English. Bloomington, IN: Indiana University Linguistics Club.

Ernst, Thomas. 2002. The syntax of adjuncts. Cambridge: CUP.

Ernst, Thomas. 2004. Principles of adverbial distribution in the lower clause. Lingua 114. 755-777.

de Haan, Ferdinand. 1997. The interaction of modality and negation. New York: Garland Press.

Hale, Kenneth, Keyser, Samuel-Jay. 1993. On argument structure and the lexical expression of syntactic relations. The view from building 20. Eds. Kenneth Hale and Samuel-Jay Keyser. MIT Press, Cambridge, MA, pp. 53–109.

Holmer, Arthur. 2004. Interposition and Formosan adverbial heads. This volume.

Jackendoff, Ray. 1972. Semantic interpretation in generative grammar. Cambridge, MA: MIT Press.

McConnell-Ginet, Sally. 1982. Adverbs and logical form: a linguistically realistic theory. Language 58. 144-84.

Pearson, Matthew. 2001. Two types of VO languages. VO and OV, ed. Peter Svenonius.

Rackowski, Andrea. 1998. Malagasy adverbs. The structure of Malagasy, volume II, ed. Ileana Paul, UCLA Working Papers in Linguistics.

Rizzi, Luigi. 1990. Relativized minimality. Cambridge, MA: MIT Press.

Rizzi, Luigi. 2002. Locality and left periphery. ms, University of Siena (to appear in Belletti, A. (Ed.) (in press).

Svenonius, Peter. 2001. The derivation of VO and OV. Amsterdam: John Benjamins.

Daniel Kaufman
Department of Linguistics
Cornell University
Ithaca, NY 14853-4701 U.S.A.
dak37@cornell.edu

Fixed Segmentism in Palauan Multiple Reduplications

Tomoko Kawamura
State University of New York at Stony Brook

1. Introduction

Palauan is a language with two reduplicative morphemes, a CVCV-reduplicant (1) and a Cɛ-reduplicant (2). The prefix /mə-/ is a verb marker. Reduplicants are marked with underlines for CVCV-reduplicants and double underlines for Cɛ-reduplicants.

(1) *CVCV-reduplication* (Josephs 1990)

unreduplicated form		reduplicated form	
məloʔəd	'break'	mə<u>ləʔ</u>əlóʔəd	'easily broken'
mətɛŋəl	'come down'	mə<u>təŋ</u>ətɛ́ŋəl	'keep coming down'
məsúʔəd	'talk harshly'	mə<u>səʔ</u>əsúʔəd	'always talk harshly'

(2) *Cɛ-reduplication* (Josephs 1990)

unreduplicated form		reduplicated form	
mətəgói	'be talked to'	mə<u>tɛ</u>təgói	'easy to talk to'
məsulául	'sleepy'	mə<u>sɛ</u>sulául	'kind of sleepy'
məʔúʔəp	'cloudy'	mə<u>ʔɛ</u>ʔúʔəp	'rather cloudy'

McCarthy and Prince (1994) have argued that partial reduplication can be described without using templates. Because faithfulness requirements on input-output correspondence do not apply to the reduplicant, what emerges in the reduplicant is the least marked structure, as defined by the phonological constraints of the language.

Languages that have different reduplicative morphemes with different shapes appear to be problematic for this claim. However, Urbanczyk (1999) has claimed that in a language which has multiple reduplications, the reduplicative morphemes are categorized as either roots or affixes and the two reduplicative morphemes are realized differently because faithfulness requirements on roots are stronger than general faithfulness requirements. Therefore, according to this claim, root reduplicants have more faithful structures than affix reduplicants where the shape of reduplicants differs.

If Urbanczyk's claim is correct, we should be able to choose one of the Palauan reduplicant as a root and the other as an affix. When the ROOT MAX BR dominates some markedness constraints on the minimal word which dominates the general MAX BR, as ROOT MAX BR » markedness (minimal word) » MAX BR, more segments are copied in the root reduplicant than in the affix reduplicant. Thus, the length of the reduplicants suggests that the CVCV-reduplicants are roots and Cɛ-reduplicants are affixes. When we look at the vowels in the reduplicants, however, the default vowel [ə] appears in the CVCV-reduplicants, but the more marked vowel [ɛ] always appears in the Cɛ-reduplicants. The fixed segment [ɛ]

in the Cɛ-reduplicant appears to be problematic not only for the root/affix account, but for Alderete et. al's (1999) analysis of fixed segments in reduplication, as well. Alderete et. al argue that phonologically fixed segments in reduplicants represent the emergence of the unmarked (TETU). Thus, the fixed segment of Palauan reduplicants should be the default vowel [ə], but in fact, the fixed segment in the Cɛ-reduplicant is [ɛ].

In this paper, I argue that this apparent problem is accounted for by the interaction of constraints. For the fixed segment [ɛ] in Cɛ-reduplication, I argue that [ɛ] is the second least marked vowel in Palauan, which appears when the default vowel [ə] cannot appear. I show that the Palauan facts are not only consistent with the proposals of Urbanczyk (1999) and Alderete et. al (1999), but they actually provide support of their claims. In the following section, I discuss Urbanczyk's (1999) arguments concerning ROOT faithfulness in reduplication and possible asymmetries between affix reduplicants and root reduplicants. In Section 3, I introduce Palauan reduplication and discuss Finer's (1986) observations on the resulting state verb (RSV) form. I show that the RSV forms support the classification that Cɛ-reduplicants are affixes, and CVCV-reduplicants are roots. In Section 4, I discuss the shape and vowel quality of the two reduplicants. The CVCV-reduplicant has three variants: CəCə, CəC and CV. I explain this variation, illustrating why [ə] appears in the first two variations. Then, I discuss the shape and vowel quality of the Cɛ-reduplicant, arguing that the fixed segment [ɛ] in Cɛ-reduplication is a special case of TETU. I show that root faithfulness constraints are crucial in determining the shape and vowel quality of the reduplicants. Section 5 is the conclusion.

2. Root/Affix Asymmetries

Steriade (1995) pointed out that roots allow more marked structure than affixes. Beckman (1997) translated this observation into correspondence theory, proposing two types of correspondence relations: general correspondence and restricted correspondence, including root faithfulness constraints. These faithfulness constraints may require roots to be more faithful than affixes.

Urbanczyk (1999), examining reduplication in Lushootseed, argued that interaction of ROOT faithfulness constraints on base-reduplicant (ROOT BR) faithfulness, general faithfulness constraints, and markedness constraints explains the realization of two types of reduplication. When a ROOT BR faithfulness constraint dominates some markedness constraint, the root reduplicant can have the marked structure. However, Root BR faithfulness constraints do not say anything about the affix reduplicant. So, if the markedness constraint which is dominated by the ROOT BR faithfulness constraint dominates the general BR faithfulness constraints, the marked structure cannot appear in the affix reduplicant and we get two distinct shapes of reduplicants. Following is her Lushootseed examples.

(3) a. *Lushootseed Distributive (DIST) Reduplication* (Urbanczyk 1999)

root		**reduplication**	
sáqʷ	'fly'	sáqʷsaqʷ	'fly here and there'
jə́səd	'foot'	jə́sjəsəd	'feet'
lə́gʷəb	'young man"	lə́gʷləgʷəb	'young men'

b. *Lushootseed Diminutive (DIM) Reduplication*

root		**reduplication**	
ʔálʔal	'house"	ʔáʔalʔal	'hut'
jə́səd	'foot"	jíjəsəd	'little foot"
tədᶻíl	'lie in bed"	títədᶻil	'lie down for a little while'

In Lushootseed, [ə] in the base is preserved in the DIST reduplicant with stress as [ə́], but it cannot appear in the DIM reduplicant, which has the stressed default vowel [í]. Urbanczyk analyzes the DIST reduplicants as roots and the DIM reduplicants as affixes. Because the ROOT BR-faithfulness constraint (ROOT Ident BR) dominates the markedness constraint *[ə́], which bans stressed schwas, [ə́] can appear in the root reduplicants, but not in the affix reduplicants.

(4) a. *Lushootseed DIST (root) reduplication*

REDᵣₒₒₜ-/jəsəd/	ROOT IDENT BR	*[ə́]	IDENT BR
☞a. jə́sjəsəd		*	
b. jísjəsəd	*!		*

b. *Lushootseed DIM (affix) reduplication*

REDₐffix-/jəsəd/	ROOT IDENT BR	*[ə́]	IDENT BR
a. jə́jəsəd		*!	
☞b. jíjəsəd			*

The difference in the morphological status of DIST and DIM reduplicants accounts for the occurrence of two types of reduplication in Lushootseed.

This analysis has one important implication. Both ROOT BR faithfulness constraints and general BR faithfulness constraints require the root reduplicant to be identical to the base. However, only general BR faithfulness constraints require the affix reduplicant to be identical with the base. Therefore, root reduplicants can be more faithful to the base than affix reduplicants. Furthermore, this analysis implies that root reduplicants can have more marked structure than affix reduplicants, but not vice versa.

If one reduplicant is an affix and the other a root in Palauan, we predict that root reduplicants will be more faithful and will allow more marked structures than affix reduplicants. In the next section, I examine the question whether Palauan reduplications are consistent with these claims.

3. Two types of Reduplication

3.1. Palauan Reduplication

Palauan is a Western Malayo-Polynesian language, spoken in the Republic of Palau by 15,000 speakers. It has 6 vowels ([ə] and 5 full vowels; [i], [u], [ɛ], [o], and [a]). [ə] is the default vowel and it cannot bear stress (Wilson 1972). As shown above, Palauan has two types of reduplication. The CVCV-type has several variants, namely CəCə, CəC, and CV.[1]

(5) *CVCV-reduplication* (Josephs 1990)
    a. *CəCə-form*

unreduplicated form		reduplicated form	
mərɛ́bək	'grope at'	mə_rəbə_rɛ́bək	'grope around'
məsúʔəd	'talk harshly'	mə_səʔə_súʔəd	'always talk harshly'

    b. *CəC-form*

unreduplicated form		reduplicated form	
mərám	'be mixed'	mə_rəm_rám	'easy to mix'
məsóŋdə	'break'	mə_səŋ_sóŋdə	'keep breaking'

    c. *CV-form*

unreduplicated form		reduplicated form	
məŋióklə	'be cooked'	mə_ŋi_ŋióklə	'easy to cook'
məsúub	'get studied'	mə_su_súub	'easy to study'

In most cases, we can predict which form appears from the shape of the base form. In Cɛ-reduplication, the reduplicant always consists of [ɛ] and a copy of the first consonant of the root.

(6) *Cɛ-reduplication* (Josephs 1990)

unreduplicated form		reduplicated form	
mətəgói	'be talked to'	mə_tɛ_təgói	'easy to talk to'
məʔúu	'shady'	mə_ʔɛ_ʔúu	'fairy shady'
məsulául	'sleepy'	mə_sɛ_sulául	'kind of sleepy'

Note that the default vowel [ə] does not appear, but the more marked vowel [ɛ] appears in the Cɛ-reduplicant.

3.2. The Resulting Stative Verb (RSV)-form

Finer (1986) observes that the resulting state verb (RSV) marker is realized differently in CVCV-reduplication and Cɛ-reduplication. In the unreduplicated form, the RSV marker /-l-/ appears just after the first consonant of the root.

(7) *RSV form: RSV morpheme /-l-/ + root* (Finer 1986,[2] Josephs 1990)

simple form		RSV form	
bóes	'gun'	b_l_óes	'shot'
ŋábək	'planing'	ŋ_l_ábək	'planed'
káud	'dam'	k_l_áud	'dammed'
lɛ́dəs	'path'	l_l_ɛ́dəs	'stretched'

166

When the RSV infix /-*l*-/ appears within the reduplicated forms, the CVCV-reduplication and Cɛ-reduplication show different patterns. In the CVCV-reduplicated form, the RSV morpheme appears after the first consonant of the reduplicant. In the Cɛ-reduplicated form, it appears after the first consonant of the base.

(8) *RSV CVCV-reduplicated form* (Josephs 1972, Finer 1986)

RSV unreduplicated form		**RSV reduplicated form**	
ŋ*l*ábək	'ironing'	ŋ*l*əbəŋábək	'scraped all over'
b*l*íiʔ	'filter'	b*l*ibíiʔ	'sorted out by type'

(9) *RSV Cɛ-reduplicated form* (Josephs 1972, Finer 1986)

RSV unreduplicated form		**RSV reduplicated form**	
ŋ*l*əmúl:ə	'cut'	nɛŋ*l*əmúl:ə	'not well cut'
b*l*ɛóp	'shaped'	bɛb*l*ɛóp	'round in shape'

Given that CVCV-reduplicants are roots and Cɛ-reduplicants are affixes, we can generalize the pattern in (8)-(9); the RSV affix appears after the first consonant of the first <u>root</u>. Finer's (1986) observations on the morphological differences of the two reduplications are then captured as a result of the root/affix asymmetry. I propose the following morphological structures for unreduplicated and reduplicated forms:[3]

(10) a. *unreduplicated form*   b. *CVCV-reduplicated form*   c. *Cɛ-reduplicated form*

```
      [məkeáld]              [məɾəbəɾébək]              [məʔɛʔúʔəp]
        GrWd                     GrWd                       GrWd
       /    |                  /       \                  /       \
   affix   root           affix        GrWd          affix        GrWd
     |       |              |         /     \           |        /     \
    VM       |             VM      root     root       VM    affix      root
   [mə]   [keáld]         [mə]     RED       |         [mə]    RED        |
                                 [ɾəbə]   [ɾébək]              [ʔɛ]    [ʔúʔəp]
```
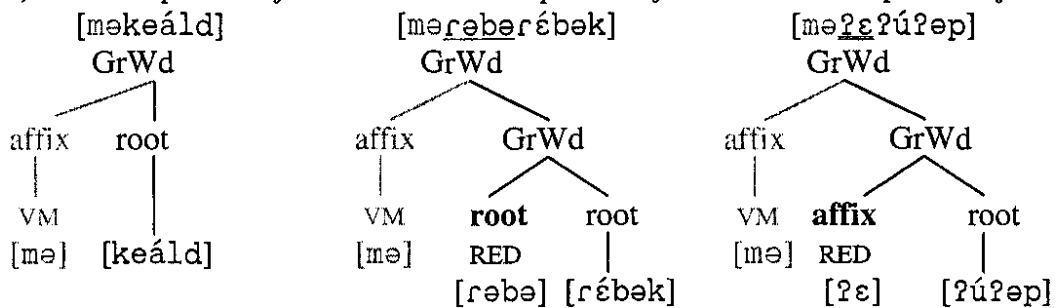
In the unreduplicated RSV form [*kleáld*], the RSV marker /-*l*-/ appears after the first consonant. In the CVCV RSV reduplicated form [*ŋləbəŋábək*], the RSV marker appears after the first consonant of the first root. In the Cɛ-reduplicated RSV form [*nɛŋləmúl:ə*], the RSV marker appears after the first consonant of the root (base).[4]

4. Apparent Problem: Vowel Quality

So far, I have argued that the CVCV-reduplicants are roots and the Cɛ-reduplicants are affixes. Urbanczyk (1999) argues that the root reduplicant may have more marked structure than the affix reduplicant. This claim implies that the Cɛ-reduplicant should have a <u>less</u> marked structure than the CVCV-reduplicants.

167

However, the least marked vowel (default vowel) [ə] appears in the CVCV-reduplicants, but it never appears in the Cɛ-reduplicants. Further, the fact that the affix reduplicant always has [ɛ] appears to be problematic for Alderete et. al's (1999) analysis of phonological fixed segments, which argues that the fixed segments in the reduplicants should be the default vowel and should be in the TETU condition.

In this section, I argue that the Cɛ-reduplicants have less marked structure than the CVCV-reduplicants and propose that [ɛ] is the second least marked vowel, which appears as a TETU effect when the default vowel [ə] cannot appear. I first discuss CVCV-reduplication and then discuss Cɛ-reduplication.

4.1. CVCV-reduplication and its Variations

As shown above, CVCV-reduplication is realized in three different forms: CəCə, CəC, and CV. The examples are repeated below:

(11) a. *CəCə-form*

unreduplicated form		reduplicated form	
mərɛ́bək	'grope at'	mərəbərɛ́bək	'grope around'
məsúʔəd	'talk harshly'	məsəʔəsúʔəd	'always talk harshly'

   b. *CəC-form*

unreduplicated form		reduplicated form	
mərám	'be mixed'	mərəmrám	'easy to mix'
məsóŋdə	'break'	məsəŋsóŋdə	'keep breaking'

   c. *CV-form*

unreduplicated form		reduplicated form	
məŋióklə	'be cooked'	məŋiŋióklə	'easy to cook'
məsúub	'get studied'	məsusúub	'easy to study'

When the first two syllables of the base are CV.CV(C)., the reduplicant is Cə.Cə., as (11a). When the first syllable of the base is CVC, this syllable is duplicated with the vowel reduced to [ə], as CəC in (11b). When the first two syllables of the base are CV.V(C), then the reduplicant is CV, copying the first consonant and either the first or second vowel, as (11c).[5]

To capture this close relation between the shape of the reduplicant and vowel quality, I propose that the default vowel [ə] and coda each carry one mora, while other full vowels, such as [i], [ɛ], [a], [u], and [o], carry two moras.[6] Assuming that the root-reduplicant must be a foot (ROOT=FOOT) and a foot contains two moras (FT-BN (μ)), the CVCV-reduplicant consists of two moras. To make the reduplicant bi-moraic, there are two options: copying two syllables, while reducing the full vowel to [ə], or copying one syllable while keeping the full vowel. So, the possible candidates for /mə+RED$_{root}$+rɛ́bək/ would be [mərəbərɛ́bək] and *[mərɛrɛ́bək]. Two syllables are copied in the first candidate and only one syllable is copied in the second. Since [*mərəbərɛ́bək*] is the correct form, the elements of the base must be copied as far as possible. These requirements on the reduplicant are translated into the following constraints:

168

(12) a. FT-BN (μ): Every foot must have exactly two moras.

    b. ROOT=FOOT: Every root must contain a foot.

    c. PARSE-σ: Every syllable must be parsed into a foot.

    d. ROOT MAX BR: Every element of the base has a correspondent in the root reduplicant.

The following tableau illustrates how these constraints work. The numbers under the segments indicate correspondence relations.

(13) /mə/-RED_{root}-/rɛ́bək/ 1 2 3 4 5	MAX IO, IDENT IO	ROOT =FOOT	FT-BN (μ)	ROOT MAX BR	PARSE-σ
a. mə(rɛbə)(rɛ́bək) 1234 12345			**!	*	*
b. mə(rɛ)(bək)(rɛ́bək) 12 345 12345	*!		*		*
c. mə(rɛ)(rɛ́bək) 12 12345			*	***!	*
☞d. mə(rəbə)(rɛ́bək) 1234 12345			*	*	*

The candidate (13a) is ruled out for two violations of FT-BN(μ) because the reduplicant has three moras: two for [ɛ] and one for [ə]. (13b) is ruled out by ROOT=FOOT, since the reduplicant equals two feet, not one foot. While the reduplicants in both (13c) and (13d) satisfy FT-BN(μ), (13d) wins because of fewer violations of ROOT MAX BR.

When the first syllable of the base has a coda, the reduplicant is realized with the CəC form. Given that the coda is moraic in Palauan (Wilson 1972), the grammar in (13) predicts the CəC form, as shown in (14).

(14) /mə/-RED_{root}-/rám/ 1 2 3	MAX IO, IDENT IO	ROOT =FOOT	FT-BN (μ)	ROOT MAX BR	PARSE-σ
a. mə(ram)(rám) 123 123			**!		*
b. mə(ra)(rám) 12 123			*	*!	*
☞c. mə(rəm)(rám) 123 123			*		*

In (14a), every element in the base is copied in the reduplicant. Since the full vowel [a] carries two moras and the coda [m] carries one mora, the reduplicant has three moras, violating Ft-Bn(μ). The reduplicants in (14b) and (14c) satisfy Ft-Bn(μ). Since more elements in the base are copied in (14c) than in (14b), the candidate (14c) wins.

In the CV-form, such as [məŋiŋióklə], the first syllable of the base is CV and the second syllable starts with vowel. Since the reduplicant must have two

moras, the possible candidates are [məɲiŋióklə] and *[məɲəəɲióklə]. If the
ONSET constraint (Syllable must have an onset) is ranked between MAX IO and
ROOT MAX BR, the onsetless syllable appears only in the reduplicant. The
grammar for [məɲiŋióklə], is shown in the following tableau:

(15) /mə/-RED_root- /ŋiókl/ 1 2 3 4 5	MAX IO, IDENT IO	ROOT =FOOT	FT-BN (μ)	ONSET	ROOT MAX BR
a. mə(ŋə)(ŋióklə) 1 2 1 2 3 4 5 6			**!	*	****
☞ b. mə(ŋi)(ŋióklə) 1 2 1 2 3 4 5 6			*	*	****
c. mə(ŋəə)(ŋióklə) 1 2 3 1 2 3 4 5 6			*	**!	***
d. mə(ŋio)(ŋióklə) 1 2 3 1 2 3 4 5 6			**!	**	***

In (15b) and (15c), the reduplicant has two moras. (15c) has an onsetless syllable
in the reduplicant. So, the constraint ONSET chooses (15b) over (15c). The
introduction of ONSET does not affect (13) and (14), since every relevant syllable
has an onset.

4.2. Cε-reduplication

I have argued that CVCV-reduplication is root-reduplication and that the
reduplicant sometimes has the least marked vowel [ə].[7] Urbanczyk (1999) argues
that affix reduplicants tend to be less marked than root reduplicants because of
root faithfulness constraints. Therefore, Urbanczyk's analysis predicts that no
vowel besides [ə] can appear in affix reduplicants. Alderete et. al (1999) also
argue that the phonological fixed segment must be the default segment. Thus the
fixed segment in the affix reduplicant should be [ə]. However, the examples in
(16) show that this is not the case.

(16) Cε-reduplication (Josephs 1990)

unreduplicated form		reduplicated form	
məʔúʔəp	'cloudy'	məʔɛʔúʔəp	'rather cloudy'
mətegói	'be talked to'	mətɛtegói	'easy to talk'
məsóðəl	'tear'	məsɛsóðəl	'easily torn'
mətoiáklə	'tall'	mətɛtoiáklə	'fairly tall'

The vowel in the Cε-reduplicant is always [ε], which is more marked than [ə].[8]

Because of the constraint ROOT=FOOT, the base is a foot. If the left edge of
the affix reduplicant must coincide with the left edge of the foot (Align L
(affix_RED, foot)), the affix reduplicant must be a foot with two moras. The possible
reduplicants for the base [ʔúʔəp] are then [ʔε], *[ʔəʔ], and *[ʔəʔə]. Since the
Cε-reduplicant has affix status, ROOT MAX BR is irrelevant here. When we
compare [ʔε] and *[ʔəʔ], we notice that the reduplicant [ʔε] does not have coda,
but *[ʔəʔ] has a coda. When we compare [ʔε] and *[ʔəʔə], we notice that [ʔε]

has fewer syllables. Thus, the constraint which bans codas and the constraint which limits the number of syllables work crucially in Cε-reduplication.

(17) a. NO CODA: Codas are not allowed. (Kager 1999)
   b. NO STRUC-σ: Syllable structure is not allowed.

(Prince and Smolensky 1993)

Since the correct reduplicant is one syllable without a coda, NO STRUC-σ and NO CODA dominate MAX BR. The ranking of constraints is shown in the following tableau.

(18) /mə/-RED_{affix}-/ʔúʔəp/	FT-BN(μ)	ROOT MAX BR	NO CODA	NO STRUC-σ	MAX BR
a. mə(ʔəʔ)(ʔúʔəp)	*		**!	****	**
b. mə(ʔə)(ʔúʔəp)	**!		*	****	***
c. mə(ʔəʔə)(ʔúʔəp)	*		*	*****!	*
☞d. mə(ʔɛ)(ʔúʔəp)	*		*	****	***

The candidate (18a) violates NO CODA twice. The reduplicant in (18b) has only one mora. Thus, they are ruled out. The reduplicants in (18c) and (18d) have two moras. Since ROOT MAX BR does not apply to the affix reduplicant, NO STRUC-σ chooses (18d). Notice that ROOT MAX BR dominates the markedness constraints NO STRUC-σ/ NO CODA. This ranking enables a coda and two syllables in the root reduplicant. At the same time, NO STRUC-σ and NO CODA dominate (general) MAX BR, so this grammar disallows [ə] to appear in the affix reduplicant.

So far, the shape of the Cε-reduplicants and the non-occurrence of [ə] are accounted for, but it is not yet explained why the vowel in the reduplicant should always be [ε]. I propose that [ε] is the second least marked vowel in Palauan and it appears under the TETU condition in which the default vowel [ə] cannot appear. [ə] does not have place features, such as [-back], [+high], [+low], or [+round]. Thus [ə] never violates featural markedness constraints, such as *[-B(ack)], *[+H(igh)], *[+L(ow)] and *[+R(ound)]. [ε], on the other hand, has the feature [-back], violating *[-B]. Thus, [ε] is more marked than [ə]. However, as we saw above, the default vowel [ə] cannot appear in the affix reduplicant. If the featural markedness constraints are ranked as *[+H], *[+L], *[+R] » *[-B], then the appearance of [ε] is considered as a TETU effect. This is shown in the following tableau. The candidates in (18a), and (18d) are repeated as (19a), and (19b) respectively.

(19) /mə/-RED_affix- /ʔúʔəp/	NO CODA	*[+R]	*[+L]	*[+H]	*[-B]	MAX BR
a. mə(ʔəʔ)(ʔúʔəp)	**!	*				**
☞ b. mə(ʔɛ)(ʔúʔəp)	*	*			*	***
c. mə(ʔi)(ʔúʔəp)	*	*		*!	*	***
d. mə(ʔa)(ʔúʔəp)	*	*	*!			***
e. mə(ʔu)(ʔúʔəp)	*	**!				***

Among full vowels, [ɛ] is the least marked vowel, and hence (19b) is the output form. This tableau shows that the emergence of [ɛ] in the Cɛ-reduplicant is a special case of TETU and it indicates that a non-default vowel can be the phonological fixed segment under some special conditions.

Note that even if the first syllable of the base is [ə], [ɛ] appears in the affix reduplicant.

(20) **unreduplicated form**   **reduplicated form**

məʔəsáŋ   'busy'   məʔɛʔəsáŋ   'kind of busy'
məʔəláoð   'content'   məʔɛʔəláoð   'rather content'
mərəʔórəʔ   'steal'   mərɛrəʔórəʔ   'keep stealing'

The realization of [ɛ] in the affix reduplicant can be explained with the above grammar. The tableau for [məʔɛʔəsáŋ] is as follows:

(21) /mə/-RED_affix- /ʔəsaŋ/	FT-BN (μ)	NO CODA	*[+R]	*[+L]	*[+H]	*[-B]	MAX BR
a. mə(ʔə)(ʔəsáŋ)	**!	*		*			***
☞ b. mə(ʔɛ)(ʔəsáŋ)	*	*		*		*	***
c. mə(ʔi)(ʔəsáŋ)	*	*		*	*!	*	***
d. mə(ʔa)(ʔəsáŋ)	*	*		**!			***
e. mə(ʔu)(ʔəsáŋ)	*	*	*!	*	*		***

This tableau shows that [ə] cannot appear in the reduplicant and instead, the next least marked vowel [ɛ] appears in the reduplicant.

In this analysis, [ɛ] is not specified in the input, yet the grammar guarantees the realization of [ɛ]. This grammar shows that the phonological fixed segment can be a non-default vowel, but it is still one case of TETU. Thus, the realization of the fixed segment [ɛ] in Palauan affix reduplication is consistent with Alderete et. al's (1999) analysis of phonological fixed segments. Further, the affix reduplicant is less marked than the root reduplicant, since it copies fewer segments. So, the root-affix asymmetries in the shape and vowel quality follow the implications of Urbanczyk's (1999) analysis.[9]

5. Conclusion

In this paper, I identified the Palauan CVCV-reduplicant as a root and the Cɛ-reduplicant as an affix. While the affix reduplicants have a fixed non-default vowel, which looks problematic for Urbanczyk's (1999) claim and Alderete et. al's (1999) analysis, I showed the emergence of a non-default vowel is also a special case of TETU and the root reduplicants are more faithful to the base than the affix reduplicants due to a root faithfulness constraint.

This analysis implies that there are only two types of reduplicant in each language. While Palauan has CɛCVCV reduplication, in addition to Cɛ-reduplication and CVCV-reduplication, this case can be considered as a combination of the two reduplicative morphemes. However, some languages, such as Korean, have more than two types of reduplication (Cho 1999). These interesting cases need further research.

Endnotes

1. Zuraw (2001) observes CCə form as well, such as [ol-*dmə-dúm*] (keep bobbing to surface of water). It seems that this is one version of the CəC variation. Zuraw argues that this is mostly conditioned by the surrounding consonants. This pattern is not covered in my analysis.

2. All data discussed by Finer (1986) are originally from Josephs (1975), McManus (1977) or Wilson (1972).

3. While Finer (1986) proposes the structures in (i), I propose the structures in (10) to apply Urbanczyk's (1999) analysis.

(i) a. unreduplicated form   b. CVCV-reduplicated form   c. Cɛ-reduplicated form

4. The RSV marker /-l-/ never appears at the left edge, and the left edge of the root is always aligned at the left edge of the prosodic word. At the same time, the RSV marker has to be just after the first consonant of the root. Therefore, the

173

RSV affix is aligned as close to the left edge of the prosodic word as possible with the constraints in (i).

(i) a. ALIGN L (root, PrWd): The left edge of a root must coincide with the left edge of the prosodic word.

   b. ALIGN L (RSV affix, PrWd): The left edge of the RSV affix must coincide with the left edge of the prosodic word.

When we look at the form [kleáld], we notice that the alignment requirement for the root is stronger than the alignment requirement for the RSV affix, favoring the affix inside the root. This indicates that ALIGN L (root, PrWd) dominates ALIGN L (RSV affix, PrWd), as shown in the tableau (ii).

(ii) /l//keáld/	ALIGN L (root,PrWd)	ALIGN L (RSVaffix,PrWd)
☞a. [(kl、eáld)]		*
b. [(、lkeáld)]	*!	
c. [(keáldl、)]		*****!
d. [(kel、áld)]		**!

For the RSV CVCV-reduplicated forms, we need the constraint in (iii) which dominates the constraints in (ii). The tableau is given in (iv).

(iii) ALIGN R (root, PrWd): The right edge of a root must coincide with the right edge of the prosodic word.

(iv)/l/-RED$_{root}$-/rɛ́bək/	ALIGN R (root,PrWd)	ALIGN L (root,PrWd)	ALIGN L (RSVaffix,PrWd)
☞a. [(rl、əbə)(rɛ́bək)]	*****	*****	*
b. [(rəbə)(rl、ébək)]	******!	****	****
c. [(l、rəbə)(rɛ́bək)]	*****	******!	

The same grammar explains the RSV Cɛ-reduplicated form.

5. It is not so clear which vowel is copied. Zuraw (2002) examines the vowel reduction and observes that the same vowel remains in the possessive form, in which vowel reduction takes place. So it seems there is a systematic way, but I do not see the mechanism of vowel reduction in the VV sequences.

6. If we assume that the full vowels have one mora and [ə] and codas do not carry any moras, we have to say that the foot can contain at most one mora for the CV case and we cannot exclude the reduplicant *CəCəC or *CVCəC or *CVC. Further, the assumption that the coda does not have a mora fails to capture Wilson's (1972) observation on stress assignment.

7. When a word has a consonant cluster word finally, sometimes [ə] is inserted.

174

(i) **input**   **simple form** (Josephs 1975)

input	simple form	(Josephs 1975)
/dakt/	dáktə	'fear'
/mələʔolb/	mələʔólbə	'bother'
/kbokb/	kpókpə	'water'
/bsibs/	psípsə	'drill'

Given that the epenthetic vowel is a default vowel (McCarthy and Prince 1994), [ə] is the default vowel in Palauan.

8. Since affix reduplicants involve more marked segments than some forms of the root reduplicants, Urbanczyk (1999) mentions that Palauan is a case of morphological overwriting (Alderete et al. 1999). The input of the reduplicated form [*məʔɛʔúʔəp*] is thus considered /mə+RED$_{affix}$+ ɛ+ ʔuʔəp/, and only the first consonant of the base is copied. While the emergence of [ɛ] is predicted, assuming /ɛ-/ in the input considerably weakens the claim that we can dispense with templates for reduplicants (McCarthy and Prince 1994). So, rather than assuming /ɛ-/ in the input, I propose an analysis in which realization of [ɛ]is one instance of TETU in the text.

9. One of the reviewers pointed out that [ɛ] appears in the TETU condition in the possessive form as well. The possessive suffix is /-k/ and the syllable which has this suffix always carries stress. When the noun stem ends with a consonant, the vowel [ɛ] is inserted.

(i) **underlying form**   **possessive form**

underlying form		possessive form	
/buŋa/	'flower'	[puŋák]	'my flower'
/ŋalək/	'child'	[ŋaləkɛ́k]	'my child'

Usually the inserted vowel is the least marked vowel. In Palauan, [ə] is the least marked vowel, but it cannot bear stress. So, the second least marked vowel [ɛ] appears. Because the stress shifts to the last syllable in the possessive form, the vowel reduction takes place in the possessive form in the same way as with the CVCV-reduplication (Wilson 1972, Finer 1986, and Zuraw 2002).

(ii) a. **simple form**   **possessed form** (Josephs 1975, Finer 1986)

	simple form		possessed form	(Josephs 1975, Finer 1986)
	ŋór	'mouth'	ŋərɛ́k	'my mouth'
	bsíbs	'drill'	bsəbsɛ́k	'my drill'
	sɛ́rs	'garden'	sərsɛ́k	'my garden'
b.	sáik	'laziness'	sikɛ́k	'my laziness'
	klakoád	'fight'	klakodɛ́k	'my fight'
	bɛróus	'dream'	bɛrusɛ́k	'my dream'

The possessive suffix has stress, and the stressed vowel in the root is reduced. When the CVCV-reduplication and the possessives are compared, we notice that vowel reduction takes place in the unstressed foot. To account for these two cases uniformly, MAX IO and IDENT IO should be considered as faithfulness constraints only for <u>the stressed foot</u> (MAX IO IN STRESSED FOOT, IDENT IO IN STRESSED FOOT), and general MAX IO and IDENT IO must be ranked lower than other

constraints. These constraints do not affect the elements in the base of reduplicated words, since the base itself is a stressed foot. The following is the tableau for the possessive form in (iia).

(iii) /ŋoɾ/+/k/	MAX/IDENT IO IN STRESSED FOOT	ROOT= FOOT	FT-BN (μ)	ONSET	ROOT MAX BR
☞ a. (ŋəɾ)(έk)			*	*	
b. (ŋoɾ)(έk)			**!	*	
c. (ŋəɾ)(ék)	*!			*	

The candidate in (iiib) is ruled out, since both feet have three moras each and it violates FT-BN (μ) twice. (iiic) is ruled out because the second foot (ék) has stress and it is not faithful to the input. In the winning candidate (iiia), the vowel in the root is reduced. This change is allowed, since the foot for the base is not stressed and MAX/IDENT IO IN STRESSED FOOT is satisfied.

In (iib), one vowel is deleted and another vowel remains without being reduced to [ə]. To explain this case, I have to assume the conjoined constraint [Max IO + Ident IO] (Smolensky 1993). This constraint is violated only if both Max IO and Ident IO are violated. It does not affect the tableau in (iii) because the correct form does not violate it.

(iv)/saik/+/έk/	MAX/IDENT IO IN STRESSED FOOT	ROOT= FOOT	[MAX IO + IDENT IO]	FT-BN (μ)	ONSET
☞ a. (sik)(έk)				**	*
b. (saik)(έk)				**	**!
c. (sək)(έk)			*!	*	*

(iv b) has two onsetless syllables, namely [ik] and [έk], violating ONSET twice. In the candidate (iv c), one of the vowels in the sequence is deleted, violating Max IO, and another vowel is reduced to [ə], violating Ident IO. Since this candidate violates both Max IO and Ident IO, it violates the conjoined constraint [MAX IO + IDENT IO]. Given this analysis, vowel reduction takes place to satisfy FT-BN (μ) in both possessive forms and reduplications.

References

ALDERETE, JOHN D., JILL BECKMAN N., LAURA BENUA, AMALIA GNANADESIKAN E., JOHN MCCARTHY J., AND SUZANNE URBANCZYK. 1999. Reduplication with fixed segmentism. Linguistic Inquiry 30. 3, 327-364.
BECKMAN, JILL N. 1997. Positional faithfulness. Amherst: University of Massachusetts Amherst dissertation.

CHO, YOUNG-MEE Y. 1999. Optimality in Korean reduplication. Harvard Studies in Korean Linguistics 8, 73-87.

FINER, DANIEL L. 1986. Reduplication and verbal morphology in Palauan. The Linguistics Review 6, 99-130.

JOSEPHS, LEWIS S. 1975. Palauan reference grammar. Honolulu: University of Hawaii Press.

JOSEPHS, LEWIS S. 1990. New Palauan-English dictionary. Honolulu: University of Hawaii Press.

KAGER, RENÉ. 1999. Optimality theory. Cambridge: Cambridge University Press.

MCCARTHY, JOHN J., AND ALAN PRINCE. 1994. The emergence of the unmarked: optimality in prosodic morphology. Proceedings of the North East Linguistic Society (NELS) 24. 333-379.

MCMANUS, EDWIN S. J. 1977. Palauan-English dictionary. Honolulu: University of Hawaii Press.

PRINCE, ALAN. AND PAUL SMOLENSKY. 1993. Optimality theory: Constraint interaction in generative grammar. Technical Report #2 of the Rutgers Center for Cognitive Science, Rutgers: Rutgers University.

STERIADE, DONCA. 1995. Underspecification and markedness. The handbook of phonological theory, ed. by Goldsmith, John, 114-174. Oxford: Blackwell.

SMOLENSKY, PAUL. 1993. Harmony, markedness, and phonological activity. [ROA-87, http://ruccs.rutgers.edu/roa.html]

URBANCZYK, SUZANNE. 1999. Reduplicative size-segmentism correlations as root-affix asymmetries. ms. New Bruswick: Rutgers University,.

WILSON, HELEN. 1972. The phonology and syntax of Palauan verb affixes. Honolulu: University of Hawaii dissertation.

ZURAW, KIE. 2002. Vowel reduction in Palauan reduplications. Proceedings of AFLA 8.

Tomoko Kawamura
SUNY at Stony Brook
Department of Linguistics, SUNY at Stony Brook, NY 11794-4376,
U.S.A.
tkawamur@ic.sunysb.edu

Malagasy Clause Structure and Language Acquisition

Edward L. Keenan,   Cecile Manorohanta
UCLA   Université Nord, Madagascar

abstract We argue that Malagasy (and related W. Austronesian languages[1]) has a positive setting for a macro-parameter RICH VOICE MORPHOLOGY which builds complex predicates that code the theta role of their argument: $S = [[_{Pred}V(\theta) + (X)] + DP]$. Manifestations of this parameter are: (1) Case and theta role are assigned in situ in nuclear clauses with no movement or co-indexing to a topic position. (2) Relative Clauses (and other "extraction" structures) satisfy the "Subjects Only" constraint, again with no movement or indexing. (3) UTAH is freely violated, as theta role assignment derives from compositional semantic interpretation. Predicates resemble lexical Ns in assigning case directly to arguments without using Prepositions and in combining directly with Dets to form DPs that include tense and negation (Keenan 1995, 2000). The major Predicate-Argument type is modeled on the Noun+Possessor one, not the Verb+Object one.

1. background Lexical verbs denote relations requiring that their arguments bear thematic (theta) relations such as *Agent, Theme, Goal, Expierencer* and *Means* to them. For example in *John chased Bill, Bill* bears the Theme relation to *chase*, meaning that the object it denotes[2] is affected by the chasing action. UTAH (Baker 1988:46) syntactically constrains the expression of theta roles:

UTAH Identical thematic relationships between items are represented by identical structural relationships between those items at the level of D-structure.

We assume UTAH is updated, replacing "at the level of D-structure" by "in the derivationally prior structures at which theta roles have been assigned".

Formally UTAH seems arbitrary (if historically comprehensible). Would deriving constituents bearing the same theta role from non-isomorphic structures lead to a contradiction? Would we wrongly predict some theta equivalences? Deriving theta equivalence compositionally yields a negative answer to both questions. As a conceptual illustration first consider (1a,b) from Propositional Logic.

(1)   a.   (P or Q)   b.   not (not P and not Q)

(1a,b) are not syntactically isomorphic but are, like infinitely many other pairs, semantically interpreted the same in all models. This logical equivalence just depends on their compositional interpretation plus the same choice of lexical items, 'P' and 'Q'.

Similarly in Malagasy nuclear Ss built from verbs with different voice morphology may be theta equivalent even though constituents with the same theta role may never have been structurally identical. So UTAH fails here, which is unsurprising: theta role assignment can be a function of structure: "same structure, same theta role", without being a one to one function: "different structures, different theta roles".

2. Theta Relations in Malagasy Here is a simple pair of theta equivalent Ss built from the same root, *enjika*. Verbs are given in morphemic decomposition followed by the derived form in standard orthography[3]. Square brackets mark the major constituent break (here, and later).

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

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

Voice affixes are not glossed (except *-ina* as *-Vna* since the vowel varies). There is massive evidence that the bracketed string in each of (2a,b) is a constituent (Keenan 1976, 1995; Guilfoyle, Hung and Travis (GHT 1992); Paul (1999), Pearson (2001, 2003). Theory neutrally we call it a Predicate Phrase (PredP). The final DP in these Ss will be called the External Argument (EA), following Pearson (2001). Here are two types of evidence that support this constituency judgment.

First, various particles, such as the yes-no question particle *ve*, are placed at the right edge of the PredP, separating it from the EA.

(3) a.   nanenjika ny jiolahy ve Rabe?   b.   nenjehin-dRabe ve ny jiolahy?
         Did Rabe chase the thief?   *Did Rabe chase the thief?*

No other locus of *ve* is grammatical. Similarly the 'even...though' construction places *na (dia)* at the left edge of the PredP and *aza* at its right. Also negation, *tsy*, occurs at the left edge and npi's like *velively* 'at all' and *intsony* 'longer' at the right edge.

Second, Keenan (1995) shows that PredPs occur in diverse syntactic contexts with the same form and meaning: relative clauses (RCs), embedded questions, Raising to Object contexts, and tensed PredP nominalizations. We illustrate with RCs:

(4) **Relative Clause Formation: N = N + PredP**

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

PredPs denote properties, functions from individuals to truth values. The property denoted by *nanenjika ny jiolahy* maps an individual x to True if x chased the thief, and to False otherwise. That denoted by *nenjehin-dRabe* maps x to True if x was chased by Rabe, and to False otherwise. In RCs PredPs are property denoting expressions. In (4a) the children referred to are those that *nanenjika ny jiolahy* 'chased the thief' maps to True. In (4b) it is those *nenjehin-dRabe* 'was chased by Rabe' maps to True. RCs are interpreted as in (5a). (5b) is an instance of (5a).

(5) a. $[\![\text{Noun} + \text{PredP}]\!]$ = $[\![\text{Noun}]\!] \cap \{x | [\![\text{PredP}]\!](x) = \text{True}\}$

    b. $[\![\text{zaza nenjehin-dRabe}]\!]$ = $[\![\text{zaza}]\!] \cap \{x | [\![\text{nenjehin-dRabe}]\!](x) = \text{True}\}$

                        = $\text{CHILD} \cap \{x | (\text{CHASED BY RABE})(x) = \text{True}\}$

Thus we compositionally interpret (4a,b)[4]. We note that demonstrative Dets and the definite article *ny* combine directly with PredPs yielding DPs: *ny nenjehin-dRabe* = "the (ones) chased by Rabe".

Beginning with Keenan (1972) linguists have ethnocentrically modeled Malagasy RC formation and interpretation on the English pattern, yielding structures like (6) in which the post-N expression contains a full S and its PredP has an external argument (EA) that moves preclausally and is co-indexed with a trace in its original position:

(6) zaza [wh_i [$_S$[$_{PredP}$ nanenjika ny jiolahy] [t_i]]].
     child who   chased   the thief

(6) can also be compositionally interpreted but carries useless structure: an embedded S and an invisible EA. But this is objectionable in Malagasy. To build and interpret RCs as per (4) and (5) the Malagasy child just uses Ns and PredPs already learned in interpreting simple Ss. What motivates the child to detect and interpret unpronounced categories? (We treat below a question of *expressive completeness* raised by this way of forming RCs).

The English bias leads us to ask why *wh* does not occur as an internal argument, moving to Pred initial position yielding the ungrammatical (7) (extraction sites marked by case marked traces)? In informal classical terms, Why do only subjects relativize? The answer we give is that the post-N phrases do not occur as independent PredPs.

(7) a. *zaza nenjehina+t_{gen} ny jiolahy   b. *zaza nanenjika t_{acc} Rabe
       child (that) the thief was chased by   child (that) Rabe chased

A more insightful perspective results from reversing the ethnocentric bias: Why does English not form RCs in the learning theoretically pleasing way that Malagasy does? A partial answer is given by comparing the subject and object RCs:

(8) a. (the) child who chased the thief   b. (the) child who Rabe chased

In (8a) the post-*who* phrase is a PredP interpretable as in Malagasy. But in (8b) *Rabe chased* is not independently a PredP. In *Rabe chased the child* it is not a constituent and not assigned a meaning. The fact that RCs (wh-questions, clefts, etc.) in English force us to interpret strings which are not independently meaningful provides an explanation for why English speakers have recourse to variable binding operators and empty categories: they represent minimal logical extensions of what actually occurs that yields a correct compositional semantic interpretation.

But one can object here that the meaning of (8b) is expressible in a Malagasy way just by using passive PredPs, as in *(the) child who was chased by Rabe*. And this is correct. But as the range of syntactic contexts is increased we see that Malagasy presents a much more extensive repertoire of passive-like PredPs supporting the expressive completeness of this strategy in Malagasy. Here are six types of PredPs fully natural in Malagasy but marginal or ungrammatical as passives in English.

[1] PredPs with semantically oblique EAs. The PredP in (9a) admits of modification with many obliques, such as Benefactives and Instrumentals. The verb in (9b,c), formed by suffixing the verb in (9a) with *-ana* (see footnote 11), takes EAs corresponding to these obliques; the PredPs in (9b,c) are identical.

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

    b.   [ø+anaovan-dRabe trano] Rasoa
           pres+aN+do+ana+Rabe.gen house Rasoa
           Rabe is making a house for/because of Rasoa

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

    d.   ny vehivavy [anaovan-dRabe trano]
           the woman makes-Rabe house
           the woman Rabe is making a house for / because of
           *?the woman who is being made by Rabe a house for/because of

In (9d) the expression following *vehivavy* 'woman' independently denotes a property, as in (9b,c), but not so for what follows 'woman' in the English translation of (9d).

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

(10) a.   manantena aho fa nanaovan-dRabe trano Rasoa
           pres+aN+hope 1.sg.nom that past+aN+do+ana+Rabe.gen house Rasoa
           I hope that Rabe made a house for Rasoa[5]

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

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

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

Complex predicates like those in (10c,d) are the major mechanism relativizing (questioning, etc.) arguments of complement clause verbs.

[3] Control context predicates whose EAs bear theta roles to the denotation of the controlled predicate, not the controlling one:

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

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

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

[4] Object control contexts yield patterns similar to the RTO cases:

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

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

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

[5] Verbs of motion induce PredPs similar to those in control contexts:

(13)a. [niakatra hianatra teny vahiny tany Antananarivo] i Soa
past+i+go+up fut+i+study lg foreign past+loc Antananarivo art Soa
Soa went up to study foreign languages in Antananarivo

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

[6] PredPs with internal Agents binding anaphors:

(14)a. [m+aN+vono (mamono) tena ho an'ny ankizy] ny ray aman-dreny
pres+aN+kill self for'the children the father and-mother
Parents kill themselves for their children

b. [ø+amono+an'ny ray aman-dreny tena] ny ankizy
pres+aN+kill+ana'the parents.gen self the children
Parents kill themselves for their children

c. ny ankizy izay [amonoan'ny ray aman-dreny tena]
the children that aN+kill+ana'the parents self
the children that parents kill themselves for
**the children that parents are killed by themselves for*

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

b. [n+if+anampy hitsara ny fanadinana] isika
past+recip+help judge the exams we.incl
We helped each other grade the exams

c. [nifanampian-tsika hitsarana] ny fanadinana
past+recip+help+ana-1.gen.pl.incl fut+i+judge+ana the exams
We helped each other grade the exams

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

Pairs like (2a,b) do resemble English active-passive pairs in one respect: they are theta equivalent, with EA Agent in one case and Theme in the other. So we sometimes translate non-AF forms as passives. But the analogy is limited. As we will see, Ss with Theme EAs are more like actives than passives.

Returning now to (2), the different verb forms take their internal argument DPs in different cases. Malagasy has a three case system labeled traditionally below. Here are the pronominal distinctions. See K&P (Keenan & Polinsky 1998) for details.

Singular:	1	2	3	Plural: 1excl	1incl	2	3
nom	aho	ianao	izy	izahay	isika	ianareo	izy[6]
acc	ahy	anao	azy	anay	antsika	anareo	azy
gen	-ko	-nao	-ny	-nay	-tsika	-nareo	-ny

nominative is the case of the EA and functions as a default (Pearson 2003). It is only distinctively marked in the pronouns. It is never selected by Vs, Ns or As in forming PredPs, NPs or APs. Only two Preps select **nom** (*noho* 'than, in comparatives)', and *afa-tsy* 'except, lit: "free-not").

accusative has distinctive pronouns and marks proper nouns, some kin terms and, optionally, demonstratives with *an-*. **acc** is selected by many Vs, some Ns (below), a few Ps (*tahaka azy* 'like him') and a few As (*feno azy* 'full (of) it'). DP_{acc}s also serve as predicate possessives: *Ahy io* "That is mine".

genitive is the case of nominal Possessors, non-external Agents of Vs and As (Phillips 2000) and most objects of Prepositions. DP_{gen}s, both pronominal and full, are morphologically bound to their heads (the latter by a process we call *n-bonding*). Except monosyllabic *-ko* and *-ny* pronouns may carry stress (phonemic). They may replace full DPs, so pronominal replacement is a test for the case of a DP.

In a text count based on two newspaper articles and selections from three novels, in Malagasy and in English, Keenan (1995) found case distributed as in (16)[7].

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

Genitive is the most widely occurring case in Malagasy, compared with accusative in English, nominative falling in second place in the two languages.

Gen The dominant expression of the Predicate-Argument relation in Malagasy is the Head+Possessor one. In English it is the Verb+Object one.

Verb voices and Nuclear Clauses in Malagasy Primary voice affixes combine with roots to build verbs with different subcategorizations which build Ss with different phrase structures. We present a compositional semantic interpretation which yields as a corollary the correct judgments of theta equivalence. We repeat (2a,b) below.

(17) a. n+aN+enjika (nanenjika) ny jiolahy Rabe
        past+aN+chase   the thief   Rabe
        Rabe chased the thief

     b. n+enjika+ina+Rabe (nenjehin-dRabe) ny jiolahy
        past+chase+Vna+Rabe.gen   the thief
        Rabe chased the thief

In (17a,b) both verbs are marked past with *n-/no-* which is constant across voices.

We often ignore tense marking (obligatory) as irrelevant to our concerns.

In (17a) *aN*- combines with the transitive root *enjika* to form a transitive verb[8], one which combines first with an accusative DP and then a nominative DP. We derive *aN*- verbs by a function AN mapping roots to verbs, illustrated in (18)[9]. We list leftmost in the subcategorization of the verb the category it combines with first (contra the convention in Keenan 1995). The resulting expression combines with a DP_{nom} forming an S. So the rightmost category is that of the EA of the verb.

$$\text{(18)} \quad \text{enjika: } RT_2 \quad \xrightarrow{\text{AN}} \quad \text{-anenjika: } S/[DP_{acc}, DP_{nom}]$$

Syntax: *enjika* has category RT_2, a root with two theta roles. AN is a function mapping such expressions to predicates of category $S/[DP_{acc}, DP_{nom}]$, which Merge combines with the appropriate case marked DPs. The derivation of (17a) is (19).

(19)
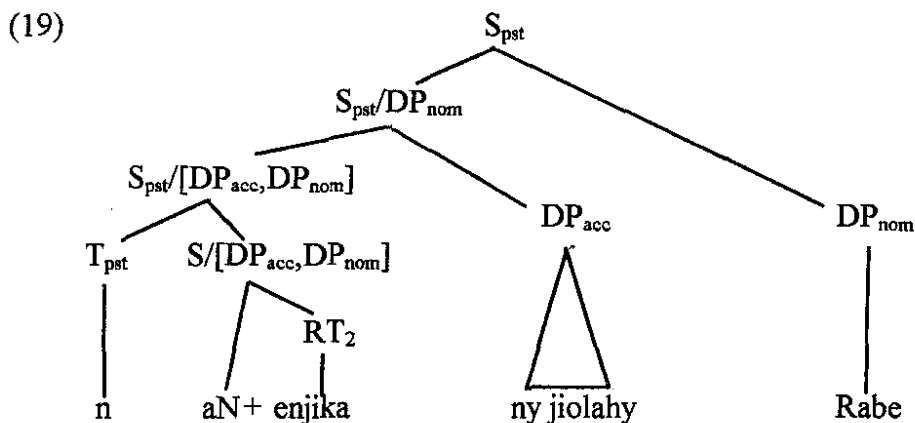
The use of 'S/...' in our category of verbs is not redundant. Nouns like 'fear' and 'compassion' subcategorize two case marked DP arguments to form a DP.

(20)

Razay's compassion for her father
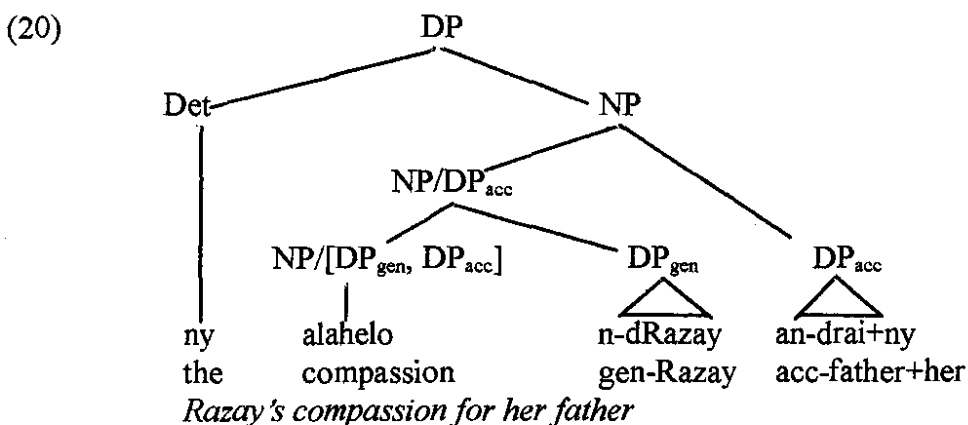
Semantics We interpret the root as a binary relation, ENJIKA, with theta roles Theme and Agent. The interpretation of the verb AN(enjika) is given by (21), which says that the pair (x,y) is in the ENJIKA (CHASE) relation, y is its THEME and x its AGENT.

(21)⟦AN(enjika)⟧(y)(x) = True iff

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

Of course we want to give the syntax and semantics of *aN-* more generally. What roots does it combine with? Does it always build verbs whose EA is Agent? In fact the verbs *aN-* builds are usually (di)transitive. Examples are: *didy* ==> *mandidy* 'cuts', *vaky* ==> *mamaky* 'reads', *vono* ==> *mamono* 'hits, kills'. Examples of ditransitives are: *seho* ==> *maneho* 'shows', *tolotra* ==> *manolotra* 'offers', *ome* ==> *manome* 'gives'. AN however builds two types of intransitives. First a very few activity ones: *leha* ==> *mandeha* 'goes', *dihy* ==> *mandihy* 'dances', and *lohalika* ==> *mandohalika* 'kneels'. Finally, there is a non-trivial class of verbs it builds which do not present an Agent: *mangetaheta* 'is thirsty' (< *hetaheta*), *mangatsiaka* 'is cold' (< *hatsiaka*) and *mangovitra* 'shivers' (< *hovitra*) and *firatra* ==> *mamirapiratra* scintillates'. Relative to the Theta Hierarchy below, (22) covers all the cases.

Theta Hierarchy Agent > Experiencer > Goal /Theme/Patient > Other

Def A verb whose EA is assigned the highest theta role in the verb's subcategorization is called *Actor Focus* (AF). Affixes which build AF verbs and nuclear Ss built from AF verbs are also called Actor Focus.

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

Language internal generalizations like (22) aid in acquisition by providing default interpretations for novel verbs.

Another AF voice prefix is *i-*. It builds many transitive verbs: *kapoka* ==> *mikapoka* 'beats', *vidy* ==> *mividy* 'buys', *orina* ==> *miorina* 'builds', *varotra* ==> *mivarotra* 'sells', *laza* ==> *milaza* 'says', *tady* ==> *mitady* 'seeks', but most often *i-* builds intransitive verbs, and that in two distinct ways. First, it combines with roots taking just one theta role: *asa* ==> *miasa* 'works', *tsangana* ==> *mitsangana* 'stands up', *tomany* ==> *mitomany* 'cries'. Second, it combines with transitive roots that also accept the *aN-* prefix, usually omitting the Theme theta role: *enjika* ==> *manenjika* 'chases', *mienjika* 'flees'; *sasa* ==> *manasa* 'washes (tr)', *misasa* 'washes (intr), *hidy* ==> *manidy* 'closes (tr)', *mihidy* 'is locked'. So the following generalization holds:

(23) *i-* is an Actor Focus affix; when *i-* and *aN-* combine with the same root the *i-* verb usually has lesser valence, lacking a Theme present in the *aN-* verb.

Consider now the verb *enjehina* in (17b), formed by suffixing the root *enjika* with *-ina*. It combines first with a DP_{gen}, interpreted as Agent, and then with a DP_{nom}, interpreted as Theme. (24) shows the syntactic action of this suffix, (25) gives its semantic interpretation, and (26) derives a nuclear S built from the INA-verb.

186

$$\text{INA}$$

(24) enjika; RT_2 ==> enjehina: $S/[DP_{gen}, DP_{nom}]$

(25) $[\![INA(enjika)]\!](u)(v)$ = True iff

$$(u,v) \in \text{ENJIKA} \land \text{THEME}(v,\text{ENJIKA}) \land \text{AGENT}(u,\text{ENJIKA})$$

(26)
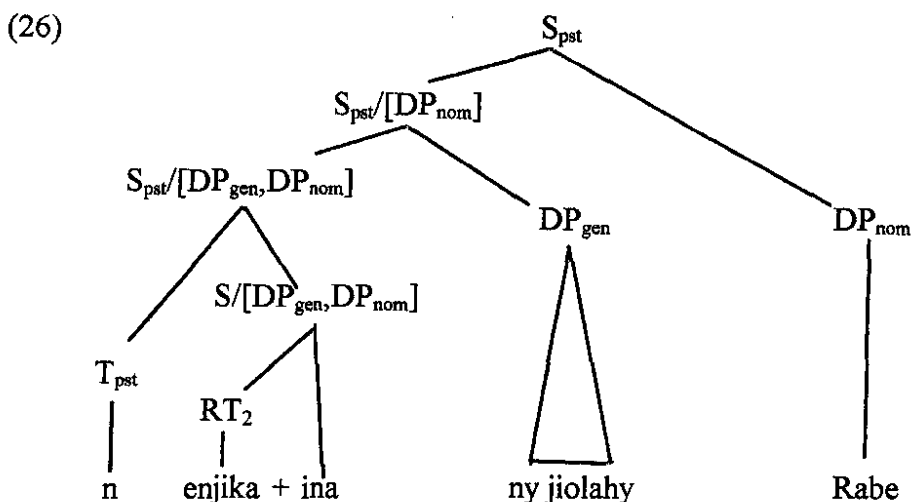
Let us call a verb τ Focus if its external argument is assigned the role τ^{10}. So *enjehina* is a Theme Focus (TF) verb. Typically *-Vna* builds TF verbs. Notice now that (27a) follows immediately from (21) and (25), and in general (27b) holds when ROOT is in the domain of both AN and INA.

(27) a. $[\![AN(enjika)]\!](y)(x)$ = $[\![INA(enjika)]\!](x)(y)$ and

b. $[\![AN(ROOT)]\!](y)(x)$ = $[\![INA(ROOT)]\!](x)(y)$ and

(28) **Theta Equivalence** Minimal Ss built from AN(ROOT) and INA(ROOT) are theta equivalent: the EA of each bears the same theta role as the internal argument of the other. But the Agent DPs in the two Ss were never in structurally comparable positions; nor were the Themes. Thus UTAH does not hold on this analysis.

Another primary TF affix is the prefix *a-*.

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

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

187

tosika does not accept a *-Vna* suffix. However ditransitive roots with both Goal and Theme roles, (30), often have both *a-* and *-Vna* forms, in which case the *a-* verb assigns Theme to the EA and *-Vna* assigns it Goal:

(30) a. n+aN+roso (nandroso) vary ny vahiny Rasoa
    past+aN+serve   rice the guest Rasoa
    Rasoa served rice to the guests

   b. no+roso+ana+Rasoa (norosoan-dRasoa)vary ny vahiny
    past+serve+Vna+Rasoa.gen   rice the guests
    The guests were served rice by Rasoa

   c. n+a+roso+Rasoa (naroson-dRasoa) ny vahiny ny vary
    past+a+serve+Rasoa.gen the guests the rice
    The rice was served to the guests by Rasoa

(31) With ditransitive roots *a-* externalizes the "Intermediary": the Theme with roots of transmission, the Instrument otherwise. In the first case *-Vna* externalizes the Goal, and in the second the Theme. See Paul (1999).

Secondary verbal affixes build (tenseless) verbs from (tenseless) verbs, combining thus with items which already have (at least) one verbal affix. Of the four major such affixes, *if-* 'Reciprocal' and *amp-* 'Causative' are semantically specific and not our primary concern here (see K&R 2004).

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

(33) a. m+aN+sasa (manasa) lamba Rabe
    pres+aN+wash clothes Rabe
    Rabe is washing clothes

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

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

(33c) shows that the primary suffix *-ina* also combines with causative verbs to form TF verbs. No other shape of *-Vna* will work here, nor do causatives accept an *a-* TF prefix. But There remains a central voice affix which is crucial to the expressive power constraints alluded to earlier. It is the suffix *-ana* in (9b,c,d). Compare:

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

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

    c. [n+aN+enjika+ana+Rabe (nanenjehan-dRabe) ny jiolahy] ny fiara
  past+aN+chase+ana+Rabe.gen   the thief   the car
  The car was used by Rabe to chase the thief

The EA in (34c) is the Means constituent is nominative, with no preposition. The verb in (34c) will be called CF, *circumstantial focus*, since its EA expresses some circumstance of the action or state denoted by the verb.

Syntactically a CF verb is built from an AF verb by suffixing *-ana*, which triggers stress shift to the right[11]. The resulting verb is not AF, and its first argument is bound to the verb as a genitive, just as with TF verbs, as in (17b). We note that all tense taking[12] AF verbs accept *-ana* suffixation yielding non-AF verbs. For example, from *enjika* 'chase' we may form (present tense, primary stress marked):

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

ana generally adds an argument for which its AF verb is not subcategorized (Paul 1999, Keenan 2000). All we know about it is that it will be interpreted as one expressible by an oblique constituent of the AF verb. *-ana* verbs are built by a function ANA, illustrated applying to a transitive verb in (34)[13].

$$\text{ANA}$$
(36)   (verb$_{AF}$: S/[DP$_{acc}$, DP$_{nom}$])   ==>   (verb+ana; S/[DP$_{gen}$, DP$_{acc}$, DP$_{nom}$])

Semantically oblique DPs and adverbials are predicate modifiers – functions F mapping predicate denotations P to predicate denotations F(P). Normally they are *restricting*: F(P) $\leq$ P, meaning that for all entities x, F(P)(x) = True then P(x) = True[14]. *for Bill* is restricting in (37a), whence (37a) entails (37b). Replacing *for Bill* by *with a screwdriver* preserves the entailment since it is also restricting.

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

    b.   Sue opened the box   **(open(the box))(Sue)**

So Prepositions are semantically functions **p** mapping entities to restricting modifiers. And the semantics for CF verbs quantifies over such functions, illustrated in (38) with the circumstantial of a transitive verb.

(38) $\llbracket ANA(verb_{AF}) \rrbracket(x)(y)(z)$ = True iff $\exists p\ p(z)(\llbracket verb_{AF} \rrbracket(y))(x)$ = True

The crucial point is that $p(x)$ is restricting, so a nuclear circumstantial S like (39a) entails the corresponding AF one (39b) which lacks an oblique DP corresponding to the EA of the circumstantial. (39a) is derived in (40) and interpreted in (41).

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

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

(40)

(41)
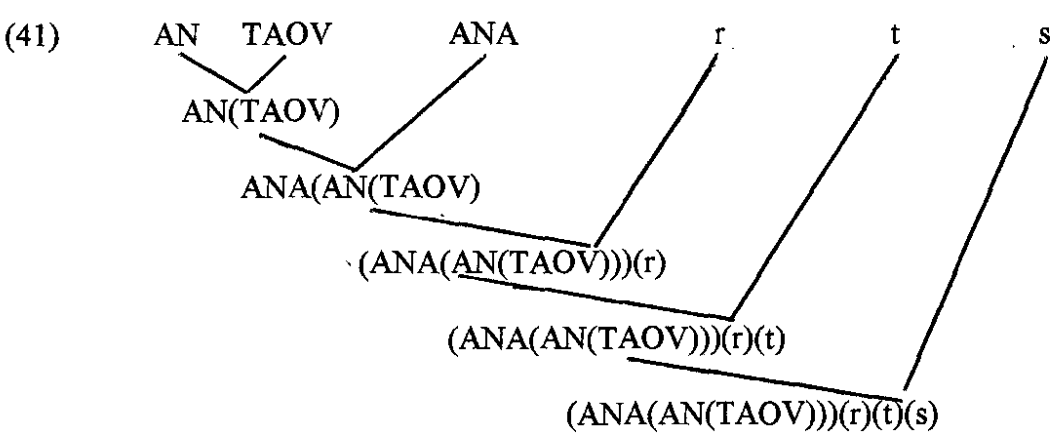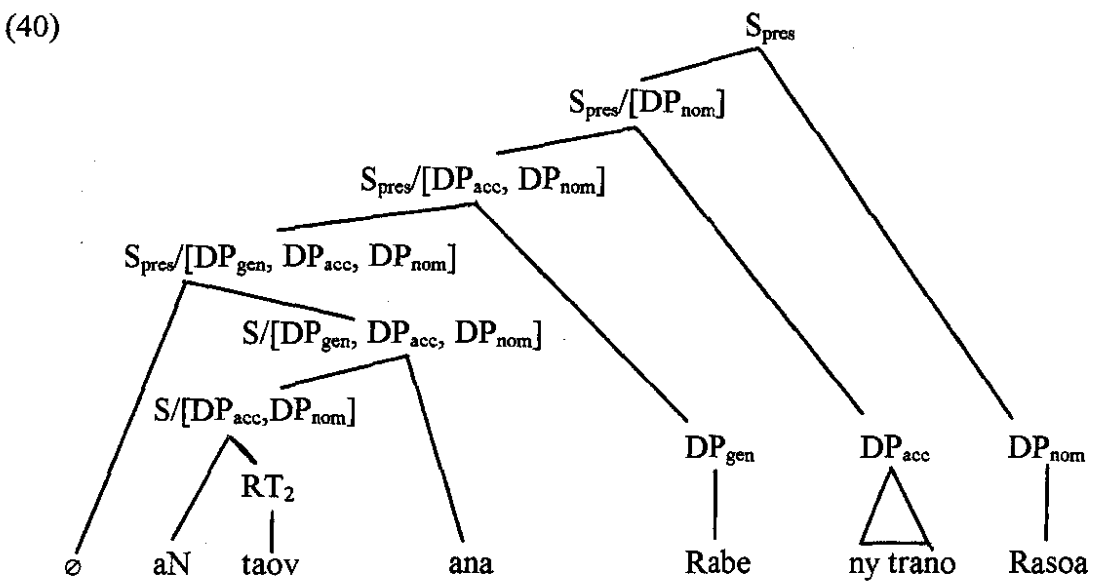
The last line is True iff for some **p**, **p**(s)AN(TAOV)(t))(r) = True. And since **p**(s) is restricting the last line entails AN(TAOV)(t)(r), the interpretation of *Rabe is making the house*. Thus our semantics for the circumstantial S entails that Rabe is Agent of *make* and the house is Theme. Consider now the predictions of theta equivalence between a circumstantial S and a corresponding AF one with an overt oblique.

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

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

Our semantic analysis shows that the genitive complement, Rabe, in (42a), has the same theta role, Agent, as the EA in (42b). Also the accusative *ny trano* 'the house' has the same theta role, Theme, in both Ss. But while *Rasoa* is a Benefactive in (42b), forced by the choice of Prep, *ho an-* 'for', *Rasoa* is not assigned a theta role on our analysis in (42a). Rather our semantics merely quantifies over the set in which Preps denote, different theta roles for Rasoa correspond to different choices of possible Prep denotations. Most choices – Instrument, Location, Manner – are pragmatically implausible in (42a). But *Rasoa* could still be a Benefactee or a Cause.

Thus on our analysis the EA of CF Ss is vague, not ambiguous, with respect to theta role. We just know that it stands in some relation to the predicate which is of the sort expressible by a Preposition. On standard views in contrast the EA (or its co-indexed operator) is derived by movement from structures in which it has different theta roles, whence the CF S is semantically ambiguous. The point is subtle, but indicative of a deeper difference in approach. Namely, the only regular semantic relations between Ss differing just by verbs in different voices are those predictable from the verb. This includes theta equivalence and logical equivalence for Ss with individual denoting DPs, as the semantics of verbs is given in terms of individuals. But the behavior of adverbials and quantifier scope is not predictable.

3. Support for the K&M analysis We review first, briefly, some "standard" approaches to Malagasy clause structure – GHT (1992), Paul (1999) and Pearson (2001, 2003). (43) is the derivation of a nuclear TF S in GHT. Here *-Vna* occurs in INFL and by stipulation assigns case to the Agent in SPEC VP. The Theme complement to V lacks case and so moves to SPEC IP, where case is assigned (by stipulation). V moves to INFL to support the *-Vna* morphology. Then the Agent attaches to it in the same way that possessors n-bond to their heads.

191

(43)

```
                        IP
          I'_____SPEC
        /    \                     Theme
     INFL    VP                      /\
      |     /  \                    /  \
           SPEC`   V'              /    \
           Agent   /\             /      \
                  /  \           /        \
                 V    NP        /          \
                 |    |        /_____\
     -Vna  Rabe  enjika  tᵢ   ny jiolahyᵢ
```

(*Enjehin-dRabe ny jiolahy*, ≈ (17b))
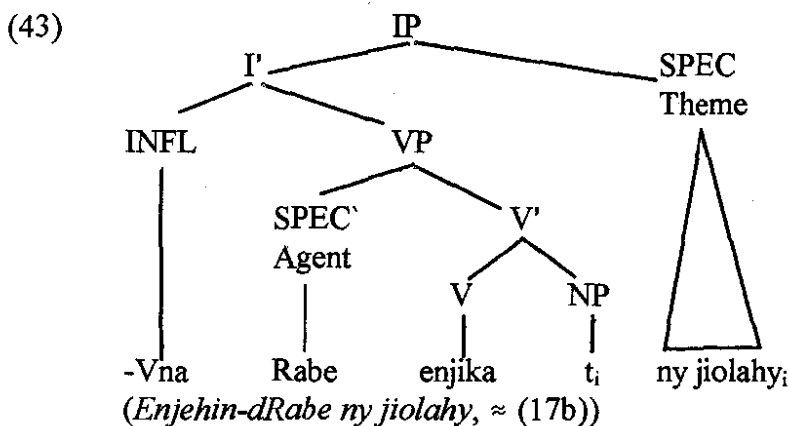
The derivation of a nuclear AF S differs from (43) in that *aN-* occurs in V and assigns case to its Theme. Nothing assigns case to the Agent in SPEC VP, so it moves to SPEC IP where it gets case. Later V moves to INFL string vacuously. (Both imperative and present tense indicative AF forms present an AF-specific *m-* in INFL, as Pearson notes. Hosting this *m-* would motivate the verb movement, but why does *m-* not assign case to the Agent in SPEC VP as *-Vna* in (43) does?)

Paul (1999)'s underlying structures are comparable to GHT's but richer. The *a*-prefix forms are explicitly treated, and *a-* is assigned a structural position different from *aN-* and *-Vna*. Movement of DPs to the EA position is motivated by case assignment as in GHT. Pearson's approach shares with Paul and GHT the assumption that case is determined in different structural positions associated with the voice morphology. But for Pearson the EA is base generated in SpecTP (Topic Phrase) and by stipulation is co-indexed with an empty wh-operator in the next lower projection, SpecWhP. That position, by stipulation, licenses wh-operators, so the empty operator is moved there from a case position. So Pearson's movement is A' movement, to a non-case position, not A movement as in Paul and GHT. We turn now to three learning theoretic merits of our, K&M, analysis. For more narrowly grammar internal advantages see Keenan (1995).

1. Our analysis correctly predicts several basic judgments of logical equivalence – (17a,b), (30a,b,c), and the judgment of entailment in (42a,b). Standard approaches lack an explicit semantic interpretation and thus cannot justify such predictions. Indeed the complexity of the structures assigned to simple Ss and the movement operations needed to derive the audible expressions present a non-trivial challenge to a compositional semantic analysis. Compositionality is basic from a learning theoretic perspective as it assures us that speakers can interpret complex expressions if they know what the component lexical items mean and how things built in that way take their meaning as a function of the meanings of their constituents. Compositionality is to date our only means of accounting for how speakers comprehend novel utterances.

2. On our view the voice affixes – *aN-, i-, ina, a-, -ana*, etc. are lexical entries with a semantic interpretation, different affixes applying to a given root yielding different meanings. So we are explicit regarding what children must learn to master

their use and meaning. And given their interpretation, and hence that of the verbs they build, it is explicit what theta role each argument has. This is not a primary concern of standard approaches. It is also clear on our view why the generalizations in (22), (23), and (31) provide useful default reasoning guides to learning novel verbs built from the affixes considered. So our analysis better explains how Malagasy learn their language quickly with limited exposure to imperfect data....

In contrast on standard approaches the voice affixes are not meaningful lexical items, but rather reflexes of case assignment from different structural positions. This means that the morphemic identity of the case assigners is redundant. They could all be null and we could still claim that case X is assigned in configuration Y. On our approach the voice morphemes are not redundant – make them null and we can't distinguish and hence interpret differently verbs in different voices.

Note that we can not associate with each primary voice affix a specific case, given that they all form verbs of differing valencies. E.g. if *aN-* always assigns accusative, how do we generate the intransitive verbs built from it? *I-* is even more problematic. Nor can we bi-uniquely associate voice affixes with theta roles, a sufficient reason being that basic (di)transitive verbs have more theta roles than affixes. On our approach case and theta selection are determined by the affix plus the root. A given affix may map different roots to verbs of different valencies, hence of different case and theta grids, and different affixes may map a given root to verbs with different case and theta grids.

3. Standard approaches and K&M differ regarding the syntactic complexity of primary AF and TF Ss. On K&M the *derivations* of (17a,b) are isomorphic, differing just by the choice of voice affix and the case of the internal argument[15]. But on standard approaches the AF S is simpler than the TF one, since the movement involved is shorter. In an AF S on the GHT/Paul view we only move the DP in SpecVP to SpecIP, "the next Spec up". But in TF Ss the complement of the V must move over SpecVP to get to SpecIP. A comparable distance gap arises with the movement of the empty operator on Pearson's analysis[16]. And in *a-* and *-Vna* forms of ditransitive Vs the EA (or empty operator) must move across two theta positions.

The symmetry between primary AF and TF verbs on the K&M approach extends beyond mere subcategorization and theta role assignment as Schachter (1976) first pointed out for Tagalog. Here are three important ones, all of which are slightly more natural using isomorphic K&M derivations, slightly less so on standard ones.

Imperatives The AF and non-AF verbs discussed all have imperative forms with the addressee phrase missed out. That will be the nominative EA for AF imperatives, (44a), but the genitive complement of non-AF forms in *a-*, *-Vna* and *-ana*, (44b,c):

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

b. tolóry (tólotra+y) vary izy
offer offer+imp rice they
Offer them rice

c. atolóry (a+tólotra+y) azy ny vary
offer a+offer+imp them the rice
Offer them the rice

As always the EA, when present, must be definite. Note that imperative suffixes differ for AF verbs and non-AF ones, being the same for *a-, -Vna* and *-ana* verbs.

Selection Both AF and non-AF forms are selected by modal auxiliaries. *mahazo* 'permitted' and *mahay* 'able' select AF verbs, their TF roots *azo* and *hay* select non-AF verbs. Replacing AF by non-AF verbs in (45) and (46) yields ungrammaticality.

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

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

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

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

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

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

b. nokasaiko (no+kasa+ina+ko) (ny) hovakina (ho+vaky+ina) io boky io
past+intend+Vna+1.sg.gen (the) fut+read+ina that book that
I intended to read that book

Assuming that the child in general takes longer to learn complex things than simpler ones standard views should predict that the child learns AF transitive verbs before TF ones, *other things being equal*. This prediction is reinforced by the fact that the child's intransitive verbs will exhibit AF morphology (there being only one argument). This prediction is well supported for the active/passive distinction in English, where actives are plentiful in early child language and passives, especially agented ones, are scarce to non-existant (Borer & Wexler 1987; Grodzinsky and Fox

194

1998). Unfortunately there are many "other things" whose equality must be controlled to render this prediction informative. For example, the EA of verbs of any voice must be definite in Malagasy. To express "Rasoa is reading a paper" the simple TF form is not possible, since the Theme is indefinite[17]. And as we individuate inanimate objects less well than humans[18] and Agents are often human and Themes often inanimate, it is likely that Agents are more often definite in discourse than Themes, so these differences will influence the distribution of AF and TF forms in Malagasy texts. That said, predictions P1 – P4 below do seem to us well founded.

P1   On the standard views the proportion of AF forms to non-AF forms should be higher in child language than in adult language. On the K&M view the proportion should be roughly equal to the adult one.

P1 assumes that factors like definiteness which constrain the choice of voice when there is a choice are similar in child and adult contexts. Then the simplicity factor favors early AF learning on the standard treatments but not on the K&M analysis.

P2 a.   CF verbs should be learned later than primary[19] AF ones
   b.   CF verbs should be learned later than primary TF[20] ones

On the K&M view, P2a is a clear prediction since the AF verb is a proper constituent of the CF verb, so the latter cannot be formed or interpreted without the former. P2b follows from P2a plus P1: primary TF verbs are equal to AF ones in complexity and so should not be learned later. Standard views *should* also predict P2a on the grounds that CF forms require some movement in addition to whatever licenses the AF morphology. (But it is not clear what forces AF morphology and blocks TF morphology forms within CF verbs on Paul's and Pearson's approaches).

P3   From **Gen** the dominant expression of the Predicate-Argument relation will be the $Pred_{TF}+DP_{gen}$ one rather than the $Pred_{AF}+DP_{acc}$ one, so the former should outnumber the latter in the acquisition data.

P4   a. K&M expect that learning to form possessive nominals will be roughly simultaneous with the emergence of agented non-AF verbs since they are both instances of Possessor Formation. No such prediction follows on standard views as these two phenomena are independent.

   b. Moreso than on standard views K&M's suggests that the genitive Agent of a non-AF verb will be present since it is a complement of the verb and forms a constituent with it. Neither is true on the standard view.

The child acquisition study We test P1 - P4 against the results of an acquisition project conducted by the second author. For a more systematic presentation of the analysis of that project see Hyams, Manorohanta, and Nthelios (2004).

Three children, TS, SO, and NI, were recorded regularly in their home settings for an 8 month period in each case. In all cases the children spoke the major dialect, Merina, that of the capital, Antananarivo, on which "Official Malagasy" is based. TS was recorded from age 24 months to 32 months, SO from18 months to 26 months, and NI from 22 months to 30 months.

General properties of the data As expected, there are very few functional elements: one occurrence of the question particle *ve*, two occurrences of the definite article *ny*, no occurrences of the focus particle *no*, no relativizers or complementizers, and almost no prepositions and no modal auxiliaries which serve to build complex predicates (A few do occur as main verbs, without a verbal complement). There are no syntactically complex constructions, such as relative clauses, sentence complements, focus/cleft constructions including wh-questions, or nominalized predicates. Negation occurs, as do imperatives and a few prohibitives. Utterances are often phonologically collapsed. The complex morphophonemics associated with possessive formation as well as compounding and object incorporation is almost entirely absent. We turn now to P1 – P4.

First, P2 is verified. We find only four CF verbs with a total of 7 occurrences. This is most compatible with the K&M analysis, but not incompatible with the standard analysis. And additional factors may contribute to the relative absence of CF forms: A CF verb must have at least two arguments, and will have three if its argument AF verb is transitive. This is pretty much a maximum in adult speech. Further several of the structures that force a CF form are syntactically complex (RCs, Clefts/Focus, nominalizations) and so independently absent. Finally, many oblique DPs whose externalization requires a CF form cannot occur sentence finally (Rajemisa-Raolison 1971) but must be focused.

The absence of CF forms has the effect that the total of AF and TF forms have greater frequencies in child speech than adult speech. K&M (2001) compute voice distribution in two romance novels. (48) compares the voice distribution in those adult sources with that in the children studied here:

(48)   Children   Adults (Novels)

AF	890	(66.2%)	5601 (62.3%)
TF	447	(33.3%)	1846 (20.5%)
CF	7	(0.5%)	1532 (17.1%)
Total	1344		8979

These data support the K&M view over the standard view since the percentage of AF forms in the children's speech, 66%, is about the same as in the adult speech, 62%. Thus the absence of CF forms in the children's speech is largely compensated by an increase in TF forms. This is the opposite of what we would expect on the standard view where AF forms are less complex than TF ones and so should dominate early learning. But on the K&M view we expect, P1, the ratio of AF to non-AF forms to be roughly equal in the two groups. And the comparison supports this expectation.

More revealing than the gross frequency comparison of AF and TF forms is the one obtained by comparing the number of *transitive* AF verbs in the child data with the number of TF ones. This comparison is closer to telling us what verb form children use when they have a grammatical choice:

(49) Child data:	Number of X	Number of occurrences of X
X = TF verbs	107	448
X = AF transitive verbs	40	131

So children choose TF verbs more than two and half times as frequently as AF ones, and they use them more than three times as frequently. Clearly TF verbs in Malagasy are not distributed like passives in English, despite some thematic similarities.

These facts also support P3 (that $Pred_{TF}+DP_{gen}$ structures dominate $Pred_{AF}+DP_{acc}$ ones) since it shows that TF verbs, which can all host a DP_{gen} complement, strongly outnumber the AF ones that can host a DP_{acc} complement. Other factors, not well understood even for the adult language, determine whether DP_{gen} and DP_{acc} complements are overt. One factor that is understood is imperative deletion.

AF verbs built with *aN-* and *i-* (and causatives and reciprocals built from them) have imperative forms built by suffixing *-a*. But equally non-AF verbs, *a-* and *-Vna* and CF ones, have imperative forms, built by suffixing *-o* (or *-y* when an *o* occurs in the root). In fact imperatives are overwhelmingly in non-AF form in adult speech.

(50) a. novonjen-dRabe aho
      past+save+Vna-Rabe.gen 1.sg.nom
      Rabe saved me

b. vonjeo aho!
    save(imp) 1.sg.nom
    Save me!

(51) a. Namonjy ahy Rabe
      past+aN+save 1.sg.acc Rabe.nom
      Rabe saved me

b. ?Namonjia (namonje) ahy!
    Save(imp)   1.sg.acc
    Save me!

(52) In the child data there are 26 imperative verbs with 144 tokens. *All* are TF forms

The existence of TF imperatives is incomprehensible if TF verbs are regarded as Passives. It helps here to realize that what is deleted in an imperative is not the "subject" but the Addressee Phrase, which must then be the Agent and thus the EA with AF verbs, but the internal, genitive, argument with non-AF verbs. We turn now to the distribution of Agents with TF verbs in order to test P4 and further evaluate P3.

First, P4a is solidly established. Each child in our sample presents both significant numbers of TF predicates with Agent, and also possessive DPs consisting of a head N and a possessor, both pronominal and full DP. For example: TS at 25 months produces Agented TF verbs, as in (53), and DP possessives, as in (54).

197

(53) a. titako for tsy hitako 'not seeTF+1.sg.gen' = not seen by me
b. labiby for lanin'ny biby 'swallowed by the animal'
c. tintsioso for tian'i Tsoso 'liked by Tsoso'

(54) a. teleko for ny teleko 'the television+1.sg.gen' (my television set)
b. bibian zaza for biberon'ny zaza 'bottle of the baby'
c. basy dopi for basin'i Droopy 'Droopy's gun'

In all cases there is simplification of the juncture between the V or N and the DP$_{gen}$, except when mere concatenation is required (as with *hitako* = *hita+ko* 'seen by me').

Similarly SO first produces possessives at 21 months: *kasetinini* for *kasetin'i Nini*, 'cassette of Nini', *zazatati* for *zazan'i Tati* 'child of Taty' and *mapa nini* for *kapan'i Nini* 'shoes of Nini'. The first verb + genitive agent occurs at the same time: *bubela nani tati* for '*du beurre lanin'i Taty* 'butter finished by Taty', and more productively by 23 months: *nalefana* for *nalain'i Fara* 'taken by Fara'. Lastly NI also presents both genitive DPs and Pred+DP$_{gen}$ by 23 months, (55).

(55) a. kakajoko for akanjoko 'clothes+1.sg.gen'
b. silipiniseheno for silipon'i Sehena 'Seheno's underpants'

(56) a. tsy ti papakoa for tsy tian'ny papako 'not liked by my Papa'
b. titko for tsy tiako 'not liked by me'

Thus P4a is supported. The early simultaneous appearance of nominal and verbal DP$_{gen}$s supports K&M's position that the two expression types are instances of the same structure. And in adult speech the two share more than complex morphology and pronoun forms. For example DP$_{gen}$s in both license empty categories:

(57) a. Adidin'i Doda ny manaiky ny ray aman-dreni+ny
duty of art Doda.gen the obey the father and mother+his
Doda's duty is to obey his father and mother

b. Kasain'i Doda ny hamaky io boky io
intend+ina'art Doda.gen the fut+aN+read that book that
It was intended by Doda to read that book

Finally, concerning P4b, we compute:

(58) Of the 447 TF verb occurrences in the child language data, 224 or 50% have overt Agents. (Imperatives are counted as +Agent here).

This figure is high compared with English and Dutch passives, which present agent phrases from 14% to 20% of the time (K&M 2001) and so supports our treatment of the DP$_{gen}$ as a complement of the verb rather than as an adjunct. Note that in possessive DPs in adult speech the genitive behaves as a complement not a specifier:

(59) a. ity trano+n-dRabe ity   b. ny vadi+n-dRanaivo
       this house+of-Rabe this   the spouse+of-Ranaivo
       This house of Rabe's   *the spouse of Ranaivo*

On the other hand in Malagasy novels agents are present in TF forms 67% of the time (K&M 2001)[21]. And in several cases counted as agentless the missing agent is controlled. So the children lag behind the adult pattern here. We have no explanation for this lag aside from the obvious: their utterances are all short with much omitted.

There is also a last relevant respect in which the profile of the children's speech does not quite conform to that of the adult: the case distribution of DPs. Here we just counted pronouns, as usually their form is distinctive (though not always in compressed speech). Their case distribution is:

(60)   Genitive: 106   Nominative: 121   Accusative: 5

As in adult speech the accusatives are the least well represented. But we have seen, (16), that in adult speech Genitives dominate Nominatives 43% to 34%. There are at least two obvious factors that contribute to this disparity. First, short utterances dominate the child corpus, many Ss are just an intransitive verb + EA, perforce nominative. Second, prepositions, which overwhelmingly take genitive complements, are almost entirely absent from the children=s speech.

In sum, the language acquisition data in Malagasy support our analysis and interpretation of simple clauses.

Endnotes

1. E.g. Tagalog, Timugon (Prentice 1971) and Kimaragang Dusun (Kroeger 1988).

2. Or quantified over if the object DP is quantificational.

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

4. *izay* sometimes separates the N and PredP in a RC. We interpret it as follows: IZAY(F)(A) = A ∩ {x|F(x) = True}. A is a N denotation, F a PredP one.

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

6. *Izy* can be augmented to force a plural reading: *izy* ireo '3 dem+pl' = they; *izy mivady* '3 spouse' = they spouses; *izy roa lahy* '3 two man' = they two men.

7. We used pronominal replacement as a test for case. In *He left on Monday*, *Monday* was not counted accusative (in English) or genitive (in Malagasy translation) as pronominal replacement is unnatural **He left on it*. Counting such cases would have increased the proportion of accusatives in English and genitives in Malagasy.

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

9. Minimally an expression is a pair (s,C), s a string, C a category name. A language is the set of expressions derivable from a Lexicon by iterated application of some structure building functions. Each voice affix, like *aN-*, is associated with such a function. We also assume a function Merge, which maps a predicate of category S/[C1,...,Cn] and an expression of category C1 to one of category S/[C2,...,Cn].

10. A verb with no Agent subcategorized may still be Actor Focus, as *matahotra* 'fears', where the EA is Experiencer. By the Theta Hierarchy an Agent Focus verb is always Actor Focus. An intransitive verb like *mangatsiaka* 'to be cold' is both Actor Focus and Theme Focus but not Agent Focus.

11. In suffixing, an extra consonant often appears, as the v in (9). Following Erwin (1996) we treat it as part of the root. It remains in derived forms until the CV template to which Malagasy conforms is applied. At that point it drops if no suffix provides a vowel to follow it. So we have indicative *manao* (AF) and *atao* (TF), but imperative *manaova* and *ataovy*. We continue to use standard orthography.

12. There are four degenerate roots which are AF verbs but which are not "tense taking", meaning they do not mark present or past tense or use an AF prefix in nuclear clauses. There are about 20 AF verbs that do fully mark tense but do not take an AF affix. They all have normal circumstantial forms. E.g. *isy* ==> *m+isy* 'exists'; *isiana* 'circumstance of existing'.

13. Of course ANA takes AF verbs of all valencies as arguments.

14. For P an $n > 0$ place predicate denotation and F restricting, $F(P) \leq P$ means that if $F(P)(x) = \text{True}$ then $P(x) = \text{True}$. In the text we treat x as a "1-tuple".

15. But, as defined explicitly in K&S, AF Ss and their TF counterparts are not isomorphic since an isomorphism would map DP_{gen}s to DP_{acc}s or vice versa, which cannot happen as they have quite different distributions.

16. The problem is more serious on the GHT/Paul view, since movement of the Theme must cross a theta position in violation of Shortest Move [Dimitris Nthelios pc]. For Pearson Shortest Move is not violated since it only applies to A movement, not A' movement.

17. To present an indefinite Agent with an AF verb an Existential S would be used: *Misy zaza mitomany* 'Exist child cry', for "There are children crying"

18. "Which two?" is a reasonable response to "I just interviewed two of your students" but not to "I just smoked two of your cigarettes".

19. We use 'primary' here to put aside complex AF and TF verbs, such as (passives of) causatives. They virtually don't occur in our child data in any event.

20. We use TF to cover Goal Focus forms as well. In fact there is only one GF verb in our data, *ome* 'give' and it has the same TF and GF forms: *omena*.

21. This figure was given as 65% in K&M (2001). Re-examination showed that a few imperatives were omitted from the +Agent category there.

References

Abinal, R.P. et V. Malzac. 1963. Dictionnaire Malgache-Français. Mission Catholique de Tananarive. Paris: Editions Maritimes et d'Outre-Mer (reprint from 1888 edition).

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

Borer, Hagit and Kenneth Wexler. 1987. The maturation of syntax. In *Parameter Setting*, Thomas Roeper and Edwin Williams (eds). Reidel. Dordrecht.

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

Fox, Danny and Yosef Grodzinsky. 1998. Children's passive: a view from the *by*-phrase. *Linguistic Inquiry* vol 29.2: 311-332.

[GHT] Guilfoyle, Eithne, Henrietta Hung and Lisa Travis. 1992. Spec of IP and Spec of VP: two subjects in Austronesian languages. *Natural Language and Linguistic Theory* 10: 375-414.

Hyams, Nina, Cecile Manorohanta and Dimitris Nthelios. 2004. First language acquisition in Malagasy: an overview. Dept. of Linguistics, UCLA.

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

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

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

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

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

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

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

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

Kroeger, Paul. 1988. Verbal focus in Kimaragang. In *Papers in Western Austronesian Linguistics* No. 3:217 S140. Pacific Linguistics, A-78.

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

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

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

Pearson, Matthew. 2003. The Malagasy subject/topic as an A' element. To appear in *Natural Language and Linguistic Theory*.

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

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

Rabenilaina, Roger-Bruno. 1985. *Lexique-grammaire du malgache: Constructions transitives et intransitives*. Thèse de doctorat d'état. Paris Université Paris VII.

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

Rajemisa-Raolison, Regis. 1971. *Grammaire malgache*. Fianarantsoa.

Sabel, Joachim. 2003. Malagasy as an optional multiple wh-fronting language. In *Multiple Wh-Fronting*, Cedric Boeckx and Kleanthes Grohmann (eds). John Benjamins, Amsterdam. pp. 229-254.

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

Consequential *'O*-clauses in Tongan and Licensing of Null Anaphors

Yuko Otsuka
University of Hawai'i

1. Introduction

Tongan (Polynesian) manifests ergativity not only in case morphology, but also with respect to certain syntactic operations: e.g., only ABS-marked NPs can undergo relativization by the gap strategy and deletion under *pea*-coordination. As is well known, not all languages with ergative case morphology show syntactic ergativity. For this reason, syntactic ergativity has been regarded as a phenomenon independent of morphological ergativity (Bittner and Hale 1996a, b, Dixon 1979, 1994).[1] Contrary to such a view, I have proposed elsewhere that syntactic ergativity should be understood as a direct consequence of morphological ergativity, showing that in Tongan an ergative pattern arises because the relevant syntactic operation is subject to case feature matching requirement (Otsuka 2002, 2003).

There is, however, one instance of syntactic ergativity in Tongan that seems to contradict this claim. In *'o*-constructions, what follows *'o* may contain a null argument. The null argument itself may be a subject (ABS or ERG) or a direct object. On the other hand, its antecedent must be an ABS-marked NP. Apparently, case matching requirement cannot account for this phenomenon, for the null argument and its antecedent may bear different cases. I argue, however, that this condition can also be viewed simply as a by-product of morphological ergativity. Specifically, I propose that the empty category in *'o*-clauses is a null anaphor (*pro*$_{SE}$) and that the binding condition for *pro*$_{SE}$ has to do with case.

In what follows, we will first put forward the claim that syntactic ergativity results from morphological ergativity by examining relativization and *pea*-coordination in Tongan (Section 2). In Sections 3 and 4, we compare *'o*-constructions with *pea*-constructions to conclude a) that unlike *pea*, *'o* should be regarded as a complementizer rather than a conjunction; and b) that the gap in *'o*-clauses is not an outcome of deletion, but a null anaphor. We will then discuss a Minimalist approach to binding proposed by Reuland (2001) and see how it accounts for the distribution and behavior of *pro*$_{SE}$ in Tongan. Some implications of the current proposal are discussed in Section 6, with section 7 in conclusion.

2. Syntactic ergativity in Tongan: Relativization and *pea*-coordination

Tongan has two relativization strategies: the gap strategy and the pronoun strategy. The former leaves a gap in the relative clause, while the latter requires a resumptive pronoun inside the relative clause. When the relativized argument is associated with an ABS-position, the gap strategy must be used. When the relativized argument is associated with an ERG-position, a resumptive pronoun is required. See (1) below.[2,3]

(1) a. e fefine [na'e tangi *t*]
        DEF woman PST cry
        'the woman (who) cried'
    b. e fefine [na'e fili   'e Sione]
        DEF woman PST choose ERG Sione
        'the woman (who) Sione chose'
    c. *e fefine [na'e fili   'a Sione]
        DEF woman PST choose ABS Sione
        'the woman (who) chose Sione'
    d. e fefine [na'a *ne* fili   'a Sione]
        DEF woman PST 3.S choose ABS Sione
        'the woman (who) she chose Sione'

Otsuka (2002) proposes that the gap in a relative clause must be licensed by C through feature-matching. The operation in question is twofold. First, following Suñer (1998), C of a relative clause is assumed to have a [+pron(ominal)] feature. When this feature is strong, it forces movement of a pronominal element (i.e., relative pronoun) to [Spec, C], leaving a gap in the relative clause. When weak, it cannot license movement and as a result, the relative pronoun remains in situ and is pronounced as a resumptive pronoun. Second, it is assumed that C is also endowed with a case feature and that only the element that has a matching case feature may undergo movement. The value of C's case feature is determined in terms of the active case: i.e., NOM in accusative languages and ABS in ergative languages (cf. Bobaljik 1993, Laka 1993). In Tongan, C's case feature is claimed to be [ABS] when its [+pron] feature is strong, and [ERG], when its [+pron] feature is weak. As a result, a resumptive pronoun is licensed only in a position associated with ERG. In short, the ergative pattern in question can be viewed as a consequence of morphological ergativity: the relativization rules refer to case.

Another phenomenon concerning syntactic ergativity involves a conjunction *pea*. *Pea* connects two clauses, the second of which may contain a gap, *e*, as illustrated in (2) below.

(2) Na'e 'ita   'a Sione pea tangi *e*.
     PST angry ABS Sione and cry
     'Sione_i was angry and (he_i) cried.'

The gap in the second clause and the coreferential argument in the first clause must bear the same case. Thus, combinations such as ABS-ERG and ERG-ABS are prohibited. See (3) below.

(3) a. *Na'e poto 'a Sione pea fili   'a Pila.
        PST smart ABS Sione and choose ABS Pila
        'Sione was smart and (he) chose Pila.'
    b. Na'e poto 'a Sione pea fili   'e Pila
        PST smart ABS Sione and choose ERG Pila
        'Sione was smart and Pila chose (him).'

204

(3a) is ungrammatical because coreference is intended between the ABS argument of the first clause and the ERG argument of the second clause.

Otsuka (2003) proposes that the gap in *pea*-constructions arises due to a PF deletion operation, which may apply only if the two items have identical sets of features. Given this requirement, an argument in the second clause can be deleted only if it has the same case as the coreferential argument; hence the ergative pattern.[4] Again, syntactic ergativity can be viewed as a reflex of morphological ergativity: the gap must be licensed by case-feature matching.

3. Properties of *'o*-clauses

'O-constructions demonstrate an ergative pattern different from that of *pea*-coordination. Case matching is not a requirement. Rather, the relevant condition is that the antecedent must bear ABS. Combination such as ABS-ERG is permitted, whereas ERG-ERG is prohibited, as shown in (4) below.

(4) a. Na'e ha'u 'a Hina$_i$ 'o taa'i e_i 'a Mele.
　　　 PST come ABS Hina and hit 　ABS Mele
　　　 'Hina came and hit Mele.'

　 b. Na'e fili 　'e Sione$_i$ 'a Mele$_j$ 'o fiefia.
　　　 PST choose ERG Sione ABS Mele and happy
　　　 'Sione chose Mele and (*he/she) was happy.'

　 c. *Na'e tamate'i 'e Lisiate$_i$ 'a Tevita 'o ma'u e_i 'a e koloa.
　　　 PST kill 　　ERG Lisiate ABS Tevita and get 　ABS DEF treasure
　　　 'Lisiate killed Tevita and got the treasure.'

'O is traditionally considered a conjunction along with *pea* and *mo*. Churchward (1953) observes the following semantic differences between these three items: *pea* is sequential ("and then"), *mo* is simultaneous ("and also"), and *'o* is consequential ("and so, as a result"). There is evidence to suggest that *'o* is a complementizer rather than a conjunction, however. First, note that unlike *pea* or *mo*, *'o* cannot conjoin anything other than clauses. As shown in (5) below, nouns, adjectives, and adverbials can be conjoined by *mo*, and PPs, by *pea*. As demonstrated in both cases, the use of *'o* results in ungrammaticality.

(5) a. 'Oku talavou mo/*pea/*'o poto 'a Mele.
　　　 PRS beautiful and 　　　 smart ABS Mele
　　　 'Mele is beautiful and smart.'

　 b. Na'e 'alu 'a Sione mo/*pea/*'o Mele ki kolo.
　　　 PST go ABS Sione and 　　　 Mele to town
　　　 'Sione and Mele went to town.'

c. Naʻe ʻalu ʻa Sione ki Fisi pea/*mo/*ʻo ki Haʻamoa.
   PST   go ABS Sione to Fiji and   to Samoa
   'Sione went to Fiji and then to Samoa.'

Second, while *pea* can be followed by a tensed full clause, the clause following *ʻo* cannot contain a tense marker.

(6)   Naʻe haʻu ʻa Sione pea/*ʻo naʻa ne fiefia.
       PST come ABS Sione and   PST 3.s happy
       'Sione came and he was happy.'

Note that overt subjects are permitted in *ʻo*-clauses as long as no tense marker is present. See (7) below.

(7)   Naʻe haʻu ʻa Sione ʻo fiefia ʻa Mele.
       PST come ABS Sione and happy ABS Mele
       'Sione came and Mele was happy.'

In this respect, *ʻo*-clauses are similar to subjunctive *ke*-clauses, which occur as sentential arguments or adjuncts. *Ke*-clauses are like English infinitival clauses (i.e., *to*-clauses) in their distribution and functions, except that they permit an overt, as well as null, subject. See (8) below.

(8) a. ʻOku sai ke (*ʻoku) ʻalu ki ai ʻa Sione.
        PRS good SUBJ   PRS   go to there ABS Sione
        'It is good that Sione goes there.'
   b.   ʻOku ako lahi ʻa Sione ke (*ʻoku) poto.
        PRS study much ABS Sione SUBJ PRS smart
        'Sione studies to be smart.'

Finally, as noted earlier, *ʻo* necessarily gives rise to a consequential interpretation. Sentences with *ʻo*-clauses are most naturally translated as "so that ...". The tense interpretation of *ʻo*-clauses can be best described as "unrealized" in the sense of Stowell (1982), who observes that this is a property of control infinitival clauses. Given that *ʻo*-clauses are generally optional constituents, they are similar to infinitival adjuncts introduced by *to*-infinitives in English. Thus, I treat *ʻo* as a complementizer introducing a consequential adjunct, which adjoins to VP.[5]

4. Null anaphor *pro*SE

In this section, we investigate the properties of the empty category that occurs in *ʻo*-clauses and claim that the empty category in question cannot be either *pro* or PRO, but should be understood as a null anaphor, which I call *pro*SE.

4.1 Empty categories in Tongan

There are three kinds of empty categories in Tongan: PRO, *pro*, and *wh*-trace.[6] An example of each is given in (9) below.

(9) a. 'Oku loto 'a   Sione [ke   ma'u PRO 'a   e   koloa]
      PRS want ABS Sione SUBJ get   ABS DEF treasure
      'Sione wants to get the treasure.'
  b. 'Oku ma'u *pro* 'a   e   koloa.
      PRS get   ABS DEF treasure
      '(He) gets the treasure.'
  c. Ko   e   hā   na'e ma'u 'e   Sione *t*?
      PRED DEF what PST get   ERG Sione
      'What did Sione get?'

Of these three, we have seen above that *wh*-traces are permissible only in positions associated with ABS. The null argument in *'o*-clauses, on the other hand, may occur in ERG-marked positions (4a). Thus, we may safely assume that the null element in *'o*-clauses is not a *wh*-trace.

The second possibility is *pro*. Note that what I call *pro* in Tongan is different from the null pronouns of the so-called *pro*-drop languages such as Spanish and Italian, which are licensed by rich agreement inflection on verbs. Rather, it is similar to the null arguments in Chinese and Japanese, which Huang (1984, 1989) treats as an instance of topic variable. Null arguments of this sort are typically context-dependent; their reference must have been established in the discourse. Accordingly, argument drop is impossible in an out-of-the-blue context. This is true of Tongan as well. Sentences with a null argument such as (9b) uttered out of context are infelicitous. In addition, *pro* in Tongan shows a unique property. It can only be third person singular.[7] The use of *pro* as a first and/or second person pronoun results in ungrammaticality, even if the reference is clearly established in the discourse. Consider (10).

(10)   Na'a ku 'alu ki he fale   kai. Na'e kai *pro* 'a   e   ika.
       PST 1.s go to DEF house eat PST   eat   ABS DEF fish
       'I went to the restaurant. (*I/he) ate fish.'

In contrast, the null argument in *'o*-clauses can be coreferential with a first/second person argument; this is not expected of *pro*.

(11)   Na'a ku/ke 'alu 'o   mohe ai.
       PST 1.s/2.s go and sleep there
       'I/you went and (I/you) slept there.'

Furthermore, recall that the antecedent of this empty argument must be in ABS. This can be also taken as evidence against the *pro* analysis, for if it were *pro*, it should be able to refer to any NP regardless of its Case. (12) below shows that *pro* in the embedded clause can refer to any argument in the matrix clause.

(12) Na'e talaange 'e   Sione kia   Mele 'oku sai   taha *pro*.
       PST tell.DIR.3 ERG Sione to.PERS Mele PRS good one
       'Sione told Mele (that) he/she is the best.'

Thus, we have reasons to believe that the empty category in *'o*-clauses is not *pro*, but something else.

Let us now turn to PRO. PRO in Tongan deviates from the general description of PRO in several respects. First, PRO occurs in a position where an overt NP is also permitted. Compare (13) below with (9a).

(13) 'Oku loto 'a   Sione [ke   ma'u 'e   Mele 'a   e   koloa]
       PRS   want ABS Sione SUBJ get   ERG Mele ABS DEF treasure
       'Sione wants Mele to get the treasure.'

This is a peculiar property of *ke*-clauses, which is discussed extensively in Otsuka (2000). *Ke*-clauses permit both PRO and an overt NP in the subject position. The reason why the empty category in (9a) is regarded as PRO is that it cannot occur in the object position, as shown in (14).

(14) *'Oku loto 'a   Sione [ke   fili   'e   Mele PRO]
        PRS want ABS Sione SUBJ choose ERG Mele
       'Sione wants Mele to choose PRO.'

This fact distinguishes PRO from *pro* in Tongan, for nothing prohibits *pro* in the object position. In fact, (14) is grammatical if the empty category is taken to be *pro*, which in turn is taken to be coreferential with the discourse topic, but not the matrix subject.

(15) 'Oku loto 'a   Sione [ke   fili   'e   Mele *pro*]
       PRS   want ABS Sione SUBJ choose ERG Mele
       'Sione$_i$ wants that Mele choose him$_{*i/j}$.'

There is another unique property of PRO in Tongan: PRO cannot occur as the subject of an intransitive verb. As shown in (16) below, when the embedded clause is intransitive, it must contain an overt pronoun if the embedded subject and the matrix subject are taken to be coreferential. On the other hand, if the embedded clause contains a null argument, it cannot be coreferential with the matrix subject. That is, control by the matrix subject is not possible. The only possible interpretation is the one where the empty category is taken to be *pro*, whose reference is determined by the discourse topic rather than the matrix argument.[8]

(16) a. 'Oku loto 'a   Sione [ke *(ne) 'alu].
          PRS want ABS Sione SUBJ 3.s go
          'Sione wants to go.'

b. 'Oku loto 'a   Sione [ke 'alu *pro*].
    PRS   want ABS Sione SUBJ go
    'Sione$_i$ wants him*$_{i/j}$ to go.'

The null argument in *'o*-clauses shows neither of these properties. First, it can occur in as a direct object, as shown in (17a). Second, it can also occur as the subject of an intransitive verb and be coreferential with the matrix argument. See (17b).

(17)   a. Na'e ngāue lahi 'a Sione 'o fili   'e he pule *e*.
           PST work much ABS Sione and choose ERG DEF boss
           'Sione worked hard so the boss should choose (him).'
       b. Na'e tokoto 'a Sione 'o   mālōlō *e*.
           PST lie.down ABS Sione and rest
           'Sione lay down in order to rest.'

In summary, we have seen that the empty category in *'o*-clauses is neither *pro* nor PRO.

4.2 Null anaphor pro$_{SE}$

To recapitulate, the null argument in *'o*-clauses has the following properties: a) unlike *pro*, it can be other than third person singular; b) unlike PRO, it can occur as O; c) its antecedent must be in the matrix clause and ABS-marked; and d) it can be replaced by an overt pronoun. As for the last property, it has an interesting effect. As long as *'o*-clauses contain an overt pronoun, the ergative constraint does not take effect: the pronoun can refer to the matrix argument that is marked as ERG.

(18)   a. Na'e fili   'e Sione 'a Mele 'o *e* fiefia.
           PST choose ERG Sione ABS Mele and happy
           'Sione chose Mele so *he/she should be happy.'
       b. Na'e fili   'e Sione 'a Mele 'o ne fiefia.
           PST choose ERG Sione ABS Mele and 3.s happy
           'Sione chose Mele so he/she should be happy.'

This suggests that what is subject to the relevant constraint is not *'o*-clauses per se, but rather the empty category in question.

I propose that the null argument in *'o*-clauses is a null anaphor, whose feature specification is [+ana(phor), −pronominal]. Its [+ana] feature makes it subject to binding principles, which require that this element be bound (in its binding domain). Hence, the null element in *'o*-clauses must have an antecedent in the same sentence. Two facts need to be explained, however: a) its antecedent must be outside the *'o*-clause; and b) its antecedent must be in ABS.

It should be noted that there are two kinds of anaphors: those that permit long-distance binding and those that do not. The former is known to exist in languages like Japanese, Chinese, and Icelandic. It is known that long-distance

anaphors are of simplex form, and generally underspecified for phi-features if the language generally realizes phi-features morphologically. For example, in Japanese, the long distance anaphor *zibun* "self" contrasts with the local anaphor *zibun-zisin* "self-self"; in Dutch, the long distance anaphor is *zich* while the local anaphor is *zichszelf*. Following Reinhart and Reuland (1991, 1993) and Reuland and Reinhart (1995), I refer to the former as SE anaphors and the latter, SELF anaphors. It is SE anaphors that permit non-local binding. The empty category in 'o-clauses exhibits a property of non-local anaphors, suggesting that it is a kind of SE anaphor. For the sake of convenience, I call this element *pro*$_{SE}$, a phonetically null reflexive pronoun of the SE-type. In the following section, we will discuss how binding relation is established in the phase based Minimalist framework (Chomsky 2000, 2001).

5. Binding in the Minimalist Program

Given that government is not regarded as a core syntactic relation in the Minimalist Program, the Binding Theory (Chomsky 1981) requires a drastic reinterpretation. Reuland (2001) proposes that binding, just like any other syntactic relations, should be understood as a consequence of basic operations such as Merge and Move. Below, I discuss how Reuland's approach can account for the properties of *pro*$_{SE}$ in Tongan.

5.1 Binding as a formal dependency

Reuland (2001) claims that binding of SE-anaphors is a formal dependency which he calls CHAIN. This dependency is established between SE and its antecedent by virtue of movement of SE's formal features. Reuland's analysis is based on the following assumptions. First, whenever a checking configuration arises, checking must take place. Second, interpretable features can delete if such deletion would not violate the principle of recoverability of deletion. That is, if no information is lost due to deletion of the relevant features. Third, feature chains exist (contra Chomsky (1995, 2000)).

Note that SE anaphor in Dutch (*zich*) is only partially specified for features: only for person (i.e., third person) and not for gender or number. Reuland (2001) argues that this property enables SE to form a CHAIN with its antecedent in the following fashion. First, SE's formal features (FF$_{SE}$) adjoin to V in order to check its case feature.[9] As V overtly raises to T in Dutch, FF$_{SE}$ raising yields a configuration in which FF$_{SE}$ is in a Spec-head relation with the NP in [Spec, T]. Note that prior to FF$_{SE}$ raising, V's phi-features were checked by the NP in its Spec and deleted, for they are uninterpretable. Meanwhile those on the NP would not delete, for they are interpretable. When SE adjoins to V in T, this creates a checking configuration (i.e., Spec-head) between SE and the NP in [Spec, T]. Given that feature checking must apply whenever a checking configuration arises, SE's phi-features are automatically checked against those of the NP in [Spec, T]. This establishes a dependency between SE's phi-features and the NP in [Spec, T] in a way similar to the dependency established between the subject and T.

In Reuland's (2001) analysis, the difference between SE anaphors and pronominals lies in their feature composition. Recall that establishment of dependency is possible in this case because SE (i.e., *zich*) is only specified for person, which is an inherent feature and therefore, context independent. Hence, deletion of an occurrence of this feature under identity with another occurrence would not result in loss of any information that is necessary for interpretation. In contrast, such a dependency cannot be formed with a pronominal and the NP in [Spec, T], for pronominals (e.g., *hem*) are fully specified for features including number and gender. As number is an optional feature and is context dependent, deletion of an occurrence of such a feature is unrecoverable.

5.2 Pro$_{SE}$ in Tongan

In Reuland's model, binding relation is taken to be a formal dependency established by checking configuration between a head and XP in its Spec (specifically, the subject and T). The bindee and its antecedent can enter into this relation by virtue of FF raising to V. This FF movement is motivated by case reasons: the object NP must check its case. Two additional conditions must be met. First, V-raising to T is overt and hence precedes FF raising to V. Second, FF of the NP in question must be only partially specified, consisting of only interpretable and inherent features such as person, as in the case of *zich* in Dutch.

The first condition is met in Tongan. V-raising is obligatory and overt, as evidenced by the V-initial word order. The second condition is also met in Tongan; for the null anaphor can be bound by NPs of any number or person (gender is not a morphosyntactic category in Tongan). This situation is slightly different from *zich*, as it seems that *pro$_{SE}$* is not specified at all for phi-features. We may postulate that *pro$_{SE}$* has a categorical feature [+D], unspecified phi-features [person, number], and a case feature. Given the requirement of Full Interpretation (FI), these features must be checked. As far as phi-features are concerned, they must be assigned some values. Binding, then, can be understood as phi-feature value assignment in this case. It is motivated by the FI, and is achieved by feature checking. Feature checking is achieved by feature match, which should be understood as feature-type match rather than feature-value match (cf. Chomsky 2000). *Pro$_{SE}$*'s phi-features are checked by matching features of the same type, i.e., phi-features of its antecedent. Through this checking process, the two sets of phi-features are identified, with those of *pro$_{SE}$* deleted as a result, establishing a formal dependency between the two elements.

Adopting Reuland's (2001) model, I propose the following analysis of *pro$_{SE}$* in Tongan. FF of *pro$_{SE}$* raises to attach to *v* of the matrix clause and enter into a checking configuration with the NP in [Spec, *v*], whereby a formal dependency (i.e., binding) is established between the two. There are, however, some issues that must be addressed regarding how the proposed FF raising is motivated and achieved in Tongan. Note that Tongan differs from Dutch in two respects. First, case marking in Tongan is ergative, which presumably involves a mechanism quite different from that of an accusative system. Second, unlike *zich* in Dutch, only non-local binding is possible for *pro$_{SE}$* in Tongan: its antecedent

must be outside 'o-clauses and, as we will see later, it cannot occur in simple sentences.

In the current study, ABS and ERG are both taken to be structural; the former is checked by v and the latter, by T (Otsuka 2000, 2002, 2003). The difference between ergative and accusative systems is essentially the choice of "active" functional head, T in accusative languages and v in ergative languages (cf. Levin and Massam 1984, Bobaljik 1993, Laka 1993). The "active" functional head is the one always available for case checking. In accusative languages, T is always available and has a case feature [NOM], whereas v is only present in transitive constructions. In contrast, in ergative languages, v with [ABS] is present in intransitive constructions as well as transitive, while T may or may not be endowed with a case feature [ERG] (Otsuka 2003). When T is not endowed with a case feature, only intransitive constructions can survive the derivation. The derivation would crash if the verb is transitive, for the subject NP would never be able to have its case feature checked.

Let us consider how pro_{SE}'s case is checked. Given that overt NPs, both ABS and ERG, can occur in 'o-clauses, we may assume that both ABS and ERG are available in 'o-clauses. That is, pro_{SE} can and therefore, must have its case checked inside 'o-clauses. If case is the motivation for FF-raising that is necessary to establish a binding relation, pro_{SE} does not have any reason to move its FF to any higher position, for the relevant feature has already been checked. Thus, the trigger for FF-movement must be some feature(s) other than case. I propose that it is phi-features that motivate FF-movement.

5.3 Analysis

The gist of the current proposal is that pro_{SE} and its antecedent establish a formal dependency through feature checking. This dependency is established by FF raising of pro_{SE}'s phi-features to the matrix v. In order for such movement to be licensed, however, the following conditions must be met. First, pro_{SE}'s phi-features must not delete even after its case feature has been checked. This is problematic given Chomsky's (2000) approach, in which a) case is considered to be a by-product of phi-feature checking and b) an NP becomes inactive to computation once its case is checked. Second, the matrix v's phi-features must also remain accessible so that Agree can apply to v and pro_{SE}. This raises a question, considering that v's phi-features are uninterpretable, which are generally assumed to obligatorily delete upon checking (Chomsky 1995, 2000). Thus, the current analysis requires two additional stipulations in order to meet these two conditions.

(19)   a.   An NP remains active as long as it has some uninterpretable feature.
            Features with unspecified values are uninterpretable.
       b.   Uninterpretable features on an active functional head remain accessible until the relevant phase is completed.

Let us consider the derivation in question step by step. First, Agree applies to pro_{SE} and the relevant functional head (v or T) inside the embedded 'o-clause.

Recall that Agree holds in terms of feature types rather than feature values. In this case, the relevant feature type is phi-features. As a result of this operation, pro_{SE}'s case feature is checked and assigned a specific value (ABS or ERG). Note, however, that its phi-features fail to receive any specific values, for v's phi-features are uninterpretable. Thus, pro_{SE} still has features whose values are unspecified, whence the FI cannot be satisfied. Given (19a), pro_{SE} remains active although its case feature has been successfully checked.

The next step builds a new phase, vP on top of the $'o$-clause. As mentioned above, $'o$-clauses are taken to be adjoined to VP to its left.[10] In this phase, the following operations take place. First, Agree applies to the phi-features of the matrix v and those of NP inside the vP (subject if V is intransitive, object if V is transitive). As a result of this operation, phi-features on v and on the NP are checked, and the NP's case feature is assigned a specific value, [ABS]. Due to (19b), however, v's phi-features still remain accessible to subsequent operations within the phase. Thus, Agree applies to v and pro_{SE}. By virtue of this operation, pro_{SE} forms a dependency with the ABS-NP compositionally; there is a dependency between pro_{SE} and v, which in turn forms a dependency with the NP through the previous application of Agree. Pro_{SE}'s phi-features identified with those of the NP and the operation forms a CHAIN <NP, pro_{SE}>. As a consequence, pro_{SE} is interpreted as coreferential with the ABS-NP in the matrix clause.[11]

6. Discussion: Long-distance anaphors and case

The current analysis explains how the null argument in $'o$-clauses is interpreted as coreferential with the ABS-marked NP of the matrix clause. It also correctly predicts that coreference between pro_{SE} and the ERG-marked NP is impossible. Note that in the present approach, in order to establish a formal dependency with the ERG-marked NP, pro_{SE}'s phi-features must adjoin to T instead of v. However, such movement is impossible for two reasons. First, note that T is not the active functional head in Tongan. Thus, T's phi-features become automatically inaccessible once they are checked against the subject NP (i.e., the ERG-marked NP). Agree cannot apply to T and pro_{SE}. Second, assuming featural cyclicity (Chomsky 1995), feature checking must take place at a first opportunity. Furthermore, given the Phase Impenetrability Constraint (Chomsky 2000), feature checking cannot take place across a phase boundary. In other words, pro_{SE} must check its phi-features when the first opportunity arises, namely, within the matrix vP. Note that T is outside the relevant phase. Thus, pro_{SE} does not have access to T unless it moves up to the next phase (i.e., CP). Yet, it cannot do so, since v provides an opportunity for feature checking.

The same argument can be used to account for the fact that pro_{SE} cannot occur in a simple sentence, as shown in (20) below.

(20)   *Na'e fili   'e   Sione$_i$ pro_{SEi}
        PST   choose ERG Sione
        Intended meaning: 'Sione chose himself.'

In order to obtain the reflexive reading, the ERG-marked NP *Sione* and *pro*$_{SE}$ must form a CHAIN. However, such a CHAIN cannot be formed. Recall that T is not the active functional head in this language. Hence, once Agree has applied to T and the subject, T's uninterpretable features delete and become inaccessible to computation.

Finally, it should also be noted that the analysis proposed here has some intriguing implication for other instances of long-distance anaphors. It is well-known that non-local binding of SE anaphors is subject-oriented. Long-distance binding by a non-subject (e.g., direct object) is not possible, as illustrated by the Japanese example in (21) below.

(21)   John$_i$-ga   Bill$_j$-ni   [Peter$_k$-ga   zibun$_{i/*j/k}$-o   kaita   to]   itta.
        John-NOM Bill -DAT Peter- NOM SE   -ACC painted that said
        'John$_i$ told Bill$_j$ that Peter$_k$ painted self$_{i/*j/k}$.'

In the current approach, this property of long distance anaphors can be readily accounted for. Recall that the active functional head is T in accusative languages. Thus, the condition (19) applies to T rather than *v*. In other words, whatever uninterpretable features *v* has, they must delete once checked. Consequently, *v*'s phi-features are never available for *pro*$_{SE}$. As *pro*$_{SE}$ cannot check its phi-features (i.e., have their values specified) inside *v*P, it moves up to the next phase. On the other hand, T's phi-features remain accessible due to the condition (19a) and therefore, T can establish the required dependency with *pro*$_{SE}$. In accusative languages, the NP that forms a dependency with T is the one bearing NOM, i.e., the subject.

7. Conclusion

I have argued that the ergative pattern concerning *'o*-constructions in Tongan can also viewed as a result of morphological ergativity, i.e., the fact that the active functional head is *v*. In the current analysis, the null argument in *'o*-clauses is taken to be a phonetically null SE anaphor. Its interpretation relies on binding, a dependency established due to phi-feature matching. I have argued that the proposed phi-feature matching is possible only with the active functional head. Thus, in Tongan, it can be coreferential only with ABS-arguments. In contrast, in accusative languages its antecedent must bear NOM, the active case in the system. This is supported by the fact that long-distance anaphors in languages like Japanese and Chinese are subject-oriented. To conclude, the current study lends further support to Otsuka's (2002, 2003) position that syntactic ergativity should be understood as a reflex of morphological ergativity.

Endnotes

1. Bittner and Hale (1996a, b) propose that syntactic ergativity arises when ABS-NPs are structurally higher than ERG-NPs and that the structural position of ABS is determined by a parameter concerning VP's transparency to government. In their system, it is assumed that C assigns ABS under government. If VPs are opaque to government, ABS-NPs must move to [Spec, IP] in order to receive case and end up higher than ERG-NPs.

2. The subject pronouns are clitic and therefore occur immediately after the tense marker, while the default word order in Tongan is VSO. Also note that the past tense marker *na 'a* has an allomorph *na 'e*, which is used when followed by a clitic pronoun.

3. Abbreviations used in this paper are as follows: ABS = absolutive, ACC = accusative, DAT = dative, DEF = definite, DIR = directional, ERG = ergative, NOM = nominative, PERS = personal, PRED = predicate, PRS = present, PST = past, S = singular, SUBJ = subjunctive, 1 = first person, 2 = second person, 3 = third person.

4. This deviates from Chomsky's (1995, 2000) claim that case features are deleted once checked. I argue that case features on NPs remain accessible after checking in the sense that their value is accessible to the computation, as I have proposed elsewhere on an independent ground concerning the distribution of resumptive pronouns (Otsuka 2002).

5. Chung (1978) also treats *'o* as a complementizer, but in a slightly different way. For her, *'o* is a morphological consequence of raising, replacing the subjunctive particle *ke*, similar to the alternation between *that* and *to* in English. Thus, *'o*-clauses are considered to be complements rather than adjuncts. The phenomenon Chung refers to as "raising" is illustrated in (i) below.

(i) a. 'oku lava ke ha'u 'a Sione.
       PRS able SUBJ come ABS Sione
       'Sione can come.'

    b. 'oku lava 'a Sione 'o ha'u.
       PRS able ABS Sione and come
       'Sione can come.'

I argue, however, that there is no derivational relation between (ia) and (ib), i.e., (ib) is not an instance of raising. *Lava 'o* is one of the two instances of idiomatic usage of *'o*. See Otsuka (2000) for discussion.

6. I have not found any evidence for NP-traces in Tongan. For one thing, passive does not exist in Tongan. Furthermore, predicates equivalent to *seem* and *likely* do not permit raising, but only occur with the expletive subject. See Otsuka (2000) for details.

7. This is true if the empty argument is used as a personal pronoun referring to a [+human] item. It should be noted that [−human] objects are always realized as a null pronoun even if they are [+plural].

8. This condition can be viewed as prohibition of PRO in ABS-marked positions. One could consider is a kind of syntactic ergativity in that the distribution of PRO distinguishes ABS from ERG, preferring the latter. Intriguing as it is, I will not discuss this phenomenon further in this paper for the interest of space.

9. Note that Reuland's (2001) analysis is based on Chomsky (1995), in which Case and phi-features are considered to be independent formal features.

10. Contra Kayne (1997), who claims that only left adjunction, is permitted in UG.

11. As Edith Aldridge points out, the proposed operation involves extraction out of adjuncts, which is generally taken to be illicit in UG. While acknowledging this as a potential problem, I would like to suggest a couple of possible solutions, albeit somewhat speculative. First, note that the long-standing assumption that extraction out of adjuncts is prohibited has been challenged elsewhere. Hornstein (1999), for example, permits such movement in his analysis of control into adjuncts. (Thanks to Eric Potsdam for bringing this point to my attention.) Alternatively, we may postulate that consequential 'o-clauses are complements in a way similar to Nilsen's (2000) analysis of circumstantial adverbials as complements. I hope to return to this issue in future work.

References

Bittner, Maria, and Ken Hale. 1996a. The structural determination of Case and agreement. Linguistic Inquiry 27: 1-68.

Bittner, Maria, and Ken Hale. 1996b. Ergativity: Toward a theory of a heterogeneous class. Linguistic Inquiry 27: 531-604.

Bobaljik, Jonathan. 1993. On ergativity and ergative unergatives. Papers on Case and agreement II, MIT Working Papers in Linguistics 19, ed. by Colin Philips, 45-88.

Chomsky, Noam. 1995. The Minimalist Program. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiry: The framework. Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik, ed. by Roger Martin, David Michaels, and Juan Uriagereka, 89-155. Cambridge, Mass.: MIT Press.

Chomsky, Noam 2001. Derivation by phase. Ken Hale: A life in language, ed. by Michael Kenstowicz, 1-52. Cambridge, Mass.: MIT Press.

Chung, Sandra. 1978. Case marking and grammatical relations in Polynesian. Austin: University of Texas Press.

Churchward, C. M. 1953. Tongan Grammar. Oxford: Oxford U. Press.

Dixon, R.M.W. 1979. Ergativity. Language 55: 59-138.

Dixon, R.M.W. 1994. Ergativity. Cambridge: Cambridge University Press.

Hornstein, Norbert. Movement and control. Linguistic Inquiry 30: 69-96.

Huang, C. T. James. (1984) On the distribution and reference of empty pronouns, Linguistic Inquiry 15: 531-574.

Huang, C. T. James. (1989) Pro-drop in Chinese: A generalized control theory. The Null Subject Parameter, ed. by Osvaldo Jaeggli and Kenneth J. Safir, 185-214. Dordrecht: Kluwer Academic.

Kayne, Richard. 1994. Antisymmetry. Cambridge, Mass.: MIT Press.

Laka, Itziar. 1993. Unergatives that assigns ergative, unaccusatives that assign accusative. Papers on Case and agreement I. MIT Working Papers in Linguistics 18, ed. by Jonathan Bobaljik and Colin Philips, 149-172.

Levin, Juliette, and Diane Massam. 1984. Surface ergativity" Case-/theta relations re-examined. Proceedings of NELS 15, 286-301. GLSA, University of Massachusetts.

Massam, Diane. 2000. VSO and VOS: Aspects of Niuean word order. The syntax of verb initial languages, ed. by Andrew Carnie and Eithne Guilfoyle, 97-116. Oxford: Oxford University Press.

Massam, Diane. 2001. Pseudo noun incorporation in Niuean. Natural Language and Linguistic Theory 19: 153-197.

Nilsen, Øystein. 2000. *The syntax of circumstantial adverbials*. Oslo: Norvus Press.

Otsuka, Yuko. 2000. Ergativity in Tongan. Doctoral dissertation. University of Oxford.

Otsuka, Yuko. 2002. Syntactic ergativity in Tongan: Resumptive pronouns revisited. Ms. University of Hawai'i.

Otsuka, Yuko. 2003. Two Derivations of VSO: A Comparative Study of Niuean and Tongan. Ms. University of Hawai'i.

Reinhart, Tanya, and Eric Reuland. 1991. Anaphors and logophors: An argument structure perspective. Long-distance anaphora, ed. by Jan Koster and Eric Reuland, 283-321. Cambridge: Cambridge University Press.

Reinhart, Tanya, and Eric Reuland. 1993. Reflexivity. Linguistic Inquiry 24: 657-720.

Reuland, Eric. 2001. Primitives of binding. Linguistic Inquiry 13: 561-570

Reuland, Eric, and Tanya Reinhart. 1995. Pronouns, anaphors and Case. Studies in comparative Germanic syntax, ed. by Hubert Haider, Susan Olsen, and Sten Vikner, 241-269. Dordrecht: Kluwer.

Stowell, Tim. 1982. The tense of infinitives. Linguistic Inquiry 13: 561-570.

Suñer, Margarita. 1998. Resumptive restrictive relatives: A crosslinguistic perspective. Language 74: 335-364.

Yuko Otsuka
University of Hawai'i at Mānoa
Department of Linguistics
1890 East-West Road, Honolulu, HI96822
USA
yotsuka@hawaii.edu

Two Types of Non-Noun-Incorporation

Ileana Paul
University of Western Ontario

1. Introduction

The goal of this paper is to investigate cases of apparent noun-incorporation in Malagasy, a western Austronesian language spoken in Madagascar.[1] Looking at examples such as those in (1), one may ask whether or not Malagasy has noun-incorporation.

(1) a. Rovi-body   ny harona.
         torn-bottom   DET basket
         'The basket has a torn bottom.'   [K&R: (4b')]
    b. Mandatsa-bato izy.
         AT.drop-stone 3(NOM)
         'She votes.'   [K&R: (21c)]

In (1), there is a bare noun (*vody* 'bottom' and *vato* 'stone' – the [b] is the result of a phonological rule) adjacent to the matrix predicate. Examination of these data reveal two key facts. First, despite the surface similarities between (1a) and (1b), the two bare nouns in (1) have very different structure: the first is an NP, the second is a DP with a null D°. Second, neither example illustrates noun-incorporation, as traditionally understood, whether lexical compounding or head adjunction. Instead, I will show that (1a) is an example of pseudo noun-incorporation, as proposed for Niuean by Massam (2001). The example in (1b), on the other hand, involves simple juxtaposition.

The organization of this paper is as follows: I begin with a general discussion of the distribution of nominals in Malagasy – with and without determiners. In section 3 I turn to the two constructions illustrated in (1) and compare and contrast them. Section 4 details the analyses of the two constructions and I conclude the paper in section 5.

2. Background facts

2.1 DPs

Malagasy is a VOS language with fairly rigid order. Regular nominals (proper names, nouns with determiners/demonstratives) can appear in any argument position, as shown in (2) for *ny vehivavy* 'the woman', which is underlined in each example.

(2) a. SUBJECT
         Mihiry   ny vehivavy.
         AT.sing   DET woman
         'The woman is singing.'

b. DIRECT OBJECT
Manaja <u>ny vehivavy</u> Rasoa.
AT.respect DET woman Rasoa
'Rasoa respects the woman.'

c. INDIRECT OBJECT
Nanome boky <u>ny vehivavy</u> Rasoa.
PS.AT.give book DET woman Rasoa
'Rasoa gave a book to the woman.'

d. OBJECT OF A PREPOSITION
Tezitra amin'<u>ny vehivavy</u> Rasoa.
angry with'DET woman Rasoa
'Rasoa is angry with the woman.'

I will henceforth refer to such nominals as DPs, given the presence of a determiner/demonstrative.[2] Object DPs can scramble rightwards past adverbs.

(3) a. Mamitaka <u>ny ankizy</u> matetika Rabe.
AT.trick DET child often Rabe
'Rabe often tricks the children.'

b. Mamitaka matetika <u>ny ankizy</u> Rabe.
AT.trick often DET child Rabe
'Rabe often tricks the children.'

Thus DPs in Malagasy have a "normal" argument distribution.

2.2 Other nominals
There are two types of bare nominal, corresponding to (1a) and (1b). I will illustrate each in turn.

2.2.1 Bare possessees
Malagasy has what appears to be possessor raising, discussed in detail by Keenan and Ralalaoherivony (2000). As shown by K&R, there are two main types: possessor raising to subject and possessor raising to object. The first is illustrated in (4) and (5): the possessor of the subject becomes the subject and the possessee is demoted to within VP.[3] Note that the possessee loses its determiner. The (b) examples are the raising versions of the (a) examples. Thus in (4a), *Rabe* is the genitive-marked possessor of *ny zanaka* 'the child', while in (4b) *Rabe* is the nominative subject.

(4) a. Marary ny zana-dRabe.
sick DET child.GEN.Rabe
'Rabe's child is sick.'

b. Marary zanaka Rabe.
sick child Rabe
'Rabe has a sick child.' [K&R: (3)]

(5) a. Rava   ny tranony.
       destroyed DET house.3(GEN)
       'Her house was destroyed.'
   b. Rava   trano izy.
       destroyed house 3(NOM)
       'She was house-wrecked.'   [K&R: (4d,d')]

K&R provide ample evidence that the clause-final DP in these examples is the subject.

Instances of possessor raising to object are not as widespread, but K&R give several examples. As illustrated in (6) and (7), the possessor of the object becomes the object, while the possessee is demoted, losing its determiner. In (6a), *ny gadra* 'the prisoner' is the genitive possessor of *ny fatorana* 'the bonds', while in (6b) *ny gadra* is an accusative object (accusative case is overtly marked on pronouns).

(6) a. Manala   ny fatoran'ny gadra   Rabe.
       AT.remove   DET bond.GEN.DET prisoner Rabe
       'Rabe removes the prisoner's bonds.'
   b. Manala   fatorana ny gadra   Rabe.
       AT.remove   bond   DET prisoner   Rabe
       'Rabe bond-removes the prisoner.'   [K&R: (60b,b')]

(7) a. Manety   ny volon-janany   Rabe
       AT.cut   DET hair.GEN.child.3(GEN) Rabe
       'Rabe cut his child's hair.'
   b. Manety   volo an-janany   Rabe.
       AT.cut   hair ACC-child.3(GEN) Rabe
       'Rabe hair-cut his child.'   [K&R: (60a,a')]

In the remainder of this paper, I refer to the bare possessee (e.g. *zanaka* 'child' in (4b) or *fatorana* 'bond' in (6b)) as BP.

2.2.2 Bare objects

The other type of bare nominal appears as the direct object (usually a patient or theme) of a verb. If the verb and the noun have the correct phonological form, the two optionally "bond" to form one phonological word. Simplifying somewhat, bonding occurs when the predicate ends in *–na*, *-ka* or *–tra*. This syllable drops and the first consonant of the noun, if it is a fricative, becomes a stop. The resulting word (written with a hyphen), has one main stress as shown in (8), where I use the acute accent to indicate main stress. (See Rajemisa-Raolison 1971 for discussion and further examples.)

(8) a. Mánana vóla izy.   →   Mànam-bóla izy.
       AT.have money 3(NOM)
       'She has money.'

b. Mangálatra   fáry   izy.   →   Mangàla-páry izy.
   AT.steal   sugar-cane 3(NOM)
   'She steals sugar cane.'

c. Mandátsaka   váto izy.   →   Mandàtsa-báto izy.
   AT.drop   stone 3(NOM)
   'She votes.'   [K&R: (21)]

It thus appears that noun has incorporated into the verb. I call these bare direct objects BO.

3. The status of bare nouns

In comparison with DPs, discussed in section 2.1, all bare nouns have a very restricted distribution. Moreover, all bare nouns look the same at first glance. But despite their surface similarities, the two types of bare noun have very different syntactic and semantic properties.

3.1 The similarities

As seen in example (1), repeated in (9), both BP and BO have similar phonological effects. That is, given a verb with the right phonological shape, the bare noun bonds with the verb.

(9) a. Rovitra vody ny harona.   →   Rovi-body ny harona.
    torn   bottom DET basket
    'The basket has a torn bottom.'   [K&R: (4b')]

b. Mandatsaka   vato izy.   →   Mandatsa-bato izy.
    AT.drop   stone 3(NOM)
    'She votes.'   [K&R: (21c)]

Moreover, both types of bare noun must be string-adjacent the verb and cannot be separated from the verb by an adverb.[4] The examples in (10) and (11) illustrate this order for BP, those in (12) for BO. Note that in this way, bare direct objects differ from DP direct objects (see example (3)).

(10) possessor raising to subject
    a. Maty <u>vady</u>   tampoka Rabe.
       dead   spouse   suddenly Rabe
       'Rabe was suddenly widowed.'
    b.* Maty tampoka <u>vady</u> Rabe.
       dead suddenly spouse Rabe   [K&R: (19b,c)]

(11) possessor raising to object
    a. Nanendaka   <u>akanjo</u>   an-keriny an-dRabe Rasoa.
       AT.tear-off   clothes   ACC-force ACC-Rabe Rasoa
       'Rasoa tore Rabe's clothes off by force.'

b.* Nanendaka an-keriny <u>akanjo</u> an-dRabe   Rasoa.
   AT.tear-off ACC-force clothes ACC-Rabe Rasoa   [K&R: (63b,c)]

(12) BO
   a. Mamitaka   <u>ankizy</u>   matetika   Rabe.
      AT.trick   child   often   Rabe
      'Rabe often tricks children.'
   b.* Mamitaka   matetika   <u>ankizy</u>   Rabe.
      AT.trick   often   child   Rabe

Thus bare nouns show a certain dependency on the verb, with both phonological and syntactic effects.

This dependency, however, may be interrupted in non-active sentences, where the genitive agent must appear next to the verb. (13) is an example of a BP: the noun *volo* 'hair' appears next to the verb in (15a), but appears after the genitive agent in (15b).

(13)a. Manety   <u>volo</u>   an-janany   Rabe.
       AT.cut   hair   ACC-child.3(GEN) Rabe
       'Rabe cut his child's hair.'
    b. Hetezan-dRabe   <u>volo</u>   ny   zanany.
       TT.cut.GEN.Rabe hair   DET child.3(GEN)
       'His child has his hair cut by Rabe.   [K&R: (60a'),(48a')]

Example (14) shows the same effect with BO: the bare noun *akanjo* 'clothes' appears after the genitive agent in (14b).

(14)a. Nividy <u>akanjo</u> ho an'ny ankizy   Rasoa.
       AT.buy clothes for ACC'DET child Rasoa
       'Rasoa bought clothes for the children.'
    b. Nividianan-dRasoa   <u>akanjo</u> ny ankizy.
       CT.buy.GEN.Rasoa   clothes DET child
       'Rasoa bought clothes for the children.'   [K&R: (23a,b)]

Note that these examples show that the bare noun and the predicate do not form a lexical compound. I return to this point in section 4.

3.2 Differences
Despite the surface similarities, there are important differences between BP and BO. First, as pointed out by K&R, BPs are non-referential. It is impossible to refer to them by a pronoun in later discourse, as shown in (15a). On the other hand, BOS <u>do</u> introduce a discourse referent that can be referred to, as shown in (15b).[5]

(15)a.?*Maty vady   Rabe. Efa   antitrantitra   (izy).
         dead spouse Rabe already   oldish   (3.NOM)
         'Rabe was widowed. She was already oldish.'   [K&R: (16b)]

b. Manam-bady   Rakoto   ary tiany   izy.
   AT.have-spouse   Rakoto   and love.3(GEN)   3(NOM)
   'Rakoto has a wife and loves her.'

Second, the nominalizations of verbs with BP and BO are different. For possessor raising, the genitive agent appears obligatorily outside of the BP, as illustrated by the contrast between (16a) and (16b).

(16)a. ny fahakingan-tsain-dRasoa
       DET NM.CT.quick-spirit.GEN.Rasoa
       'Rasoa's intelligence'
   b.* ny fahakingan-dRasoa   saina
       DET NM.CT.quick.GEN.Rasoa   spirit
       'Rasoa's intelligence'   [K&R: (26d,e)]

In the nominalization of a verb with a BO, on the other hand, the genitive agent comes between the verb and the BO (similar to the non-active sentence in (14b)), as seen in (17).

(17)a. ny   fangalaran-dRasoa   fary
       DET   NM.CT.steal.GEN.Rasoa   sugar
       'the theft of sugar cane by Rasoa'
   b.* ny   fangalara-parin-dRasoa
       DET   NM.CT.steal-sugar.GEN.Rasoa
       'the theft of sugar cane by Rasoa'

As a final difference, a BP may be modified, but such modification is limited, as discussed by K&R. A BO, however, can easily be modified, taking relative clause modifier in (18b).

(18)a. Maty zanaka hendry   Rabe.
       dead child   wise   Rabe
       'Rabe suffers the death of his well-behaved child.'   [K&R: (32a)]
   b. Manam-bola   nangalarinao   aho.
      AT.have-money PST.CT.steal.2SG(GEN)   1SG(NOM)
      'I have the money that you stole.'

In sum, BPS and BOS appear similar on the surface. A closer look at syntactic and semantic properties, however, reveals important differences between the two.

4. The syntax of bare nouns
In order to account for the differences between BPs and BOs, I suggest that they have different structures. In particular, BPs are NPs, while BOs are DPs.

4.1 Bare possessees

Following recent research on applicatives and related constructions (Pylkkänen 2002, Lee-Schoenfeld 2003), I suggest that possessor raising in Malagasy is a type of applicative. There is an applicative head projected in the structure, as the complement to the main verb.[6] The possessor is a DP projected in the   specifier position and the possessee is the NP complement of the applicative head.

Turning first to possessor raising to subject, the verb is intransitive and has no case features. The possessor DP requires case and raises to the first available case position, the subject position. The possessee, being an NP does not require case and remains in-situ. This is precisely what Massam (2001) calls "pseudo noun-incorporation". The tree in (19b) provides the structure of (19a).

(19)a. Marary   zanaka   Rabe.
       sick   child   Rabe
       'Rabe has a sick child.'

b.
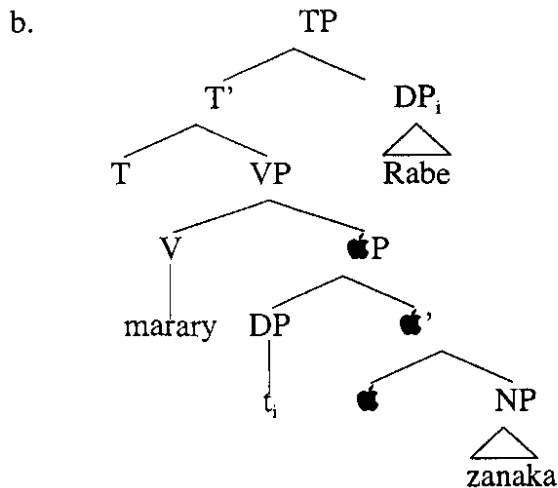
For possessor raising to object, the base structure is the same. But the host verb is transitive and therefore has the ability to assign accusative case. The possessor can therefore raise to a case position within the verbal projection, which I assume to be right-adjoined to vP (we saw earlier that objects scramble rightwards in Malagasy). As with the previous example, the possessee does not move. The example in (20) illustrates this possibility.

(20)a. Manety volo   an-janany   Rabe.
       AT.cut   hair   ACC-child.3(GEN)   Rabe
       'Rabe cuts his child's hair.'

b.

```
                        TP
                      /    \
                    T'      DP_i
                   /  \      Rabe
                  T    vP
                      /  \
                    t_i   v'
                         /  \
                        v    VP
                     manety / \
                          V'   DP_j
                         /  \   an-janany
                        V    ￼P
                       / \
                     DP   ￼'
                      |  /  \
                     t_j ￼   NP
                             volo
```
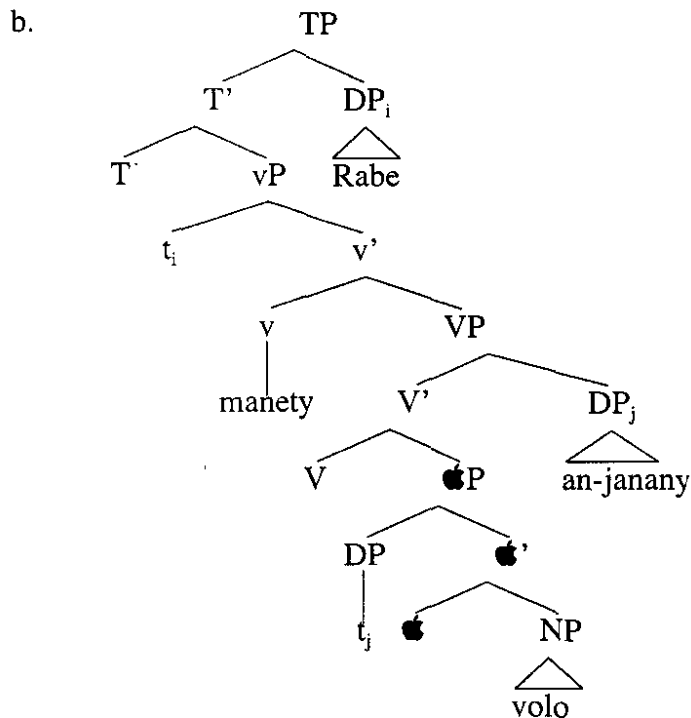
The proposed analysis accounts for the properties of the bare nouns (BPs) in possessor raising. Recall that BPs are interpreted as non-referential. Example (21a) shows that the possessee may not appear with a determiner. Example (21b) (repeating (15a)) shows that possessee does not introduce a discourse referent. Finally, (21c) illustrates the narrow scope of a BP: it obligatorily scopes under adverbs such as *indroa* 'twice'.

(21)a.* Maty   ny vady   Rabe.
        dead   DET spouse   Rabe   [K&R: (15a)]

    b.?*Maty vady   Rabe. Efa   antitrantitra (izy).
        dead spouse Rabe already   oldish   (3.NOM)
        'Rabe was widowed. She was already oldish.'   [K&R: (16b)]

    c. Maty vady   indroa   Rabe.
       dead spouse   twice   Rabe
       'Rabe was twice widowed.'
       ≠ 'Rabe's wife died twice.'

These data suggest that a BP is non-referential and I therefore conclude that it is an NP, lacking the DP layer that corresponds to referentiality.

As mentioned above, possessor raising, under this approach, is a kind of pseudo noun-incorporation (Massam 2001). Massam argues that what has been called noun incorporation in Niuean does not involve true incorporation. She shows that the incorporated element can be bigger than just a noun, but smaller than a full DP. Massam therefore concludes that the seemingly incorporated element is an NP, lacking case features and inert for syntactic movement. Her conclusions for Niuean fit nicely with the Malagasy possessor raising facts. The

BP is in fact syntactically inert and cannot be extracted, as shown by the ungrammaticality of (22).

(22)*Volo no   hetezan-dRabe   an-zanany.
    hair FOC   TT.cut.GEN.Rabe   ACC-child.3(GEN)
    'It is his child's hair that Rabe is cutting.' (and not his nails)

On the other hand, a BP is not syntactically incorporated into the verb, nor does it form a lexical compound with the verb. Thus elements such as genitive agents can intervene between the verb and the BP (as seen in (13b), repeated in (23a)) and a BP can be coordinated (23b), ruling out a head movement analysis (movement would violate the Coordinate Structure Constraint).

(23)a. Hetezan-dRabe   volo   ny   zanany.
      TT.cut.GEN.Rabe   hair   DET child.3(GEN)
      'His child has his hair cut by Rabe.   [K&R: (48a')]
    b. Maty vady   aman-janaka   Rakoto.
      dead spouse   with-child   Rakoto
      'Rabe suffers the loss of his wife and child.'

Thus the pseudo noun-incorporation analysis of possessor raising best accounts for the range of data.

4.2 Bare objects

Turning now to BO, I follow a suggestion by Zribi-Hertz and Mbolatianavalona (1997) that Malagasy has a null determiner, the covert counterpart of *ny* 'the'. In other words, BOs are in fact full DP arguments. The null determiner accounts for the fact that BOs can be interpreted as either definite or indefinite and for the possibility of relative clause modifiers, if relative clauses attach to the DP layer (24a). Moreover, as we have already seen, BOs introduce referents into the discourse (24b). Finally, because BOs are regular DP arguments, they are syntactically active and can be extracted (24c).

(24)a. Manam-bola   nangalarinao   aho.
      AT.have-money   PST.TT.steal.2SG(GEN) 1SG(NOM)
      'I have the money that you stole.'
    b. Manam-bady   Rakoto ary   tiany   izy.
      AT.have-spouse   Rakoto and   love.3(GEN)   3(NOM)
      'Rakoto has a wife and loves her.'
    c. Vola   no   nangalarinao.
      money FOC   PST.TT.steal.2SG(GEN)
      'It was money that you stole.'

To account for the ordering of bos directly after the predicate, I suggest that the null determiner must be licensed under adjacency to the predicate (Longobardi 1994). An adverb interrupts this adjacency (see (12)), but a genitive agent does

not (see (14)). This difference may arise due to the syntactic bonding that occurs between a genitive agent and a non-active predicate (perhaps via Chung's (1998) subject lowering), but I set this question aside for future research.

In sum, positing a null determiner in the above examples accounts for their syntactic and semantic properties, as well as distinguishing BPs from BOs.

4.3 What about phonology?

Although the proposed distinction between BPs and BOs accounts for their differences, it leaves open the question of their similarities. In particular, why do both trigger the bonding process?

(25) a. Rovitra vody   ny harona.   →   Rovi-body ny harona.
     torn   bottom   DET basket
     'The basket has a torn bottom.'   [K&R: (4b')]

   b. Manana vola   izy.   →   Manam-bola izy.
     AT.have money 3(NOM)
     'She has money.'   [K&R: (21a)]

I suggest that this bonding is a purely phonological process that is insensitive to syntactic structure. As mentioned earlier, bonding is optional and is determined by the phonological shape of the words rather than by syntax. Other examples of bonding show that this process is active in many different contexts, not just predicate+noun (see Rajemisa-Raolison 1971 for more examples). The example in (26a) illustrates bonding between a noun and a conjunction and (26b) shows bonding between an adjective and the noun it modifies.

(26) a. maraina sy   hariva   →   marain-tsy hariva
     morning and   evening
     'morning and evening'

   b. satroka   fotsy   →   satro-potsy
     hat   white
     'white hat'

Bonding therefore tells us nothing about the structure of BPs or BOs.

5. Conclusion

This paper has examined what appeared to be bare noun arguments in Malagasy. I have argued that in the case of possessor raising, the bare possessee is an NP. This is therefore an instance of pseudo noun incorporation (Massam 2001). Bare direct objects, however, have been shown to be full DPs, with a null D°. Thus other than the presence of a null rather than overt determiner, these bare nouns are in fact regular DP arguments.

Endnotes

1. Most of the data in this paper are from Keenan and   Ralalaoherivony (2000), henceforth K&R. Other data are from my own interviews with native speakers of Malagasy. I have slightly modified some of K&R's glosses and translations.
2. Proper names also have a determiner in Malagasy, usually *Ra* or *i*.
3. I will discuss the exact position of the possessee in section 4.1.
4. Some exceptions to this adjacency will be discussed below.
5. The referentiality of the BO needs to be examined in a wider range of examples.
6. I use the term 'verb' very broadly here to include non-verbal predicates, such as adjectives.

References

Chung, Sandra. 1998. The design of agreement: Evidence from Chamorro. Chicago: The University of Chicago Press.

Keenan, Edward and Baholisoa Ralalaoherivony. 2000. Raising from NP in Malagasy. Lingvisticae Investigationes 23: 1-44.

Lee-Schoenfeld, Vera. 2003. German possessor datives: Raised *and* affected. Ms., UCSC.

Longobardi, Giuseppe. 1994. Reference and proper names: A theory of N-movement in syntax and logical form. Linguistic Inquiry 25: 609-695.

Massam, Diane. 2001. Pseudo noun-incorporation. Natural Language and Linguistic Theory 19: 153-197.

Pylkkänen, Liina. 2002. Introducing arguments. PhD thesis, MIT.

Rajemisa-Raolison, Régis. 1971. Grammaire malgache. Fianarantsoa: Ambozontany.

Zribi-Hertz, Anne and Liliane Mbolatianavalona. 1997. De la structure à la reference : les pronoms du malgache. Les pronoms : morphologie, syntaxe et typologie, ed. by Anne Zribi-Hertz. Saint-Denis: Presses Universitaires de Vincennes.

Ileana Paul
Department of French
University of Western Ontario
London, ON N6A3K7
Canada
ileana@uwo.ca

Voice Morphology, Case, and Argument Structure in Malagasy

Matt Pearson
Reed College

1. Introduction

The canonical Malagasy clause is divided into two constituents, the *predicate phrase* (PredP), and a definite noun phrase here called the *trigger*. This is schematized in (1) (cf. Keenan 1995 for evidence for this bipartite clause structure).

(1)   [PredP V ...] [Trigger DP]

When the predicate phrase is headed by a verb, that verb is marked for *voice* to indicate the grammatical function of the trigger. Examples are given in (2) (here and throughout, the trigger is underlined; the morphological breakdown of each verb is given in brackets, with the voice morphemes boldfaced). (2a) illustrates the *actor-trigger* (AT) form, used when the external argument—here, the agent—is the trigger of the clause. The *theme-trigger* (TT) form in (2b) is used when an internal argument—here, the patient—functions as the trigger. Finally, the *circumstantial-trigger* (CT) form in (2c) is used when the trigger is an oblique element—here, the instrument with which the event is carried out. In each case, the trigger follows the predicate phrase, occurring at the right periphery of the sentence. Within the predicate phrase, the order of elements is VSOX: non-trigger external arguments appear right-adjacent to the verb, preceding internal arguments, which precede obliques.

(2) a. Mamono [**m-*aN*-**vono] akoho   amin'ny antsy ny   mpamboly
         AT.kill   **m-**Pfx-kill   chicken with-Det knife Det   farmer
         'The farmer kills chickens with the knife'
    b. Vonoin' [vono-**in**] ny   mpamboly amin'ny antsy ny   akoho
         TT.kill   kill-**in**   Det   farmer   with-Det knife Det   chicken
         'The farmer kills the chickens with the knife'
    c. Amonoan' [*aN*-vono-**an**] ny   mpamboly akoho   ny   antsy
         CT.kill   Pfx-kill-**an**   Det   farmer   chicken Det   knife
         'The farmer kills chickens with the knife'

In previous work (Pearson, to appear), I used evidence from binding, extraction, and other domains to argue that the trigger is not the subject of the sentence, as usually assumed, but is instead base-generated in the specifier of an A'-projection, *TopicP*, and linked to a null operator in the specifier of a lower projection, *WhP*, as shown in (3) (abstracting away from surface word order; on the right-peripheral position of the trigger, see Pearson (2001)).

(3)   [TopP Trigger$_i$ Top [WhP *Op*$_i$ [TP V ... t$_i$ ...]]]

The analysis in (3) suggests a novel approach to the voice alternations in (2): Generally, voice morphology is taken to encode active/passive-like alternations in

the mapping of thematic roles to the subject (nominative Case-checking) position. However, if the trigger is not the subject of the clause but a topic linked to an A'-chain, it is possible that voice morphology instead indicates the position from which A'-extraction takes place, with the AT form in (2a) marking subject extraction, the TT form in (2b) marking object extraction, and the CT form in (2c) marking oblique extraction.

Under this approach, voice in Malagasy is less like voice in English and more like *wh-agreement*, of the sort which Chung (1998) documents for Chamorro. In A'-extraction contexts in Chamorro, regular subject agreement of the sort shown in (4) is replaced by special morphology indicating whether the extracted element is a subject, object, or oblique, as in (5). In Pearson (to appear) I suggested that Malagasy voice marking is a 'generalized' version of this type of marking: While in Chamorro wh-agreement is confined to questions, relative clauses, and the like, in Malagasy it appears in all clause types due to a requirement that the specifier of WhP be filled in every clause.

(4)   **Ha**-fa'gasi si   Juan i   kareta
        3s-wash   Det Juan Det car
        'Juan washed the car'

(5) a. Hayi f<**um**>a'gasi i   kareta?   (*subj. extraction*)
        who **um**-wash   Det car
        'Who washed the car?'
    b. Hafa f<**in**>a'gasése-ñña   si   Henry pära hagu?   (*obj. extraction*)
        what **in**-wash.Prog-3s.Lnk Det Henry for   you
        'What is Henry washing for you?'
    c. Hafa pära   fa'gase-mmu ni   kareta?   (*obl. extraction*)
        what Fut   Ø-wash-2s   Obl   car
        'What are you going to wash the car with?'

In this paper I focus on the voice affixes themselves and propose an account of their distribution. Specifically, I argue that they are realizations of light verbs and Case-checking heads, which combine with the root through head-to-head movement. The distribution of the affixes is determined by the positions from which, and through which, the null operator in (3) moves on its way to the specifier of WhP. For example, the actor-topic prefix *m-* is treated as a nominative Case-checking head, which gets spelled out just in case the operator raises through its specifier. (My analysis is thus in the spirit of Guilfoyle, Hung, & Travis (1992), who also associate voice morphemes with Case licensing.)

2. The Morphology of voice
Before proceeding with this analysis, I offer a quick overview of Malagasy voice morphology. Verbs in Malagasy are formed from *roots*, some of which function independently in the language, either as nouns or as stative predicates. To form a verb stem, the root combines with one of a small set of *verbal prefixes* (glossed

'Pfx' in the examples), of which the two most common are *aN-* and *i-*. Some examples of roots and the verb stems formed from them are given in (6).

(6) ROOT   VERB STEM

    asa   'work, task'   *i*-asa   'work'
    fefy   'enclosure'   *i*-fefy   'be fenced'
    vavaka   'prayer'   *i*-vavaka   'pray'

    feno   'full'   *aN*-feno [ameno]   'fill'
    soratra   'writing'   *aN*-soratra [anoratra]   'write'
    tapaka   'broken'   *aN*-tapaka [anapaka]   'break'

Although the choice between *aN-* and *i-* is lexically determined, it does correlate to some degree with transitivity: *aN-* stems are almost always transitive, while *i-* stems tend to be intransitive. As illustrated in (7), there are a large number of roots which can take either prefix, where *aN-* forms a transitive stem and *i-* forms its intransitive counterpart:

(7)   *i*-haja   'be respected'   *aN*-haja [anaja]   'respect (tr.)'
    i-sasa   'wash oneself'   *aN*-sasa [anasa]   'wash (tr.)'
    i-voha   'be open'   *aN*-voha [amoha]   'open (tr.)'

Additional affixes are attached to roots and stems to mark voice. There are five morphologically distinct voice forms in Malagasy, listed in the table below an illustrated using *aN-tafi* 'wrap, dress' and *aN-velar* 'unroll, spread out' (two of a handful of stems which can occur in all five forms). The *actor-trigger* (AT) voice is formed by prefixing *m-* to the stem; while the *circumstantial-trigger* (CT) voice is formed with the suffix *-an*. The other three voices are usually grouped together as the *theme-topic* (TT) voices. These are formed by adding the suffix *-an* or *-in*, or the prefix *a-*. Notice that the verbal prefix is absent in the TT forms, a fact to which I return in the next section.[1]

VOICE	TEMPLATE				EXAMPLES		
AT	**m-**	*Pfx-*	ROOT		**m-***aN*-tafi	>	manafy
					m-*aN*-velar	>	mamelatra
CT		*Pfx-*	ROOT	**-an**	*aN*-tafi-**an**	>	anafiana
					aN-velar-**an**	>	amelarana
TT			ROOT	**-an**	tafi-**an**	>	tafiana
					velar-**an**	>	velarana
			ROOT	**-in**	tafi-**in**	>	tafina
					velar-**in**	>	velarina
	a-		ROOT		**a-**tafi	>	atafy
					a-velar	>	avelatra

Examples of these five forms are given in (8)-(10). The AT form is used when the trigger is the subject of the clause—that is, the sole 'core' argument of an intran-

sitive verb (8a), or the external argument of a transitive verb (8b). The TT voices are used when the trigger is the direct or indirect object of a transitive verb (9) (see sections 4 and 5 for more on the distribution of these affixes).[2] (10) illustrates the CT form, which is used when the trigger bears some oblique role such as instrument, location, or benefactee.

(8) a. Mipetraka [**m-*i*-**petrak] any   Antsirabe ny   vehivavy
      AT.live   **m-*Pfx*-**live there Antsirabe Det woman
      'The woman lives in Antsirabe'

   b. Manoratra [**m-*aN*-**sorat] taratasy ny   mpianatra
      AT.write   **m-*Pfx*-**write letter   Det student
      'The student is writing a letter'

(9) a. Tapahin' [tapak-**in**] ny   lehilahy ny   vahitady
      TT.cut   cut-**in**   Det men   Det vine.rope
      'The men cut the vine rope'

   b. Soratan' [sorat-**an**] ny   mpianatra ny   taratasy
      TT.write   write-**an**   Det student   Det letter
      'The student is writing the letter'

   c. Atao   [**a**-taov] ny   fiomanana rehetra
      TT.make **a**-make   Det preparation all
      'All the preparations are being made'

(10) a. Amonoan' [*aN*-vono-**an**] ny   mpamboly akoho   ny   antsy
       CT.kill   Pfx-kill-**an**   Det farmer   chicken Det knife
       'The farmer kills chickens with the knife'

    b. Itoeran' [*i*-toer-**an**   ] ny   ankizy   io   trano   io
       CT.live   Pfx-live-**an**   Det children this house this
       'The children live in that house'

    c. Amonoan' [*aN*-vono-**an**] ny   mpamboly akoho   ny   vahiny
       CT.kill   Pfx-kill-**an**   Det farmer   chicken Det guest
       'The farmer is killing chickens for the guests'

I now consider each of the morphemes in the above table in turn. I start with the verbal prefixes, and then turn to the AT prefix *m-* and the TT suffix *-in*. Next I consider the TT prefix *a-*. Finally, I discuss *-an*, found in the remaining TT form, and in the CT form.

3. The verbal prefixes

As shown in the table above, the verbal prefixes *aN-* and *i-* appear in AT clauses, where the subject is extracted, and in CT clauses, where an oblique element is extracted. However, these prefixes are absent in TT clauses, where an object is extracted. What's the reason for this? Recall that *aN-* and *i-* form verb stems from roots, many of which can occur independently: For instance, (12) shows that *aN-* may attach to a one-place stative root to form a two-place verb, adding an agent argument in the process.

(11) a. Feno ny tavoahangy
full Det bottle
'The bottle is full'

b. Mameno [**m**-*aN*-feno] ny tavoahangy ny vehivavy
AT.fill **m**-*Pfx*-full Det bottle Det woman
'The woman is filling the bottle'

Thinking in terms of a Hale & Keyser approach to argument structure, this suggests that the verbal prefixes are *light verbs*, which take a root or one-place predicate as their complement, and project an external argument in their specifier. This is schematized in (12) (here and below, I abstract away from head-to-head movement; I assume that verb adjoins successively to each head in its extended projection, raising as high as T before Spell Out). Assuming that the prefixes are light verbs (and assuming furthermore that obliques are merged outside of vP; cf. section 6), the proper generalization is as follows: The light verb v is suppressed (rendered covert) when an operator undergoes A'-extraction from its complement.

(12) $[_{TP} [_{vP} DP [_{v'} v [_{VP} DP [_{V'} V]]]]]$
          Pfx-   ROOT

The connection between extraction and the spell-out of the light verb receives a straightforward explanation in terms of *phases*. In his account of successive cyclic movement, Chomsky (1999, 2000) divides the clause into two domains, a *lower phase*, identified with vP, and an *upper phase*, identified with CP. In accordance with the Phase Impenetrability Condition (13) (cf. Chomsky 2000, p. 108), a constituent can move out of a phase only by first moving to its left edge, merging as a specifier of the highest head in the phase. If this is correct, then in order for an operator to extract from the complement of the light verb in Malagasy, it must first raise to become a specifier of vP, as shown in (14). In other words, Spec$v$P is an escape hatch for A'-movement to SpecWhP in TT clauses:

(13) In phase α with head H, the domain of H is not accessible to operations outside α; only H and its *edge* (i.e., the specifier(s) of HP) are accessible to such operations.

(14) $[_{WhP} \textbf{\textit{Op}}_i [_{TP} [_{vP} t_i [_{v'} DP [_{v'} v [_{VP} \ldots t_i \ldots]]]]]]$

I propose that the prefixes *aN*- and *i*- alternate with a Ø realization of the light verb. The empty light verb has a feature which attracts an operator into its specifier. By contrast, the overt light verbs *aN*- and *i*- lack this feature, and so may not host an operator. In other words, extraction from the lower phase is possible only when the head of the lower phase v is spelled out as the Ø allomorph; when v is an overt light verb, this renders it impenetrable for A'-extraction.

Alternations of this sort are well known in the head of the *upper* phase—that is, the complementizer head. In languages such as Irish, for example, the form of the complementizer depends in part on whether there is an A'-operator in the specifier of CP. The usual finite complementizer is *go*, as shown in (15); how-

ever, when an operator raises to (or through) SpecCP, the complementizer *a* is used instead, as shown in (16). I suggest that the alternation in Malagasy between the verbal prefixes *aN-* or *i-* and Ø is a lower phase analogue of this complementizer alternation: In each case, the availability of a landing site or escape hatch for A-bar movement correlates with how the head of the phase is realized. (Examples taken from Carnie, Harley, & Pyatt 2000, Guilfoyle 2000, and Chung & McCloskey 1987, respectively.)

(15) Ceapaim [**go**   bhfaca sé an madra]
      think.1s   Comp saw   he the dog
      'I think that he saw the dog'

(16) a. *Caidé* **a**   chuir sin i do   cheann?
         what Comp put   that in your head
         'What put that in your head?'
     b. an bhean [*Op* **a**   chuir t-isteach air]
        the woman   Comp put   in   on.it
        'the woman who applied for it'

4. Intransitive and monotransitive verbs

Next I turn to the AT prefix *m-* and the TT suffix *-in*. Recall that *m-* appears on the verb when the transitive or intransitive subject is the trigger, as in (17) and (18a), while *-in* appears when the transitive object is the trigger, as in (18b):

(17) Mitomany [**m-i-tomani**] ny   zazavavy
      AT.cry   **m-Pfx-cry**   Det girl
      'The girl is crying'

(18) a. Mamaky [**m-aN-vaki**] boky ny   zazavavy
         AT.read   **m-Pfx-read**   book Det girl
         'The girl is reading a book'
     b. Vakin' [**vaki-in**] ny   zazavavy ny   boky
        TT.read   read-in   Det girl   Det book
        'The girl is reading the book'

Assuming that Malagasy is a nominative-accusative language, it seems that AT marking correlates with nominative Case, while TT marking correlates with accusative Case. I propose that *m-* and *-in* are realizations of the functional heads which license nominative and accusative Case, respectively: When the operator which raises to SpecWhP has a nominative Case feature to check, the head which checks it is spelled out on the verb as *m-*, and when the operator has an accusative Case feature, the head which checks that feature is spelled out as *-in*.

First of all, consider the sentences in (19) and (20), featuring a TT verb prefixed with *a-*. As these examples show, the predicate-internal subject may be overt, as in the (b) sentences; or it may be covert, as in the (a) sentences, in which case the agent is construed as unknown or arbitrary. Notice that overt subjects im-

mediately follow the verb, which is suffixed with the *linking morpheme*, glossed 'Lnk' in the examples (this linking morpheme is realized as -*y* between two non-continuant consonants, otherwise as -*n*'; I assume that the underlying form of this morpheme is -*ny*).

(19) a. Atosika [**a**-tosek] <u>ny   vato</u>
      TT.push **a**-push   Det stone
      'Someone is pushing <u>the stone</u>' or '<u>The stone</u> is being pushed'
   b. Atosiky [**a**-tosek-*ny*] **ny** **vehivavy** <u>ny   vato</u>
      TT.push   **a**-push-***Lnk*** Det woman   Det meal
      'The woman is pushing <u>the stone</u>'

(20) a. Aroso   [**a**-roso] ny   vahiny <u>ny   sakafo</u>
      TT.serve   **a**-serve Det guest Det meal
      'Someone serves the guests <u>the meal</u>' or '<u>The meal</u> is served to the guests'
   b. Aroson' [**a**-roso-*ny*] **ny** **vehivavy** ny vahiny <u>ny   sakafo</u>
      TT.serve   **a**-serve-***Lnk*** Det woman   Det guest Det meal
      'The woman is serving the guests <u>the meal</u>'

Although the linking morpheme surfaces only on stems taking the *a*- prefix, I will assume that it is also present on stems ending in -*in* and -*an* when they have overt subjects, but fails to surface because of fusion with the *n* of the voice suffix:

(21) a. Vakina [vaki-**in**] <u>ny   boky</u>
      TT.read read-**in** Det book
      '<u>The book</u> is being read'
   b. Vakin' [vaki-**in**-*ny*] **ny** **zazavavy** <u>ny   boky</u>
      TT.read   read-**in**-***Lnk*** Det girl   Det book
      'The girl is reading <u>the book</u>'

Since the linking morpheme is required to license an overt subject in the predicate phrase, I will assume that it is located in the functional head which checks nominative Case. Adapting an analysis due to Travis (1994), I identify this as the *event head* E, which selects *v*P as its complement, as shown in (22). In addition to checking nominative Case, the event head introduces (or binds) the event argument of the clause, and is in turn selected by the tense head. As an examination of the forms above shows, the linking morpheme is in complementary distribution with the AT prefix *m*-: I will therefore assume (following Travis (1994)) that the *m*- prefix is also in E. So *m*- and -*ny* are alternate realizations of the nominative Case-checking head.

(22)   [$_{TP}$ T [$_{EP}$ E-[NOM] [$_{vP}$ DP [$_{v'}$ v ...]]]]
                      -ny / m-

What determines how E will be realized? Given my theory of triggers, this correlates with whether the subject (i.e., the nominative Case-marked argument) undergoes A'-movement or not: When the subject is an operator which raises to

SpecWhP, the verb carries the prefix *m-*, and when the subject remains inside TP, the verb carries the suffix *-ny*. To capture this alternation, I propose that how the E head is realized depends on whether or not it contains an EPP feature in addition to its Case feature—that is, whether or not it projects a specifier, triggering raising of the nominative DP (cf. Chomsky 1999, 2000): When E carries an EPP feature, it is spelled out as *m-*, and when it lacks an EPP feature, it is spelled out as *-ny*. I will refer to a head containing an EPP feature as *strong*, and a head which lacks an EPP feature as *weak*.

When the subject is an operator which raises to SpecWhP to check a feature of Wh, it must pass through the specifier of EP in order to check its Case feature. Consequently, E will have to be strong, and so it will be spelled out as *m-*. The structure is shown in (23) (abstracting away from head-to-head movement). However, when the subject does not raise to SpecWhP, it does not need to pass through SpecEP in the overt syntax. In principle E could strong or weak in such cases. Suppose that economy dictates in favor of the weak form, and so the verb carries the suffix *-ny*. In this case, the subject is spelled out in its base position, Spec*v*P, as in (22), and checks its Case via an Agree relation, without movement.

(23) [WhP ***Op***$_i$ [TP [EP t$_i$ [E' E [$_{v}$P t$_i$ [$_{v'}$ *v* ... V]]]]]]
        m-   *Pfx-*   ROOT

My analysis of the TT suffix *-in* takes essentially the same form. Alongside the nominative Case-checking head E, I assume there is a second functional head responsible for checking accusative Case, located between the light verb and the root. Following Travis (1991), I identify this as the aspect head, which projects an AspP, selected as the complement of the light verb, as in (24).

(24) [$_{v}$P (DP) [$_{v'}$ *v* [AspP Asp-[ACC] [vP DP [v' V]]]]]
              (*Pfx-*)   **Ø**   ROOT

Like the event head, the aspect head is spelled out differently depending on whether it is strong or weak. When it is strong, and triggers raising of the object for accusative Case checking, it is spelled out as *-in*; otherwise it surfaces as Ø. When the verb is intransitive, and Asp lacks an accusative Case feature altogether, the Ø variant will be selected. When the verb is transitive, either variant may be selected. Normally the Ø variant will be selected, and the direct object will remain in situ, checking its Case feature without movement, as in (24). However, when the direct object is an operator, its Case is checked via overt movement to SpecAspP. In such cases the aspect head must be strong, and so the suffix *-in* will be inserted. From SpecAspP the operator raises to the edge of the lower phase, attracted by the phonetically empty light verb, before raising out of the lower phase to Spec-WhP. This is summarized in (25):

(25) [WhP ***Op***$_i$ ... [$_{v}$P t$_i$ [$_{v'}$ (DP) *v* [AspP t$_i$ [Asp' Asp [vP t$_i$ V]]]]]]
                              Ø   **-in**   ROOT

5. Ditransitive constructions

Although the majority of monotransitive verbs mark the TT voice using the suffix *-in*, there are a number of verbs which take the suffix *-an* instead, or the prefix *a-*, as illustrated in (9) above. More often, however, *-an* and *a-* are used with ditransitive verbs. As Paul (1999) discusses, ditransitives in Malagasy fall into various semantic classes: One class, illustrated by *roso* 'serve' in (26), includes verbs of transmission, selecting a goal or recipient and a locatum: Here *-an* is used when the goal is the trigger (26b), while *a-* is used when the locatum is the trigger (26c). A second class of ditransitives is illustrated by *didi* 'cut' in (27): This class includes verbs which select a patient, and an additional argument denoting an instrument acted upon by the agent to bring about a change in the patient. With verbs of this type, *-an* is used when the patient is the trigger (27b), while *a-* is used when the instrument is the trigger (27c):

(26) a. Mandroso [**m-aN-roso**] sakafo ny   vahiny ny   mpamboly
        AT.serve   **m-Pfx-serve**   meal   Det guest Det farmer
        'The farmer serves the guests a meal'

    b. Rosoana [**roso-an**] sakafo ny   vahiny
        TT.serve   serve-**an**   meal   Det guest
        'The guests are served a meal'

    c. Aroso   [**a-roso**] ny vahiny ny   sakafo
        TT.serve   **a**-serve   Det guest Det meal
        'The meal is served to the guests'

(27) a. Mandidy [**m-aN-didi**] antsy ny   hena ny   mpamboly
        AT.cut   **m-Pfx-cut**   knife Det meat Det farmer
        'The farmer cuts the meat with a knife'

    b. Didiana [**didi-an**] antsy ny   hena
        TT.cut   cut-**an**   knife Det meat
        'The meat is cut with the knife'

    c. Adidy   [**a-didi**] ny hena ny   antsy
        TT.cut   **a**-cut   Det meat Det knife
        'The knife is used to cut the meat'

The verbs which take both *-an* and *a-* TT forms constitute the *double object* verbs of Malagasy—that is, those verbs which allow two internal DP arguments. Comparing the sentences in (26) above, we see that the trigger of the *a-* clause in (c) corresponds to the leftmost object in the actor-topic sentence in (a) (usually indefinite and adjacent to the verb), while the trigger of the *-an* clause in (b) corresponds to the rightmost object in (a). This pattern holds for (27). Following Dryer (1986), I will refer to the object farther from the verb as the *primary object* (PO), and the object closer to the verb as the *secondary object* (SO), as in (28). Hence the proper generalization is that the *-an* form indicates that the PO of a double object construction has undergone raising to SpecWhP, while the *a-* form indicates that the SO has undergone raising.

(28)   **SO**   **PO**

Mandroso   sakafo   ny vahiny   ny mpamboly
AT.serve   meal   Det guest   Det farmer
'The farmer serves the guests a meal'
Mandidy   antsy   ny hena   ny mpamboly
AT.cut   knife   Det meat   Det farmer
'The farmer cuts the meat with a knife'

Following Marantz (1993), I assume that double object constructions are really *applicative* constructions in which the applicative morpheme is null. As schematized in (29) below, Marantz treats the applicative morpheme as a kind of light verb, which selects a phrase containing the lexical root as its complement (cf. Larson 1988). The PO merges as the specifier of the applicative head, while the SO merges within its complement. Based on data from Bantu applicatives, Marantz argues that the structure in (29) compositionally represents the complex event denoted by an applicative construction, such that the PO is interpreted as affected by the sub-event denoted by the complement of the applicative head. Hence, for verbs taking a goal and a locatum, it is the goal which will be mapped to the PO function, while verbs taking a patient and an instrument will map the patient to the PO function.

(29)

primary (affected) object → DP   V'
patient, goal/recipient

                            V   VP
                  Appl

secondary object → DP   V'
locatum, instrument

                           V   ROOT
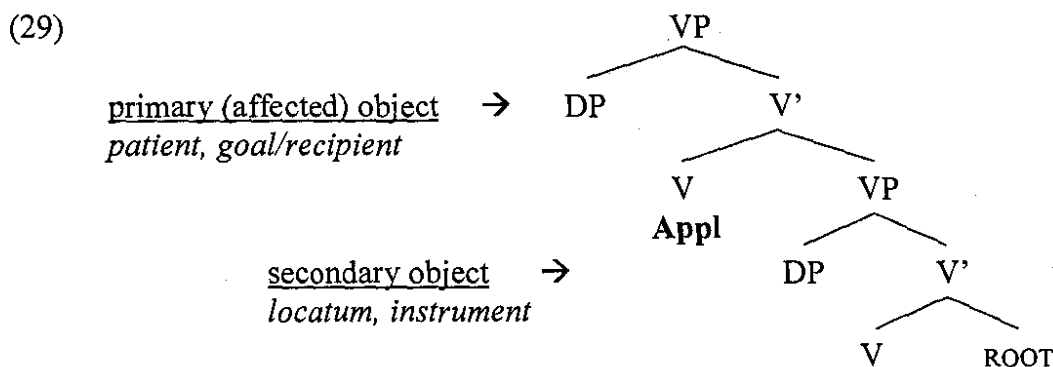
Under this analysis, the *a-* prefix receives a similar treatment to *m-* and *-in*, discussed earlier. Suppose that in double object constructions, the Case feature of the PO is checked by the Asp head, which selects the structure in (29); while the Case feature of the SO (partitive?) is checked by a lower functional head F, which takes the VP containing the root as its complement and projects an FP, selected as the complement of the applicative head. As with E and Asp, assume that the F head may be strong or weak, where the prefix *a-* spells out a strong F head. When the SO is an operator attracted to SpecWhP, its Case is checked in Spec of FP by a strong F, and the prefix *a-* appears on the verb; otherwise the SO remains in situ, and the F head is null. The structure is given in (30): The SO raises first to Spec-FP to check Case. It then raises further to the edge of the lower phase (forcing *v* to be null) before raising on to SpecWhP.

(30)   [WhP *Op*ᵢ ... [vP tᵢ [v' (DP) *v* [AspP [VP DP [FP tᵢ [F' F [VP tᵢ ... V ...
                                 Ø   **a-**   ROOT

As for -an, I treat this suffix as the applicative morpheme in (29). When the operator in SpecWhP raises from the PO position, the applicative head which takes the PO as its specifier is spelled out as -an, as schematized in (31). When the PO remains in situ, the applicative head is null.[3]

(31) [$_{WhP}$ **Op**$_i$ ... [$_{vP}$ t$_i$ [$_{v'}$ (DP) v [$_{AspP}$ t$_i$ Asp [$_{VP}$ t$_i$ V [$_{FP}$ ... V]]]]]]
                                    Ø   -an   ROOT

In support of treating -an as a spell-out of the applicative morpheme, observe the following: First, it seems that a ditransitive verb in Malagasy accepts the -an TT form if and only if it can occur in a double object construction. Consider the verb stem aN-petrak 'put' in (32), which arguably selects two arguments, a locatum and a goal. As shown in (32b), 'put' does not license a double object construction; rather, the goal must take the form of a PP, as in (32a). (32c) shows that, as expected, the -an TT form is also disallowed. This is presumably because the goal of a putting event is not normally understood to be affected by the action.

(32) a. Mametraka [**m-aN-petrak**] boky eo   ambonin'ny seza ny zaza
            AT.put   **m**-Pfx-put   book there on-Det   chair Det child
            'The child is putting books on the chair'
     b. *Mametraka [**m-aN-petrak**] boky n y seza   ny zaza
            AT.put   **m**-Pfx-put   book Det chair Det child
            'The child is putting books on the chair'
     c. *Petrahana [petrak-**an**] boky ny seza
            TT.put   put-**an**   book Det chair
            'That chair is having books put on it'

On the other hand, the -an TT form can sometimes be used to convert an intransitive stem into a transitive stem by 'promoting' a locative adjunct to the role of affected object. Consider the stem i-petrak 'sit', the intransitive counterpart of aN-petrak: Being intransitive, this stem would not be expected to take TT morphology. However, my consultant allows the -an TT form when the location ny seza 'the chair' is the trigger, as in (33b)—but crucially only if the chair is understood as affected by the act of sitting on it; otherwise the CT form would be used. I take the patterns in (32) and (33) as evidence that the TT form with -an marks A'-extraction of an affected object in an applicative construction.

(33) a. Mipetraka [**m-i-petrak**] amin'ilay seza ny zaza
            AT.sit   **m**-Pfx-sit   on-that   chair Det child
            'The child is sitting on that chair'
     b. Petrahana [petrak-**an**] ilay seza
            TT.sit   sit-**an**   that chair
            'That chair is being sat in'

One question remains, namely: Why is -an is spelled out only if its specifier raises out? Here I follow Sportiche (1992) in adopting a generalized version of the

'doubly-filled COMP' filter (34). In accordance with this filter, *-an* is spelled out when its specifier contains a trace; otherwise the applicative head is null. Note that (34) is meant to hold generally: As an examination of the structures in this paper will show, each of the projections I posit (with one exception, which I return to below) has an empty head, an empty specifier, or both, at Spell Out.

(34)   If H is a head containing some feature F, * [$_{HP}$ XP [$_{H'}$ H^0 ...]] when XP
        and H^0 both overtly encode F.

6. The circumstantial voice

I have now provided analyses of the AT form and the three TT forms in the table in section 2. This leaves the CT voice, used when the trigger is an *oblique*—that is, an element which requires a preposition in clauses where it does not function as the trigger. Examples are given in (35b) and (36b), showing CT clauses with locative and instrumental triggers, respectively.

(35) a. Mitoetra [**m-*i*-toer**] ao   amin'io trano io   ny   ankizy
        AT.live   **m-*Pfx*-live** here in-this house this Det children
        'The children live in that house'
     b. Itoeran' [*i*-toer-**an**-ny   ] ny   ankizy io   trano io
        CT.live   *Pfx*-live-**an**-Lnk   Det   children this house this
        'The children live in that house'

(36) a. Mamono [**m-*aN*-vono**] akoho   amin'ny antsy ny   mpamboly
        AT.kill   **m-Pfx-kill**   chicken with-Det knife Det farmer
        'The farmer kills chickens with the knife'
     b. Amonoan' [*aN*-vono-**an**-ny] ny   mpamboly akoho   ny   antsy
        CT.kill   Pfx-kill-**an**-Lnk   Det   farmer   chicken Det knife
        'The farmer kills chickens with the knife'

Notice that the CT form does not include any new morphemes, but rather a combination of morphemes found in other voice forms. Like the TT forms, the CT takes the linking morpheme *-ny*, indicating that the E head is weak and the subject is Case-licensed in SpecvP. The CT also includes the suffix *-an*, which means that the operator raises from the specifier of an applicative head. But unlike the TT form with *-an*, the root carries a verbal prefix. Since by assumption the lower phase is opaque to A'-extraction when a verbal prefix is present, it follows that the operator in SpecWhP in a CT clause has raised from somewhere outside of *v*P.

    Putting these pieces together, I conclude that the applicative head in Malagasy may be introduced in either the lower phase or the higher phase. In 'low' applicatives, the applicative head merges with the VP containing the root. Its specifier is a DP which checks accusative Case under agreement with the Asp head, and is interpreted as an 'affected object'. In 'high' applicatives, by contrast, the applicative head merges with some larger projection, possibly *v*P. Its specifier does not contain a DP bearing accusative Case and interpreted as affected, but rather a PP. A possible structure is given in (37).[4]

(37)   [$_{VP}$ PP [$_{v'}$ V   [$_{vP}$ DP v AspP]]]
                 Appl

Extraction of an operator from this higher applicative position yields the combination of morphemes associated with the CT voice, as shown in (38): The subject and object remain in situ, their Case-licensing heads spelled out as -*ny* and Ø, respectively. The operator raises from the specifier of the higher applicative head, causing its head to be spelled out as -*an*, in accordance with the 'doubly-filled COMP' filter. No extraction takes place from vP, and so the light verb surfaces as a prefix on the verb stem. To explain the absence of a preposition in CT clauses, we may assume (a) that the operator is of category PP, and/or (b) that the oblique operator in SpecWhP (comparable to *dont* in French) does not need Case.

(38)   [$_{WhP}$ Op_i [$_{TP}$ [$_{EP}$ E   [$_{VP}$ t$_i$ V   [$_{vP}$ (DP) v [AspP ... V]]]]]]]
                           -ny   **-an**   *Pfx-*   ROOT

One apparent problem with (38) is that the vP projection violates the 'doubly-filled COMP' filter discussed above: The head of vP is spelled out as a verbal prefix, while the specifier of vP contains the in situ subject. There are various possible solutions to this problem. For example, we could simply expand the tree so that the subject and the verbal prefix end up in different projections. A second, perhaps more interesting possibility is that vP, by virtue of being a phase, is somehow exempt from the 'doubly-filled COMP' filter. Adopting Chomsky's (1999) idea that Spell Out operates cyclically, we might speculate that the 'doubly-filled COMP' filter (plausibly a condition on Spell Out) applies only within a given cycle. Suppose that the head of a phase and its specifier are spelled out in different cycles, then it follows that vP would be able to have an overt head co-occurring with an overt specifier.

7. Conclusion

In this paper I have argued that the morphemes which combine to form the different voices in Malagasy are realizations of heads which attach to the root via head-to-head adjunction in the syntax. The verb stem-forming prefixes are light verbs, which alternate with a null variant when A'-extraction from the lower phase takes place. The AT prefix *m-*, the TT prefix *a-*, and the TT suffix -*in* are strong Case-licensing heads which surface just in case an operator raises through their specifiers. Finally, the suffix -*an*, found on the CT and on one of the TT forms, is an applicative head, which surfaces when the applicative argument raises out of its specifier, in accordance with a generalized 'doubly-filled COMP' filter.

Endnotes

241

1. For discussion of the rules for deriving the surface forms from their underlying representations, see Keenan & Polinsky (1998), Erwin (1996), Paul (1996a/b). Here I will mention only that when a consonant occurs at the end of a phonological word, an epenthetic *a* is inserted after it, in accordance with a surface ban on closed syllables (e.g., *velar-an* > *velarana*).

2. Note that the TT suffixes *-in* and *-an* are generally treated not as separate voice markers, but as lexically-conditioned allomorphs of a single marker. However, Rahajarizafy (1960) notes that there are a handful of verbs which can take either suffix, with a concomitant difference in argument structure (see Pearson 2001 for discussion). Ileana Paul (p.c.) informs me that her consultants reject Rahajarizafy's examples, allowing a stem to take *-in* or *-an* in the TT form, but not both. Since my speaker accepts Rahajarizafy's examples, I will treat *-in* and *-an* as separate morphemes—while acknowledging that the distinction between them may be disappearing for some speakers.

3. Notice that in the case of (35), the Asp head is null, rather than surfacing as *-in*, as expected. To capture this, I posit a morphological filter which blocks *-in* from attaching to a verb stem which already carries a suffix.

4. Because a clause can contain multiple PPs, I must assume that the higher applicative projection can be recursive. (It is possible that the lower applicative projection is recursive as well, although no more than projection will contain a DP specifier interpreted as an affected object.)

References

Carnie, Andrew, Heidi Harley, and Elizabeth Pyatt. 2000. VSO ordering as raising out of IP? Some evidence from Old Irish. The syntax of verb initial languages, ed. by Andrew Carnie and Eithne Guilfoyle, 39-59. Oxford: Oxford University Press.

Chomsky, Noam. 1999. Derivation by phase. MIT Occasional Papers in Linguistics 18. Cambridge, MA: MITWPL.

Chomsky, Noam. 2000. Minimalist inquiries: the framework. Step by step: Essays in minimalist syntax in honor of Howard Lasnik, ed. by Roger Martin, David Michaels, and Juan Uriagereka, 89-155. Cambridge, MA: MIT Press.

Chung, Sandra. 1998. The design of agreement: evidence from Chamorro. Chicago: University of Chicago Press.

Chung, Sandra. and James McCloskey. 1987. Government, barriers and small-clauses in Modern Irish. Linguistic Inquiry 18.173-237.

Dryer, Matthew. 1986. Primary objects, secondary objects, and antidative. Language 62.808-845.

Erwin, Sean. 1996. Quantity and moras: an amicable separation. The structure of Malagasy, vol. I, ed. by Matthew Pearson and Ileana Paul, 2-30. Los Angeles: UCLA Department of Linguistics.

Guilfoyle, Eithne. 2000. Tense and N-features in Irish. The syntax of verb initial languages, ed. by Andrew Carnie and Eithne Guilfoyle, 61-73. Oxford: Oxford University Press.

Guilfoyle, Eithne, Henriette Hung, and Lisa Travis. 1992. Spec of IP and Spec of VP: two subjects in Austronesian languages. Natural Language and Linguistic Theory 10.375-414.

Keenan, Edward. 1995. Predicate-argument structure in Malagasy. Grammatical relations: Theoretical approaches to empirical questions, ed. by Clifford Burgess, Katarzyna Dziwirek, and Donna Gerdts, 171-216. Stanford: CSLI.

Keenan, Edward, and Maria Polinsky. 1998. Malagasy morphology. Handbook of morphology, ed. by Andrew Spencer and Arnold Zwicky, 563-623. Oxford: Oxford University Press.

Larson, Richard. 1988. On the double object construction. Linguistic Inquiry 19.335-392.

Marantz, Alec. 1993. Implications of asymmetries in double object constructions. Theoretical Aspects of Bantu Grammar, ed. by Sam Mchombo, 113-150. Stanford: CSLI.

Paul, Ileana. 1996a. The active marker and nasals in Malagasy. The Structure of Malagasy, vol. I, ed. by Matthew Pearson and Ileana Paul, 49-57. Los Angeles: UCLA Department of Linguistics.

Paul, Ileana. 1996b. The Malagasy genitive. The Structure of Malagasy, vol. I, ed. by Matthew Pearson and Ileana Paul, 76-91. Los Angeles: UCLA Department of Linguistics.

Paul, Ileana. 1999. Malagasy clause structure. Doctoral dissertation, McGill University.

Pearson, Matt. 2001. The clause structure of Malagasy: A minimalist approach. Doctoral dissertation, UCLA.

Pearson, Matt. to appear. The Malagasy subject/topic as an A'-element. Natural Language and Linguistic Theory.

Rahajarizafy, Antoine. 1960. Essai de grammaire malgache. Antanimena Tananarive: Imprimerie Catholique.

Sportiche, Dominique. 1992. Clitics, voice, and spec-head licensing. GLOW Newsletter 28.46-47.

Travis, Lisa. 1991. Derived objects, inner aspect, and the structure of VP. Paper presented at NELS 22.

Travis, Lisa. 1994. Event phrase and a theory of functional categories. Proceedings of the 1994 Annual Conference of the Canadian Linguistics Association, ed. by Päivi Koskinen, 559-569. Toronto: University of Toronto.

Matt Pearson
Linguistics Program, Reed College
3203 SE Woodstock Blvd
Portland, OR 97202
USA
matthew.pearson@reed.edu

Wh-Questions in Malagasy

Eric Potsdam
University of Florida

1   Introduction

Wh-questions in Malagasy consist of a clause-initial wh-phrase followed by an invariant particle and then the remainder of the clause. This paper considers the structural analysis of Malagasy wh-questions and argues for a biclausal cleft analysis in which the initial wh-phrase is a predicate and the remaining material is a headless relative in subject position. The paper is organized as follows: Section 2 introduces some basic facts about Malagasy clause structure and wh-questions. Section 3 lays out two competing structural analyses of wh-questions: the cleft analysis and a fronting analysis in which Malagasy wh-questions are derived by wh-movement. Section 4 introduces various evidence in favor of the pseudocleft analysis and against the fronting analysis. Section 5 concludes.

2   Malagasy Syntax and Wh-Questions

Malagasy is an Austronesian language spoken on the island of Madagascar. It is well-known for having basic VOS word order, (1a), and a robust voice system which advances thematically diverse elements to the clause-final subject position. Corresponding to the active sentence in (1a), the PASSIVE sentence in (1b) has the direct object as the clause-final subject and the CIRCUMSTANTIAL sentence in (1c) has an oblique element as its subject. Note that the agent in non-active clauses appears immediately following the verb. [1,2]

(1) a. n-i-vidy   ny akoho   i Bao   ACTIVE
       PAST-ACT-buy the chicken Bao
       'Bao bought the chicken.'
    b. no-vidi-n'   i Bao ny akoho   PASSIVE
       PAST-buy-PASS Bao   the chicken
       'The chicken was bought by Bao.'
    b. n-i-vidi-anan'   i Bao ny akoho   i Soa   CIRCUMSTANTIAL
       PAST-ACT-buy-CIRC Bao   the chicken Soa
       'Soa was bought a chicken by Bao.'

For concreteness I adopt structures from Guilfoyle, Hung, and Travis 1992. In active clauses, the clause-final subject occupies a right specifier of IP and the verb raises to I•, yielding VOS word order, (2a). PredP is a projection in whose specifier the external argument is introduced (Bowers 1993, see also Chomsky's (1995) vP). For non-active clauses, the immediately post-verbal agent is in the predicate-internal subject position, spec,Pred, with the clause-final subject again in spec,I, (2b).

(2) a.

```
                IP
              /    \
            I'      DPᵢ
          /   \     i Bao
         I    PredP
     nividyₖ  /   \
     'buy'  tᵢ    Pred'
               /    \
            Pred     VP
             tₖ     /  \
                   V    DP
                   tₖ  ny akoho
                      'the chicken'
```

b.

```
                IP
              /    \
            I'      DPᵢ
          /   \    ny akoho
         I    PredP 'the chicken'
   novidin'ₖ  /   \
   'buy.PASS' DP  Pred'
             i Bao /  \
                Pred    VP
                 tₖ    /  \
                      V    DP
                      tₖ   tᵢ
```

Descriptively, wh-questions in Malagasy are formed by preposing the wh-phrase and following it immediately with the particle *no* (glossed as PRT) and then the remainder of the clause:

(3) a. iza   no   nividy   ny akoho?
         who   PRT   buy.ACT   the chicken
         'Who bought the chicken?'
     b. inona no   novidin' i Bao?
         what   PRT   buy.PASS Bao
         'What was bought by Bao?'

It is widely cited that only subjects can be questioned with this strategy (Keenan 1976, 1995, MacLaughlin 1995, Paul 2000, 2002, Pearson 2001, Sabel 2002, and others). Non-subject questions are ungrammatical:

(4) a. *inona no   nividy   i Bao?
          what   PRT   buy.ACT   Bao
          ('What did Bao buy?')
     b. *iza   no   novidina ny akoho?
          who   PRT   buy.PASS the chicken
          ('Who was the chicken bought by?')

An exception to this claim is that some adjuncts, including temporal, locative, and instrumental adverbials, can be questioned without first advancing to subject position (Keenan 1976, Rabenilaina 1998, Paul 2000, 2001, 2002, Pearson 2001, Sabel 2002, and others). (5) shows that a question with *where* is compatible with any verbal voice form. Only in the example with circumstantial voice, (5c), does the adverbial correspond to a subject.

(5) a. taiza no   nanafina ny lakileko ny zaza
         where PRT   hide.ACT the key.1SG the child
     b. taiza no   nafenin' ny zaza   ny lakileko
         where PRT   hide.PASS the child the key.1SG
     c. taiza no   nanafenan' ny zaza   ny lakileko
         where PRT   hide.CIRC   the child the key.1SG
         'Where did the child hide my key?'

This restriction, that only subjects and some adjuncts can be extracted, will be important as a descriptive generalization in discussions to follow (see MacLaughlin 1995, Sabel 2002, Paul 2002, and Pearson, to appear for possible analyses of the restriction).

3 Two Hypotheses

This section introduces two structural analyses of Malagasy wh-questions from the literature. The FRONTING ANALYSIS (Sabel 2002, 2003) likens the structure of wh-questions in Malagasy to that of wh-movement structures in better studied languages. Wh-questions, such as (6) below, resemble English wh-fronting examples and can be assigned a familiar structure, (7a). The wh-phrase moves to spec,C and *no* is question complementizer in C°.

(6)   iza   no   nihomehy?
       who   PRT   laugh.ACT
       'Who laughed?'

FRONTING ANALYSIS   CLEFT ANALYSIS
(7) a.   IP   b.   IP
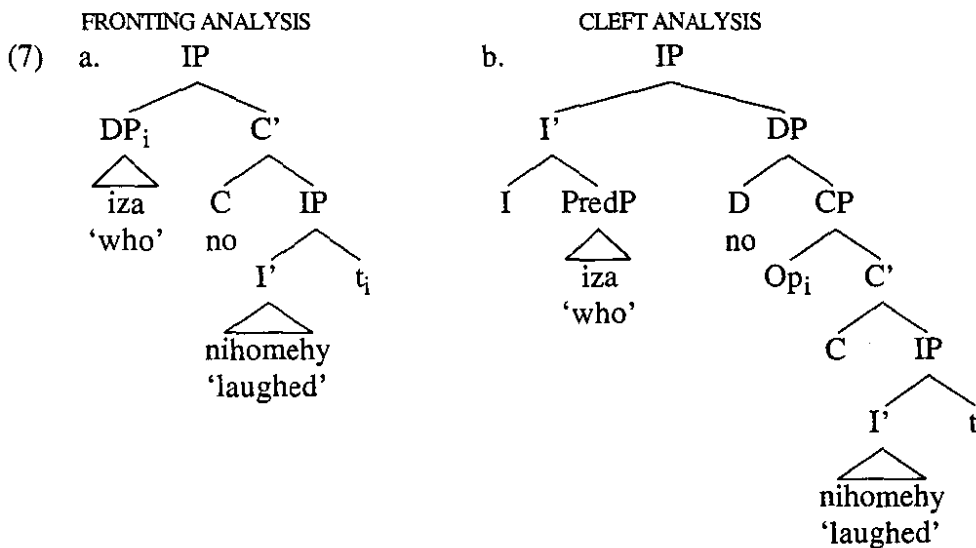
The CLEFT ANALYSIS to be defended here builds on proposals in Dahl 1986, Pearson 1996, Paul 2001, 2003b, and others. Under the cleft analysis, wh-questions are pseudocleft structures in which the wh-phrase is a non-verbal predicate and the subject is a headless relative clause involving internal movement of a null operator,

Op. The wh-question in (6) is assigned the structure in (7b). For concreteness, I take the focus particle *no* under the cleft analysis to be a determiner and the relative clause its complement, as shown. Paul 2001 points out that some related Austronesian languages are also argued to employ a pseudocleft structure for questions, for example, Palauan (Georgopoulos 1991), Tagalog, and Seediq (Aldridge 2002).

4   Argumentation Against the Fronting Analysis

This section provides evidence in favor of the pseudocleft analysis from parallels with a cleft focus construction, word order facts, and multiple wh-questions.

4.1   parallels with the focus construction

Malagasy has a focus construction illustrated in (8) that is formally similar to wh-questions. Dahl 1986 first proposed that this focus construction is a kind of cleft and Paul 2001, 2003a develops this idea, assigning (8a) the pseudocleft structure in (9). The focused element is the predicate of the clause and the subject is a headless relative clause.

(8) a. Rasoa   no   nihomehy
         Rasoa   PRT   laugh.ACT
         'It was Rasoa who laughed'
     b. ny mofo   no   novidin-   dRasoa
         the bread PRT   buy.PASS   Rasoa
         'It was the bread that was bought by Rasoa'

(9)   [[$_{predicate}$ Rasoa$_i$]   [$_{subject/headless\ relative\ clause}$ no Op$_i$ nihomehy t$_i$]]
                         Rasoa   PRT   laughed
        lit. "The one who laughed was Rasoa"

There are a number of non-trivial parallels between wh-questions and the focus construction, which suggest that they should receive the same structural analysis. First, both are formed by preposing a constituent and following it immediately with the particle *no*. Second, the two constructions have a similar focus interpretation of the initial XP. Wh-phrases indicate a request for new information in the same way that focused XPs supply new information. Third, the two constructions are subject to the same extraction restriction, that only subjects and adjuncts can be extracted. In (10), adjuncts are focused regardless of the voice of the verb (compare to the wh-questions in (5)). In (11), we see the ungrammaticality that results from focusing a non-subject argument (compare (4)).

(10) a. ao   ambanin' ny fandriana no   nanafina   ny lakile ny zaza
         there under   the bed   PRT   hide.ACT   the key   the child
     b. ao   ambanin' ny fandriana no   nafenin'   ny zaza   ny lakile
         there under   the bed   PRT   hide.PASS the child the key

c. ao   ambanin' ny fandriana no   nanafenan' ny zaza   ny lakile
    there under   the bed   PRT   hide.CIRC   the child the key
    'It's under the bed that the child hid my key'

(11) a. *ny   mofo no   nividy   i Bao
         the bread PRT   buy.ACT   Bao
         ('It's the bread that Bao bought')
     b. *Rabe   no   novidina ny akoho?
         Rabe   PRT   buy.PASS the chicken
         ('It's Rabe that the chicken was bought by')

Such similarities are perhaps unexpected under the fronting analysis in which the two constructions have rather different derivations.

4.2   predicational properties of the wh-phrase

One distinction between the two analyses concerns the characterization of the initial wh-phrase. Under the pseudocleft analysis, it is a predicate, while under the fronting analysis, it is a preposed constituent. We can thus differentiate the two analyses by seeing whether a fronted wh-phrase has properties characteristic of predicates or not. This subsection demonstrates that the position of predicate-related particles treats initial wh-phrases as predicates, not as fronted elements.

Malagasy has a number of particles that immediately follow the predicate in ordinary clauses. These post-predicate particles include the quantifiers *daholo* 'all' and *avy* 'each' (Keenan 1976, 1995), the exclamative particle *anie* (Keenan 1976, 1995), and VP adverbs such as *foana* 'always' (Pearson 1998, Rackowski 1998). There is abundant evidence that the verb and its complements constitute a predicate to the exclusion of the subject in Malagasy (Keenan 1976, 1995) and in VOS clauses these elements appear immediately after the predicate and before the subject but not elsewhere:

(12) a. nihinana vary   (daholo) ny vahiny (*daholo)
         eat.ACT   vary   all   the guest all
         'All the guests ate rice'
     b. manapaka bozaka (anie)   Rasoa   (*anie)
         cut.ACT   grass   EXCL   Rasoa   EXCL
         'Rasoa is really cutting the grass!'
     c. mihomehy (foana) Rasoa   (*foana)
         laugh.ACT always Rasoa   always
         'Rasoa is always laughing'

For concreteness, I assume that these particles right adjoin to PredP. This correctly places them between the object and the subject in VOS clauses. The two analyses under consideration make different predictions about where these particles will appear in wh-questions. As shown in (13a), the cleft analysis places them immediately after the wh-phrase. The fronting analysis, in contrast, does not and predicts that they must be lower in the structure, near the end of the clause, (13b).

248

CLEFT ANALYSIS   FRONTING ANALYSIS

(13) a.   IP   b.   CP

         I'   DP   wh$_i$ C'

      I   PredP   no ...   C   IP
                                            no
          PredP   *particle*   I'   t$_i$

            wh   I   PredP

                                              PredP   *particle*

                                                 ...
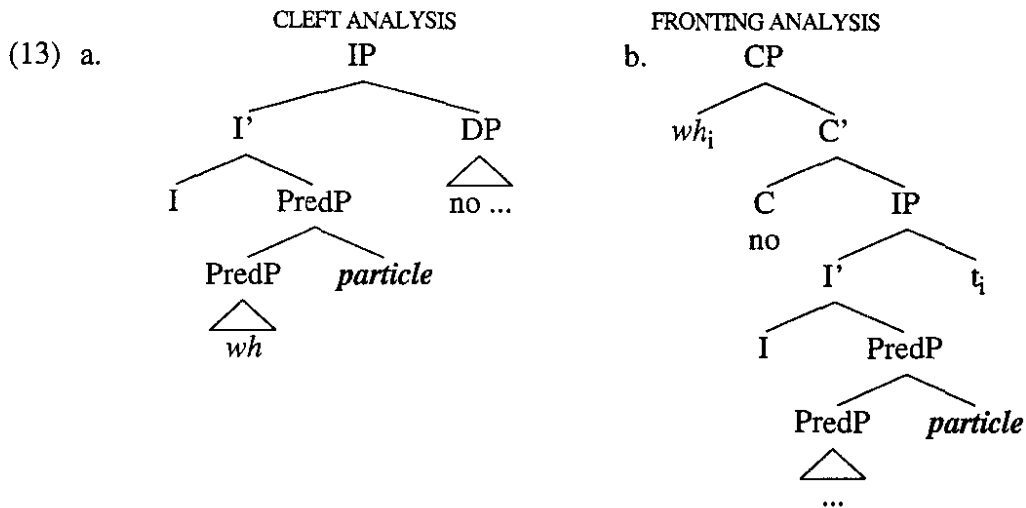
In all cases, the prediction of the cleft analysis is correct. Post-predicate particles immediately follow the wh-phrase:

(14) a. iza (daholo) no   nihinana vary (*daholo)?
        who all   PRT   eat.ACT   vary all
        'Who all ate rice?'
     b. iza (anie)   no   manapaka bozaka (*anie)?
        who EXCL   PRT   cut.ACT   grass   EXCL
        'Who is really cutting the grass?'
     c. iza (foana)   no   mihomehy (*foana)?
        who always   PRT   laugh.ACT always
        'Who is always laughing?'

Malagasy also has a number of particles that immediately precede the predicate in VOS clauses. These pre-predicate particles include *toa* 'seem', *tokony* 'should' (Paul 2001), and *tena* 'indeed (affirmative emphasis)':

(15) a. **tokony** mamangy   an-dRabe   Rasoa
        should   visit.ACT   ACC-Rabe   Rasoa
        'Rasoa should visit Rabe'
     b. **tena** hovidin' ny zaza   ny fiaramanidina
        EMPH   buy.PASS the child   the airplane
        'The child will indeed buy the airplane'

The cleft analysis of questions correctly predicts that these particles also immediately precede the wh-phrase in a wh-question, (16). Under the fronting analysis, such data are unexpected because the wh-phrase is not a predicate and the particle should necessarily appear farther to the right.[3]

(16) a. **tokony** iza no mamangy an-dRabe?
  should who PRT visit.ACT ACC-Rabe
  'Who should visit Rabe?'

 b. **tena** inona no hovidin' ny zaza?
  EMPH what PRT buy.PASS the child
  'What will the child indeed buy?'

One might object that these particles are not pre-predicate particles but clause-initial particles. If that were the correct description, then the fronting analysis would also account for the above data. There is evidence however that the particles do not simply occur clause-initially. There is a topicalization construction in Malagasy in which a constituent is preposed, followed by the topic particle *dia* (Keenan 1976):

(17) a. Rasoa dia manoroka an-dRabe
  Rasoa TOPIC kiss ACC-Rabe
  'Rasoa, she kisses Rabe'

Paul 2001 and Flegg 2003 suggest that the initial constituent in this construction is not a predicate and, indeed, the pre-predicate particles do not occur clause-initial, but predicate-initial, in this construction, (18).[4]

(18) a. (*tokony) Rasoa dia (**tokony**) manoroka an-dRabe
  should Rasoa TOPIC should kiss ACC-Rabe
  'Rasoa, she should kiss Rabe'

 b. (*tena) ny fiaramanidina dia (**tena**) hovidin' ny zaza
  EMPH the airplane TOPIC EMPH buy.PASS the child
  'The airplane, the child will indeed buy it'

In summary, the fronting analysis does not allow the grammatical placement of a wide variety of clause-internal particles, while the cleft analysis does. The wh-phrase in a wh-question behaves like a predicate with respect to the position of these particles, as predicted by the cleft analysis.[5]

4.3 multiple wh-questions

The final argument for the cleft analysis comes from a consideration of multiple wh-questions introduced in Sabel 2003, (19), in which there are two wh-phrases at the front of the clause.[6] I will claim that the cleft analysis provides a better account of such data. To see this we need to lay out how each hypothesis would handle such data. I then provide arguments in favor of the cleft-based approach.

(19) a. aiza iza no mividy ny vary?   Sabel 2003:245-246
  where who PRT buy.ACT the rice
  'Who buys rice where?'

 b. aiza (ny) inona no vidinao?
  where the what PRT buy.PASS.2SG
  'What do you buy where?'

250

4.3.1 Sabel's (2003) fronting analysis

Sabel 2003 proposes a clever analysis of the multiple wh-question examples within the fronting approach. That work argues that Malagasy is an optional multiple wh-fronting language like some of the Slavic languages. In such languages, all wh-phrases move to clause-initial position as in the Bulgarian data in (20) (Rudin 1988, Pesetsky 2000, Boskovic 2002, Richards 2002, and references therein).

(20) a. koj   kakvo na kogo   dade?   Pesetsky 2000:19
         who   what   to whom   gave
         'Who gave what to whom?'
    b. *koj   dade kakvo na kogo?
         who   gave what   to whom

Rudin 1988 first argued that there are two types of multiple wh-fronting languages. In Polish/Serbo-Croatian type languages, fronted wh-phrases move to separate CP specifiers. In Romanian/Bulgarian type languages, the fronted wh-phrases move to a single spec,C, where they form a constituent. Sabel 2003 proposes that Malagasy is of this latter type. The structure of the multiple wh-question in (19a) is (21).

(21)
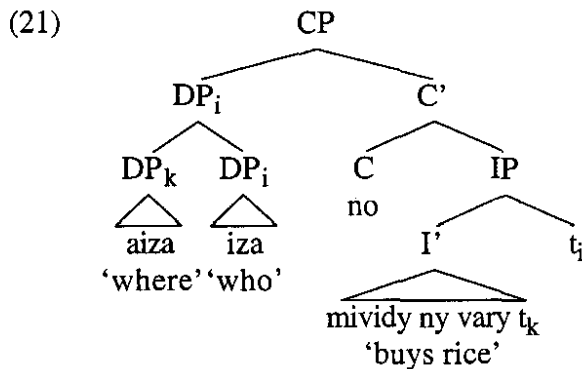
4.3.2 Paul's (2003a) cleft analysis

Paul 2003a indirectly provides an alternative analysis of the multiple wh-questions within the cleft approach. The analysis rests on the existence of another construction in Malagasy which Keenan 1976 called the BODYGUARD CONSTRUCTION. Keenan 1976 first observed that when a non-subject is questioned or focused, the subject may optionally appear immediately after the wh-phrase/focus, and before *no*. The subject serves as a "bodyguard" for the fronted element. (22b) illustrates the bodyguard variant of the wh-question in (22a) and (23b) illustrates the bodyguard variant of the focus construction in (23a). The bodyguard construction is only possible when the first XP is an adjunct and the second XP is the subject.

(22) a. aiza   no   mividy   mofo   **Rasoa**?
         where   PRT   buy.ACT   bread   Rasoa
    b. aiza   **Rasoa**   no   mividy   mofo?
         where   Rasoa   PRT   buy.ACT   bread
         'Where does Rasoa buy bread?'

(23) a. omaly   no   nividy   vary   **Rabe**
        yesterday   PRT buy.ACT rice Rabe
  b. omaly   **Rabe**   no   nividy   vary
        yesterday Rabe   PRT buy.ACT rice
     'It was yesterday that Rabe bought rice'

Paul 2003a analyzes the bodyguard construction, arguing that the bodyguard phrase is in the specifier position of the DP subject. It is a kind of possessor. This correctly places the bodyguard between the focus and the particle *no*. To the focus construction with bodyguard in (23b) Paul assigns the structure in (24).

(24)
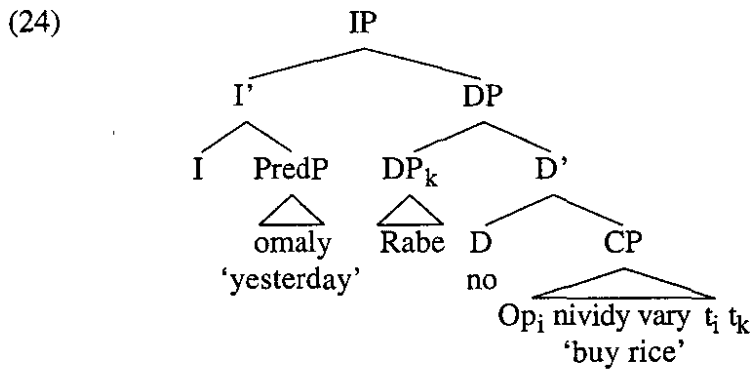
Given Paul's (2003a) analysis, we have a natural account of the multiple wh-questions. They are wh-questions in which the bodyguard in spec,D is also a wh-phrase:

(25)
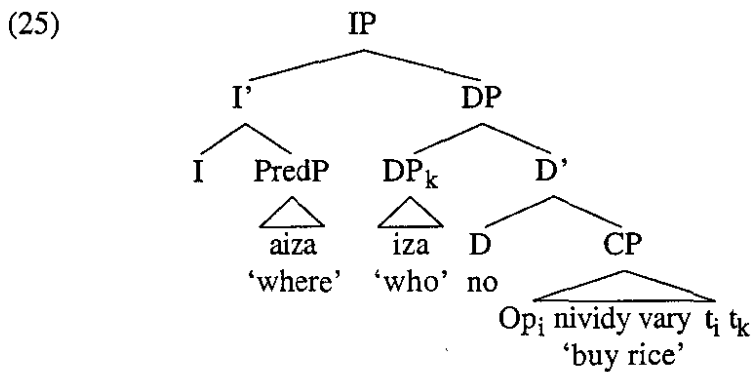
4.3.3 arguments against the fronting analysis of multiple wh-questions

In this subsection I present four arguments favoring the cleft treatment of the bodyguard construction and multiple wh-questions.

The first argument comes from a consideration of the ordering of the initial phrases and superiority. Pesetsky 2000 and Boskovic 2002, following Rudin 1988, show that Bulgarian multiple wh-fronting examples exhibit Superiority effects. The wh-phrase that originates structurally higher must appear before the wh-phrase that originates structurally lower:

(26) a. koj kuˇde udari Ivan? Pesetsky 2000:22
       who where hit   Ivan
       'Who hit Ivan where?'
    b. *kuˇde koj udari Ivan
       where who hit   Ivan
       (*'Where did who hit Ivan?')

Sabel 2003 shows that Malagasy patterns with Bulgarian in restricting the order of fronted multiple wh-phrases, (27).

(27) a. aiza iza no   mividy   ny vary? Sabel 2003:245
       where who PRT buy.ACT the rice
       'Who buys rice where?'
    b. *iza aiza no   mividy   ny vary?
       who where PRT buy.ACT the rice
       (*'Where does who buy rice?')

There is a difference between the two languages, however, in that they show the opposite licit orderings. In Bulgarian we see *who where* but in Malagasy we see *where who*. The question that arises under the fronting analysis is why Malagasy shows anti-superiority. Given that the wh-phrases constitute a constituent, under the fronting analysis, it must be the case that constituent formation in spec, C proceeds via left adjunction in Malagasy but right adjunction in Bulgarian. The theory-internal issue is how to capture this parametric variation. It can be stipulated, of course, but this is unsatisfactory and I see no principled explanation. Under the cleft analysis, by contrast, the correct ordering in (27a) is fixed by the phrase structure. The bodyguard must be a subject and it must follow the adjunct wh-predicate.

A second argument for the cleft analysis of multiple wh-questions comes from examples with multiple fronted wh-adjuncts. We have already seen that wh-questions in Malagasy are sharply restricted: only subjects and certain adjuncts may be preposed. When one of each fronts, the subject must be rightmost. Even taking these restrictions into account, if Malagasy is a multiple wh-fronting language, it should still be acceptable to front other combinations of wh-phrases, specifically, two wh-adverbials. This is not possible however, regardless of whether or not a subject bodyguard is present:

(28) a. *taiza oviana no   nividy   mofo Rasoa?
       where when   PRT buy.ACT bread Rasoa
    b. *oviana taiza no   nividy   mofo Rasoa?
       when   where PRT buy.ACT bread Rasoa
    c. *oviana taiza Rasoa no   nividy   mofo?
       when   where Rasoa PRT buy.ACT bread
       ('Where did Rasoa buy bread when?')

By contrast, true multiple wh-fronting languages like Bulgarian, allow multiple adjunct questions:[7]

(29)   koga   ku˅de   jade   Ana
        when   where   eat   Ana
        'When did Ana eat where?'

This lack of parallelism between Malagasy and Bulgarian is surprising if the two are analytically the same kind multiple wh-fronting language. If there is a well-formed derivation for (29), it is unclear why there is not also one for (28). The pseudocleft analysis, in contrast, explains (28). The sentences cannot be generated since the second wh-adverbial can be neither a second wh-predicate nor a bodyguard.

The third problem for the fronting analysis comes from a consideration of the discourse status of the bodyguard when it is not a wh-phrase. Paul 2003a shows that the bodyguard is not part of the focus but, rather, is topical. As such, the bodyguard must be definite, (30), it cannot be the answer to an information question, (31), and it is a preferential position for pronouns.

(30)   tany an-tokotany *(ny)   zazavavy   no   nilalao   baolina   Paul 2003a
        there ACC-yard   the   girl   PRT   play.ACT   ball
        'It was in the yard that the girls were playing ball'

(31) a. iza   no   nanapaka   bozaka   oviana?   Paul 2003a
        who   PRT   cut.ACT   grass   when
        'Who cut grass when?'
     b. #omaly   Rasoa   no   nanapaka   bozaka
        yesterday Rasoa   PRT   cut   grass
        'It was Rasoa who cut grass yesterday'

This observation is not compatible with the fronting analysis in which the bodyguard is part of the focus in spec,C. Under the cleft analysis, the bodyguard is not part of the predicate focus; it is part of the subject topic and has a uniform derivation and discourse status.[8]

Fourth, there is constituency evidence against the claim of the fronting analysis that the two initial elements form a constituent, as schematized in (32a). Data below will support the pseudocleft analysis, in which the wh-phrase and the bodyguard do not form a constituent, (32b).

(32) a. [$_{CP}$ [WH-PHRASE BODYGUARD] [$_{C'}$ no VP]]   FRONTING ANALYSIS
     b. [$_{IP}$ [$_{VP}$ WH-PHRASE] [$_{DP}$ BODYGUARD no VP]]   CLEFT ANALYSIS

First, (33) shows that the predicate plus bodyguard can be interrupted by the matrix clause level parenthetical *hono* 'so they say'.

(33)   oviana   hono   Rasoa/iza   no   nividy   mofo?
        when   PAREN   Rasoa/who   PRT   buy   bread
        'When, do they say, did Rasoa/who buy bread?'

254

The ability of a parenthetical to appear between the wh-phrase and the bodyguard is surprising under the fronting analysis. With the pseudocleft analysis, there is a major constituent break between the predicate and the subject where the parenthetical is and (33) is expected to be grammatical.

Second, coordination facts lead to the same conclusion. The bodyguard and following material can be coordinated to the exclusion of the initial wh-phrase:

(34) oviana Rasoa no nijinja vary ary Rabe no nanapaka bozaka?
     when Rasoa PRT harvest rice and Rabe PRT cut   grass
     'When did Rasoa harvest rice and Rabe cut grass?'

Such examples are straightforwardly accounted for with the pseudocleft structure since the coordinated material is a constituent, as reflected in (32b). Under the fronting analysis however, there is no constituent being coordinated, see (32a), and the result should be ungrammatical, assuming that only constituents can coordinate. Further, the fronting analysis predicts that one should be able to coordinate the wh-phrase+bodyguard to the exclusion of *no* and the remaining material. This is completely impossible:

(35) *[oviana Rabe] sy/ary [taiza Rasoa] no nividy mofo?
      when   Rabe   and   where Rasoa PRT buy   bread
     ('When did Rabe buy bread and where did Rasoa buy bread?')

Third, post-predicate particles treat the initial wh-phrase structurally separate from the bodyguard. When both are present, the particle follows the wh-phrase, not the bodyguard:

(36) oviana (**anie**) Rasoa (*anie) no nanapaka bozaka?
     when   EXCL   Rasoa EXCL   PRT cut   grass
     'When really did Rasoa cut the grass?'

In conclusion, there is good empirical evidence that Paul is correct in proposing the general structures in (24) and (25) for the bodyguard construction and, by extension, multiple wh-questions. The bodyguard is part of the subject constituent and not the predicate. No operation treats the two initial phrases as a constituent, contra the fronting analysis but in line with the pseudocleft analysis.

5   Summary

This paper has argued that wh-questions in Malagasy have a pseudocleft structure in which the initial wh-phrase is the predicate and the remaining material, a DP, is the subject. Despite superficial appearances, wh-questions in Malagasy are not derived by wh-movement. The cleft structure for wh-questions in Malagasy is superior in uniting the analysis of wh-questions and the focus construction, in predicting the position of predicate-related particles in wh-questions, and in providing an analysis of multiple wh-questions and the bodyguard construction.

At this stage, a number of more detailed analytical questions arise. Two in particular are, what is the proper treatment of the particle *no* and why does Malagasy not appear to have any overt A'-movement under the adopted conception of clause structure? Is the latter an accidental observation or artifact of the simple clause structure assumed or is it some deeper property of the Malagasy grammar. I leave these and other questions for future investigation.

Endnotes

1. I use the following abbreviations in glossing: 1/2/3-person, ACC-accusative, ACT-active voice, CIRC-circumstantial voice, EXCL-exclamative, EMPH-emphatic, NEG-negative, PAREN-parenthetical, PASS-passive voice, PRT-particle, SG/PL-number.
2. There is considerable debate as to whether the clause-final DP is a subject or an A' topic-like element. I continue to refer to it as a subject for convenience, without taking a stand on the issue. See Pearson, to appear for discussion.
3. The pre-predicate particles may also appear farther to the right, (i). This is expected under both analyses however because the verb constitutes a predicate under both analyses. Such examples do not help to differentiate the two analyses.

(i) a. iza   no   **tokony** mamangy   an-dRabe
         who   PRT   should   visit.ACT   ACC-Rabe
         'Who should visit Rabe?'

4. One unexpected piece of data in this realm concerns the negative particle *tsy* which also occurs predicate-initial. It occurs before the VP in VOS clauses, (ia), and it may occur before the XP predicate in the focus construction, (ib).

(i) a. **tsy**   nihinana   ny vary   Rasoa
         NEG   eat   the rice   Rasoa
         'Rasoa didn't eat the rice'
     b. **tsy** Rasoa   no   nihinana   ny vary
         NEG   Rasoa   PRT   eat   the rice
         'It isn't Rasoa who ate the rice'

Perhaps surprisingly, negation cannot precede a wh-phrase in a wh-question, (iia). The negation must occur before the verbal predicate, (iib).

(ii) a. ***tsy** iza   no   nihinana   ny vary?
          NEG   who   PRT   eat   the rice
          ('Who didn't eat the rice?')

b. iza   no   **tsy**   nihinana   ny vary?
   who   PRT   NEG   eat   the rice
   'Who didn't eat the rice?

While I don't have a full explanation for the ungrammaticality of (iia), I believe that it is semantic in nature. The corresponding English cleft is also ill-formed: *Who isn't it who ate the rice?* While there does not seem to be anything syntactically wrong with the English cleft, it is nonetheless unacceptable. A full explanation however must await further work.

5. The focus construction shows the same predicate behavior of the initial constituent with respect to particle placement (see Paul 2001). Post-predicate particles immediately follow the focused element and pre-predicate particles immediately precede the initial element. This further illustrates the parallel between the focus construction and wh-questions appealed to in section 4.1.

6. My consultants did not accept such examples. I am grateful to Joachim Sabel and Elisabeth Ravaoarimalala for help with the additional multiple wh-examples.

7. Thanks to Veronica Gerassimova and Maria Jordan for help with the relevant data.

8. Given the topic status of the bodyguard, Sabel's (2003) multiple wh-question examples are perhaps unexpected since the wh-phrase bodyguard does not seem to have topic properties under the given translation. The optional presence of the determiner in (19b) does suggest however that the wh-phrase bodyguard is somehow familiar. I leave for future work whether it is more appropriate to translate the bodyguard with a topic-like, d-linked wh-phrase such as *which person.*

References

Aldridge, Edith. 2002. Nominalization and wh-movement in Seediq and Tagalog. *Language and Linguistics* 3, 393-427.

Bowers, John. 1993. The syntax of predication. *Linguistic Inquiry* 24, 591-656.

Boskovic, Zeljko. 2002. On multiple wh-fronting. *Linguistic Inquiry* 33, 351-383.

Chomsky, Noam. 1995. *The Minimalist Program.* Cambridge, Ma.: MIT Press.

Dahl, Otto Christian. 1986. Focus in Malagasy and Proto-Austronesian. In *FOCAL 1,* 21-45. Pacific Linguistics.

Flegg, Jill Heather. 2003. Topics and clitic left dislocation in Malagasy. Paper presented at the AFLA X, University of Hawaii.

Georgopoulos, Carol. 1991. *Syntactic variables: Resumptive pronouns and A' binding in Palauan.* Dordrecht: Kluwer.

Guilfoyle, Eithne, Henrietta Hung, and Lisa Travis. 1992. Spec of IP and Spec of VP: Two subjects in Austronesian languages. *Natural Language and Linguistic Theory* 10, 375-414.

Keenan, Edward L. 1976. Remarkable subjects in Malagasy. In C. N. Li (ed.). *Subject and Topic.* New York: Academic Press, 247-301.

Keenan, Edward L. 1995. Predicate-argument structure in Malagasy. In C. S. Burgess, K. Dziwirek, and D. Gerdts (eds.). *Grammatical Relations: Theoretical Approaches to Empirical Questions*. Stanford: CSLI, 171-216.

MacLaughlin, Dawn. 1995. Wh-movement in Malagasy: An extraction asymmetry. In A. Akinlabi (ed.). *Theoretical approaches to African linguistics*. Trenton, NJ: Africa World Press, 117-128.

Paul, Ileana. 2000. Malagasy clause Structure. Doctoral dissertation, McGill Univ.

Paul, Ileana. 2001. Concealed pseudo-clefts. *Lingua* 111, 707-727.

Paul, Ileana. 2002. On extraction asymmetries. In *MIT Working Papers in Linguistics 44: The Proceedings of AFLA XIII*. Cambridge, Ma.: MITWPL.

Paul, Ileana. 2003a. Multiple topics: Evidence from Malagasy. In *Cornell Working Papers in Linguistics 19: The Proceedings of AFLA IX*. Ithaca, NY: CLC Publications.

Paul, Ileana. 2003b. On the lack of wh-movement in Malagasy. Paper presented at AFLA X, University of Hawaii.

Pearson, Matthew. 1998. Rightward object shift and the syntax of adverbs. In Ileana Paul (ed.). *UCLA Occasional Papers in Linguistics 20: The Structure of Malagasy II*. Los Angeles: UCLA Department of Linguistics, 34-49.

Pearson, Matthew. 2001. *The Clause Structure of Malagasy: A Minimalist Approach*. Doctoral dissertation, UCLA.

Pearson, to appear. The Malagasy subject/topic as an A'-element. *Natural Language and Linguistic Theory*.

Pesetsky, David. 2000. *Phrasal movement and its kin*. Cambridge, Ma.: MIT Press.

Rabenilaina, Roger-Bruno. 1998. Voice and diathesis in Malagasy: An overview. In *UCLA Occasional Papers in Linguistics 20: The Structure of Malagasy II*, ed. by Ileana Paul, 2-10. Los Angeles: UCLA Department of Linguistics.

Rackowski, Andrea. 1998. Malagasy adverbs. In Ileana Paul (ed.). *UCLA Occasional Papers in Linguistics 20: The Structure of Malagasy II*. Los Angeles: UCLA Department of Linguistics, 11-33.

Richards, Norvin. 2002. *Movement in Language*. Oxford: Oxford University Press.

Rudin, Catherine. 1988. On multiple wh-questions and multiple wh-fronting. *Natural Language and Linguistic Theory* 6, 445-501.

Sabel, Joachim. 2002. Wh-questions and extraction asymmetries in Malagasy. *MIT Working Papers in Linguistics 44: The Proceedings of the Eighth Austronesian Formal Linguistics Association*. Cambridge, Ma.: MITWPL.

Sabel, Joachim. 2003. Malagasy as an optional multiple wh-fronting language. In *Multiple wh-fronting*, ed. by Cedric Boeckx and Kleanthes Grohmann, 229-254. Amsterdam: John Benjamins.

Eric Potsdam
University of Florida
P.O. Box 115454
Gainesville, Florida 32611
United States
potsdam@ufl.edu

Two Systems of *Remnant Movement II* and Extraction from Specifier Position

Craig Thiersch
University of Potsdam and University of Tilburg

1. Introduction

In the wake of Kayne's Antisymmetry Hypothesis and Linear Correspondence Axiom (LCA), there has been much fruitful research attempting to adjust syntactic analyses to those permitted by Kayne's restrictive system. In doing so, analyses which at first seem counter-intuitive may turn out to provide solutions to old problems. Two cases in point are the analysis of Malagasy involving extensive *Remnant Movement* [henceforth RM] described in Rackowski & Travis (2000), Pearson (2001), and elsewhere; on the one hand, and the analysis of Hungarian and Dutch verbal clusters in Koopman & Szabolcsi (2000) [henceforth R&T, Pearson, and K&Sz].[1]

The original motivation (in part) for examining L&Sz and subsequently R&T was that it is the extensive use of iterated RM which increases the computational complexity of languages generatable in Stabler's "Strict Minimalist Grammar" formalism over that of context-free grammars. It has also been noted that allowing extraction from complex specifiers created by Merge (as opposed to Move) increases the level of complexity even further (Jens Michaelis, *p.c.*). Both R&T and K&Sz make extensive use of RM; R&T allow extraction from complex specifiers, while K&Sz do not. Although the specifiers in both cases are created by Move, not Merge, we nevertheless feel that there is enough intrinsic linguistic interest in trying to limit extraction possibilities to pursue the comparison of these two systems with regard to this point. Cf. Thiersch (*in prep*.b) for discussion.[2]

This is intuitively plausible, as we know extraction from deeply embedded position within specifiers is usually ungrammatical:

(1)   a.   Who$_i$ did Fred say that Susan hoped for Sam to kiss e_i ?

          b.   * Who$_i$ did Fred say that for Sam to kiss e_i would create a scandal ?

As noted above extraction from complex specifiers is explicitly forbidden in K&Sz and the prohibition forms a crucial part of their analysis. It appears however to be necessary for R&T. In this article we compare the two systems in the hope of shedding some light of how two rather different systems involving RM have coped with the same problem: the position of nominal objects.

As a terminological matter we distinguish between two types of RM, I and II, following Müller (2002) and Thiersch (2002):

(2) RM I

        a.   Erschossen hat er Bin Ladin schon gestern;

        b.   Analysis: [$_{VP}$ e_i erschossen]$_j$ hat er [Bin Ladin]$_i$ schon gestern e_j

(3) RM II

        a.   John reads no novels

        b.   John [reads e_i]$_j$ [$_\alpha$ [no novels]$_i$ e_j]

Although formally identical, in that they involve displacing a constituent α from which some element β has been extracted so that the antecedent, β, no longer C-commands its trace in α, they tend to have different characteristics. For example, in RM I both operations exist as independent constructions in the language (in (2) "scrambling" and Vorfeld-topicalization), and it cannot be iterated; RM II generally can apply repeatedly, and as (3b) shows, neither operation occurs independently, and its effects are often invisible — except for semantic effects such as the two readings of (4a); cf. Kayne (1998).[3]

(4) a. They forced him to marry no one

    b. They forced him [[to marry e_i]$_j$ [no one]$_i$ e_j]

    c. They [forced him to marry e_i]$_j$ [no one]$_i$ e_j

Before we examine the problem relating to the position of nominal DPs, we briefly sketch some relevant aspects of Malagasy syntax and three sample empirical problems the RM analysis apparently solves. For a more detailed account of Malagasy syntax the reader is referred to Pearson (2001) and the references therein.

2. The basics: the Malagasy voicing system

Roughly, Malagasy is a verb initial language (although there are constructions in which certain "fronted" constituents may precede the verb). It is sometimes cited as being typologically VOS, but this oversimplifies what is really going on. Basically the arguments line up following the verb, and depending on the voice of the verb, one of them is "promoted" to the right most position. Schematically,

(5) V_{voice2} arg$_1$ (arg$_2$) arg$_3$ arg$_2$

That is, if the verb is in "second" voice, the second argument appears to the right.[4] The five voices are traditionally called Nominative, Accusative, Dative, Transitive, Circumstantial). Cf. Pearson (2001), Chap.2 [his (65)]:

(6) a. Mamono akoho amin'ny antsy ny mpamboly
        NomP.kill chicken with-Det knife Det farmer
        'The farmer kills chickens with the knife'

    b. Vonoin'ny mpamboly amin'ny antsy ny akoho
        AccP.kill-Det farmer with-Det knife Det chicken
        'The chickens are killed by the farmer with the knife'
        or 'The chickens, the farmer is killing (them) with the knife'

    c. Amonoan'ny mpamboly akoho ny antsy
        CrcP.kill-Det farmer chicken Det knife
        'The knife is being used by the farmer to kill chickens'
        or 'The knife, the farmer is killing chickens (with it)'

Indeed, the voicing system has sometimes been compared with Indo-European passive, and the right-most argument called the subject, although the construction has rather different properties, as can be seen in (6b), the non-promoted semantic subject does not become an optional *chômer* in a PP, but remains, presumably *in situ*.[5] Earlier analyses often assumed that the structure was "right

260

branching"; cf. Guilfoyle, Hung & Travis (1992). That is, the external argument moved up to a right-hand specifier position. Pearson (and R&T) argue that this is not the case, but that the basic structure is consistantly Spec-Head-Comp, as in Kayne's antisymmetry proposal, and that the external argument has moved up to the *left*, and a constituent containing the verb with the remaining arguments (and some adjuncts) has then moved leftward around the externalized argument. Schematically, the sentence in (6b) would have a structure roughly like

(7) [Vonoin'ny   mpamboly e_i amin'ny antsy]$_j$ [ny akoho]$_i$ e_j
     AccP.kill-Det farmer   *e.c.* with-Det knife   Det chicken *e.c.*

This is grossly over-simplified and ignores the functional categories as well as the considerable difference in hypothesized nodes between, say, Pearson and R&T, but captures the spirit of the analyses under consideration. Pearson in particular motivates a number of functional categories and hence movements, giving a more complete and differentiated account of the facts.

We briefly summarize in the following sections how this type of analysis accounts for some empirical observations, indicating that the RM approach is on the right track, before turning to the problem of placing the DP arguments.

3. Some empirical problems solved

Focus "extraction": Various focus and topic particles can allow an element to be "extracted" to the left, but as has been long noted in the literature, this can only be the constituent which would have been "externalized" to the rightmost position according to the voice of the verb.[6] E.g., [Pearson, Chap.2 (36)]:

(8) a. * Ny akoho no *namono*   tamin'ny   antsy ny
         Det chicken Foc Pst-NomP.kill Pst-with-Det knife Det
         mpamboly
         farmer
         'It's the chicken that the farmer killed with the knife'

     b.   Ny akoho no *novonoin*'ny   mpamboly tamin'ny
         Det chicken Foc Pst-AccP.kill-Det farmer   Pst-with-Det
         antsy
         knife

     c. * Ny akoho no *namonoan*'ny   mpamboly ny antsy
         Det chicken Foc Pst-CrcP.kill-Det farmer   Det knife

This is surprising under the Guilfoyle, Hung & Travis analysis, as one generally expects objects to be more easily extractible than subjects; cf. Sabel (2003) for discussion and an alternative to the RM analysis.[7] In a structure like (7), however, the external object is on the main rightward projection line with no barrier-like categories in inbetween, whereas the other arguments are embedded in the fronted TP/VP, i.e. a complex specifier and presumably an island (a point to which we return below).

Placement of discourse markers: Pearson discusses another phenomenon (noted in passing in R&T), namely that various discourse markers come in peculiar places and don't seem to have any particular pattern. For example, the yes/no

question marker *ve* appears in the penultimate position (left of the external-ized argument) in ordinary sentences, but in the second position (between the fronted DP and the particle *no!*) in focus sentences [his Chap.2: (14a) & Chap.4: (102a)]:

(9) a. Vonoin'ny   mpamboly amin'ny antsy ve ny akoho?
         AccP.kill-Det farmer   with-Det knife Qu Det chicken
         'The chickens, is the farmer killing (them) with the knife?'

     b. I   Bakoly ve no manapaka bozaka?
         Det Bakoly Qu Foc NomP.cut grass
         'Is it Bakoly who is cutting the grass?'

Pearson points out that we only need to assume that *ve* is a second-position functional head which raises the specifier of its complement in order to ac-count for these facts; in (9a) the whole TP/IP is raised, leaving the externalized argument to the right of *ve*. in (9b), however, only the focused DP is raised (see the structure in footnote 7).

Adverb order. Finally we come to one of the phenomena which forms the centerpiece of the R&T article, namely the order of adverbials. The basic as-sumption is the universal order of adverbials related to a hypothesized univer-sal heirarchy of functional projections proposed in Cinque (1999). Language specific deviations were to be accounted for, as usual, by language particu-lar properties (e.g., lexical). As previously discussed with a similar analysis in Rackowski (1998), the order of adverbials in Malagasy deviates from the proposed universal order in a surprising way; aside from some minor peculiar-ities, the pre-verbal adverbs mimic the Cinque order, whereas the post-verbal adverbs are in the mirror-image order:

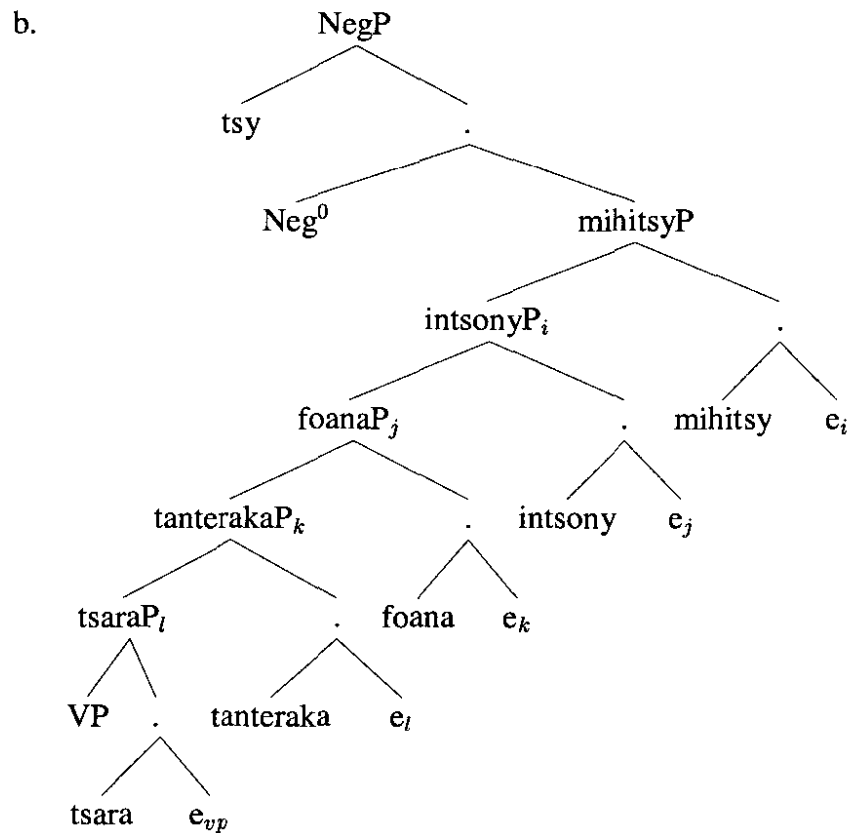(10) a. Cinque's order:

     1   2   3   4   5   6
     (speech act) > Generally > Neg > Already > Still > (at-all) >
     7   8   9   10
     Anymore > Always > Completely > Well

     b. Malagasy order:

     ?   2   3   4   5   (3)   V
     Na(dia)   > Matetika   > Tsy > Efa   > Mbola > Tsy > (Verb) >
     Even   > generally > Neg > Already   > Still   > Neg > Verb   >
     10   9   8   7   6   ?
     Tsara   > Tanteraka   > Foana > Intsony   > Mihitsy > Aza   >
     Well   > Completely > Always > Anymore   > At-all   > though >
     1
     Ve
     Speech Act

This immediately suggests a "roll-up" operation like the operation proposed in K&Sz to reverse the order of Hungarian verbs. The language particular stipula-tion they need to make is that the upper adverbs (up to and including Neg "tsy") are generated in Spec of their functional projections; the lower adverbs (6–10) generated as heads of their projections[8] Their suggested derivation works as follows:

(11) a. Repeated movement to Spec,Adv_xP (roll-up) reverses the order,

b. a blocking trigger (e.g., an element in Spec,NegP) stops the roll-up, and

c. a high resulting constituent (e.g., TP) is moved to a projection above the external argument (subject).

The following tree illustrates (11a–b), step (11c) was illustrated (for a sentence without adverbs) in (7) [cf. their (12)–(13), p.122]:

(12) a. Tsy manasa tsara tanteraka foana intsony ny lamba mihitsy
not wash wel completely always anymore the clothes at-all
Rakoto
Rakoto

b.

```
                          NegP
                         /    \
                      tsy       .
                             /      \
                        Neg^0        mihitsyP
                                    /        \
                            intsonyP_i          .
                           /        \          / \
                      foanaP_j        .    mihitsy  e_i
                     /       \        / \
               tanterakaP_k    .  intsony  e_j
              /         \      / \
         tsaraP_l        .  foana  e_k
         /    \          / \
       VP      . tanteraka  e_l
      /  \
   tsara  e_vp
```

Note that the tree in (12b) already involves repeated RM, even without DP arguments. While one can quibble with many of the details and assumptions, the capturing of these three phenomena (other arguments are given in Pearson and R&T) seems sufficient motivation that the RM approach is on the right track, and we turn to the details of DP placement.

4. The order of DP objects

The problem is illustrated by two of their examples. Definite direct objects may "optionally appear among or after postverbal adverbials." [p.125] This is illustrated in (13) [their (21a/b)][9]:

(13) a. Tsy manasa foana ny lamba mihitsy Rakoto.
NEG PRES.AT.wash always det cloths at-all R.
'Rakoto does not always wash the clothes at all'

263

b. Tsy manasa   intsony mihitsy ny lamba Rakoto.
   NEG PRES.AT.wash no-longer at-all   det cloths R.
   'Rakoto does not wash the clothes at all any more'

The adverbs *mihitsy* and *foana* are respectively Adv6 and Adv8 in the Cinque heirarchy, here appearing in 8 > 6 order due to the (remnant) movement of the verb phrase. How would the derivation procede? The first merge would be with the direct object, yielding

(14) [$_{VP}$ manasa [$_{DP}$ ny lamba]]

This is merged with the lower adverb, *foana*, and the VP moves to the specifier of *foana*, yielding

(15) [$_{Adv8P}$ [$_{VP_i}$ manasa [$_{DP}$ ny lamba]] foana e_i]

If we simply merge again and move the Adv8P out, we get

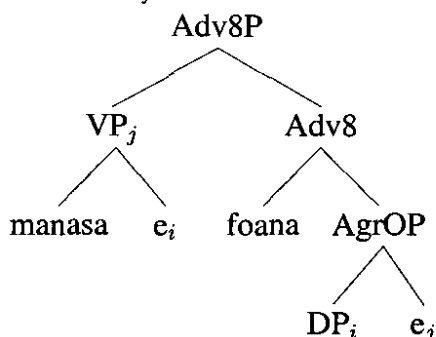(16) [$_{Adv6P}$ [$_{Adv8P_j}$ [$_{VP_i}$ manasa$_V$ [$_{DP}$ ny lamba]] foana$_{Adv8}$ e_i] mihitsy$_{Adv6}$ e_j]

This generates neither (13a) or (13b), but is grammatical.[10] Supposing we had immediately extracted the DP to Spec,AgrOP yielding

(17) [$_{AgrOP}$ [$_{DP_i}$ ny lamba] [$_{AgrO'}$ AgrO^0 [$_{VP}$ manasa e_i]]]]

Merging with Adv6P and raising the remnant gives (18a) shown as tree in (18b):

(18) a. [$_{Adv8P}$ [$_{VP_j}$ manasa e_i] [foana$_{Adv8}$ [$_{AgrOP}$ [$_{DP_i}$ ny lamba] e_j]]]
      b.   Adv8P
                 /   \
              VP$_j$ Adv8
             /   \   /   \
          manasa e_i foana AgrOP
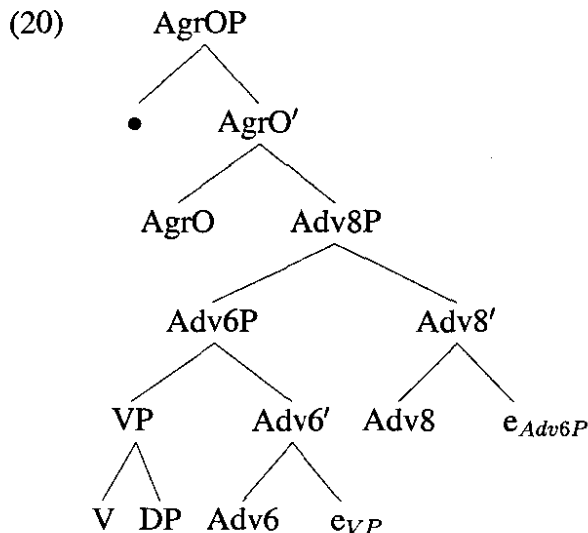   /   \
                              DP$_i$ e_j

And after merging with Adv6 and raising Adv8P, we get the structure for (13a):

(19) [$_{Adv6P}$ [$_{Adv8P_k}$ [$_{VP_j}$ manasa e_i] [foana$_{Adv8}$ [$_{AgrOP}$ [$_{DP_i}$ ny lamba] e_j]]] mihitsy$_{Adv6}$ e_k]
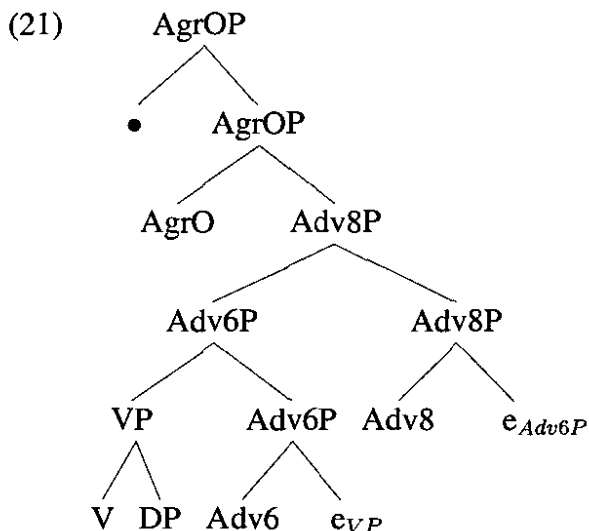
What about (13b)? Here there at least two possibilities, each with attendant problems. One possibility is to extract the DP immediately, as in the derivation for (13a). But then we would need to extract it again to get (13b). However, we have already "used up" AgrOP and checked the appropriate feature. This means we would have to postulate another functional category. This has the disadvantage of needing to postulate a potentially unlimited number of functional categories (and features) just to effect the extraction. (This is in fact the solution adopted by K&Sz; see below.)

The alternative, adopted by R&T, is to take the DP along, as in (16), and wait until the appropriate moment in the derivation, *then* merge once with AgrOP and extract the DP:

(20) AgrOP

```
          AgrOP
         /    \
        •    AgrO'
            /    \
        AgrO    Adv8P
               /      \
           Adv6P      Adv8'
          /    \      /    \
        VP    Adv6' Adv8  e_{Adv6P}
       /  \   /   \
      V  DP Adv6  e_{VP}
```

Example (20), however, presents us with a potential problem: we have here extraction from a complex specifier, and given a long series of adverbs, the DP could be indefinitely deeply embedded. Suppose we permit unlimited extraction from complex specifiers. We are then in a quandry with respect to the focus extraction: the explanation for the inaccessibility of the DPs for extraction, except for the externalized one, was that they were in a specifier island. In the article R&T do invoke the island-hood of the fronted XP: "The predicate-fronting analysis predicts that, since the subject is the only argument not contained in some kind of island, it should behave differently from the other arguments." [p.124] How could we differentiate between the two cases? (Pearson has a different explanation, see below).
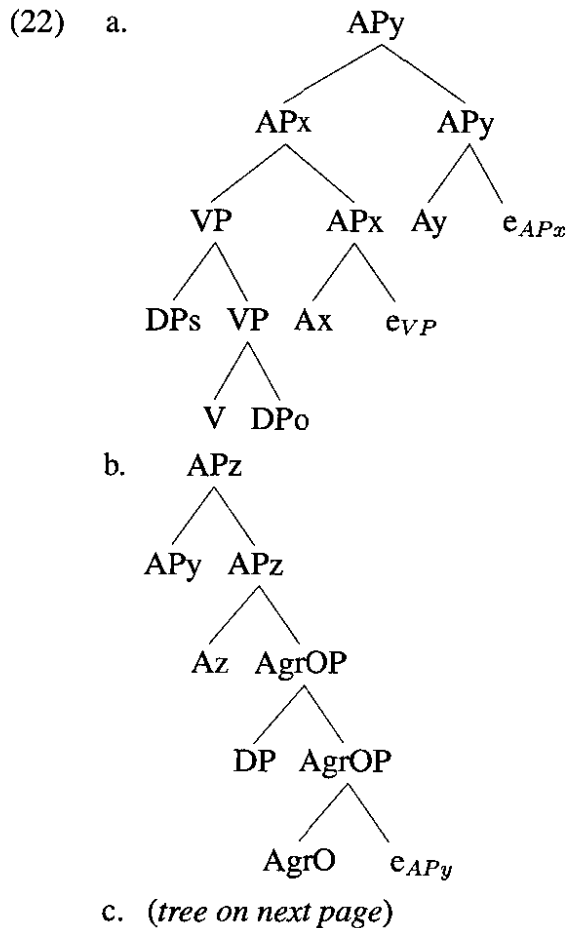
A potential solution comes from the former version of this analysis, namely the version in Rackowski (1998); here she adopts a strict Kaynean phrase structure with only one specifier per projection and differenties between *categories* and *segments*. Under these assumptions her structure for (20) was

(21) AgrOP

```
          AgrOP
         /    \
        •    AgrOP
            /    \
        AgrO    Adv8P
               /      \
           Adv6P      Adv8P
          /    \      /    \
        VP    Adv6P Adv8  e_{Adv6P}
       /  \   /   \
      V  DP Adv6  e_{VP}
```

In (21) the only *category* which dominates DP is VP; the rest of the nodes are only segments. In traditional terms (i.e., *Barriers*), as Rackowski notes, there is nothing to block the properly governed DP from being extracted.
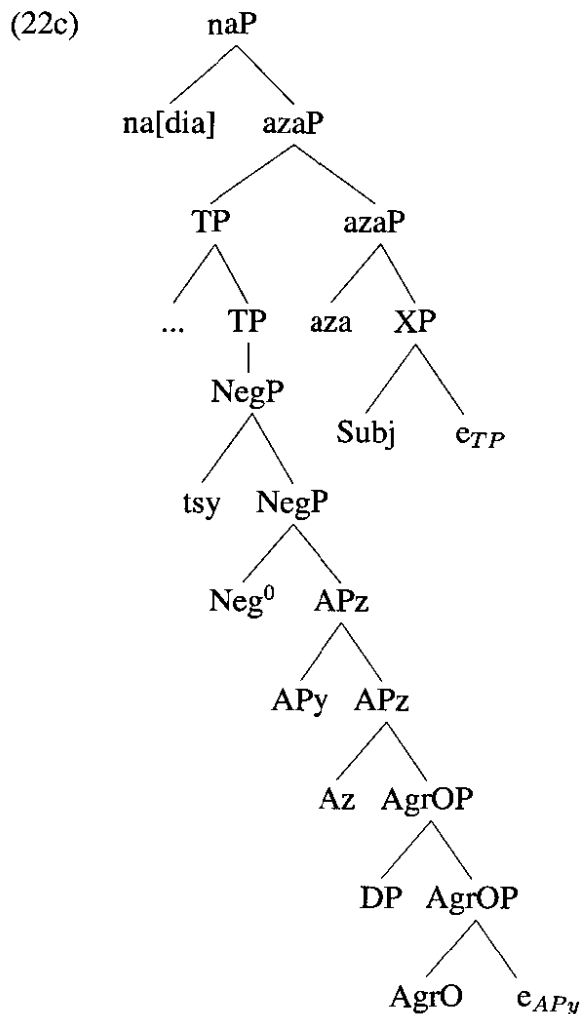
We then need to ask if this explanation is sufficient to nevertheless block extraction in, say, the focus or topic constructions. This would seem to be

the case. Let us look at a "complete" derivation, much simplified. Note the position of DP in the Spec under VP (22a) $\approx$ (21) from which it moves to AgrOP (22b). If TP subsequently moves (22c), the DP is then dominated by the *category* AP_z and all cateogries above AP_z, freezing it in place (*na (dia)...aza* translates roughly as 'even...though'):

(22) a.

APy
├── APx
│   ├── VP
│   │   ├── DPs
│   │   └── VP
│   │   ├── V
│   │   └── DPo
│   └── APx
│   ├── Ax
│   └── e$_{VP}$
└── APy
    ├── Ay
    └── e$_{APx}$

b.   APz
├── APy
└── APz
    ├── Az
    └── AgrOP
        ├── DP
        └── AgrOP
            ├── AgrO
            └── e$_{APy}$

c.   (*tree on next page*)

While this is a very neat solution, it depends on the segment/category distinction, which does not seem to be invoked in later work by Travis (2004), which uses the traditional X-bar theory. It also depends upon *Barriers*-style extraction restrictions and it remains to be seen whether we can translate this into a "Phase" explanation; see Thiersch (*in prep*.b).

We note that this approach is explicity rejected by Pearson, who claims to need multiple specifiers in his analysis, redefining C-command derivationally along the lines of Epstein *et al.* (1998) to achieve this. He links the extraction constraint to feature conflict similarly to the exclusion of simultaneous WH and Topic movement in Germanic; cf. discussion at the end of § 3.4.1, p.133.

(22c)   naP
              /   \
         na[dia]   azaP
                   /   \
               TP   azaP
              / \   / \
            ... TP   aza   XP
                  | / \
                NegP   Subj e_{TP}
                /   \
             tsy   NegP
                   /   \
               Neg^0 APz
                        /   \
                     APy   APz
                            /   \
                          Az   AgrOP
                                /   \
                              DP   AgrOP
                                     /   \
                                  AgrO e_{APy}

5. The K&Sz system

While it is beyond the scope of this preliminary article to discuss the intricate system developed in K&Sz for Hungarian and Dutch verb clusters (cf. Thiersch *in prep*.a), we would like to look briefly at their solution to the above problem, with an eye to a comparison of the issues involved. They are of course faced with the same problem as R&T, namely the DP arguments associated with verbs have to be "gotten out" of the remnant constituents being rolled up. Indeed, since the data (verbal clusters) treated in K&Sz disregards adverbial constituents, the majority of the remnant character of their roll-up comes from repeated evacuation of an argument DP from a VP.

First of all, K&Sz deal with extraction from complex specifiers by simply forbidding it by fiat.[11]

Glossing over many details (cf. Thiersch (*in prep*.a) the core of their system is as follows: although with five verbs the Hungarian verb cluster could theoretically have 5!=120 orders, in fact it has only four; e.g. with one (finite) true auxiliary, 2 semi-auxiliaries, a main verb and a VM (particle):

(23) a. 1-2-3-5-4   c. 1-5-4-3-2
      b. 1-2-5-4-3   d. 5-1-2-3-4

The verbs are base generated in the "English" (1-2-3-4-5) order following Kayne, the first pair (V+VM) must be inverted, and if the inversion stops, it cannot restart. Alternatively, the VM may raise (23d). Since they assume (i)

only XP movement is involved, (ii) obligatory feature checking and (iii) no LF movement or counter-cyclic movement, the surface "English" order (23a) must also be derived. How do they prevent this from being circular, and how do they get the alternative orders with obligatory checking? The answer is optional pied-piping. We know pied-piping as a surface phenomenon is optional, so that if we encounter a structure like (24) [their (72)=(79)] where the VP$^+$ needs to move to Spec,CP (•) to check a feature, it cannot (due to the prohibition of extraction from Spec), so we can choose either InflP$^+$ or PredP, pied-piping the rest.

(24)   (CP)

        /\
   (•)   ...
       /\
     .   [PredP]
          /\
    [InflP$^+$] LP[dp]
     /\   /\
   VP$^+$ ... ...
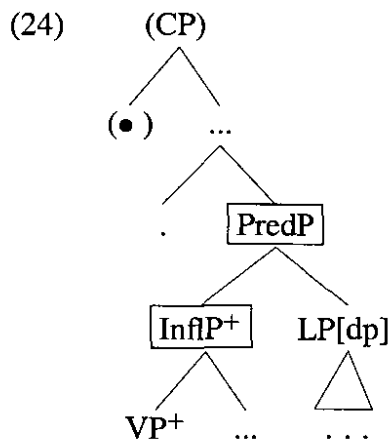
(An XP$^+$ is an extended projections based on Koopmans *Generalized Doubly Filled Comp* filter; cf. Thiersch (*in prep.*a).) This will eventually yield divergent derivations, giving the two orders.

Clearly this will run into problems if a projection contains a DP, just as in R&T, and due to their prohibition on extraction from Spec, they do not have the escape-hatch of R&T regardless of how we interpret their X-bar structure. Their solution is twofold:
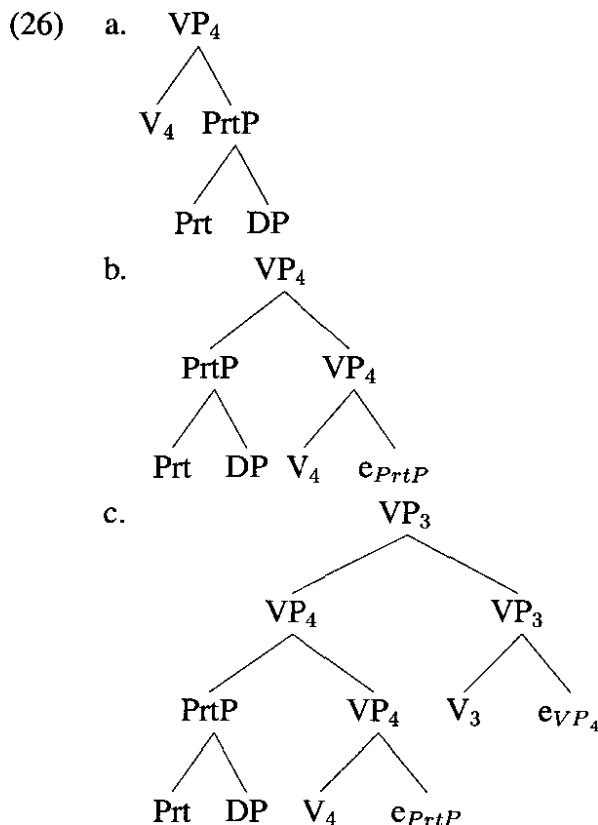
(25) a. They forbid "moving two constituents at the same time", by which they mean that in a structure like [$_{XP}$ ... YP ... ZP ...], XP counts as two constituents under "certain circumstances"; i.a., when they are different projections, e.g., a VP containing a DP, and phonetically realized, but *not* when XP is the extended projection of, say, ZP.

     b. They allow an arbitrary number of functional "pushing categories" (LPs, or licensing phrases) similar to AgrO, but with vacuous content.

Hence the DP *must* leave the VP at each step, due to (25a), and *can* leave due to (25b). (They note that this is not an entirely satisfactory solution.)

Of course we have not done justice to the complexity and subtlety of their system; but although they are able to generate the correct orders in Hungarian and parameterize the system for Dutch, there are numerous problems which they themselves note, and some they don't — e.g., many multiple derivations for the same string. Cf. Thiersch (*in prep.*a). The question arises as to whether they could relax the prohibition on extraction from specifiers, and use the Kaynean phrase structure to allow the DP to extract. Using very simplified structures we demonstrate that this seems unlikely.
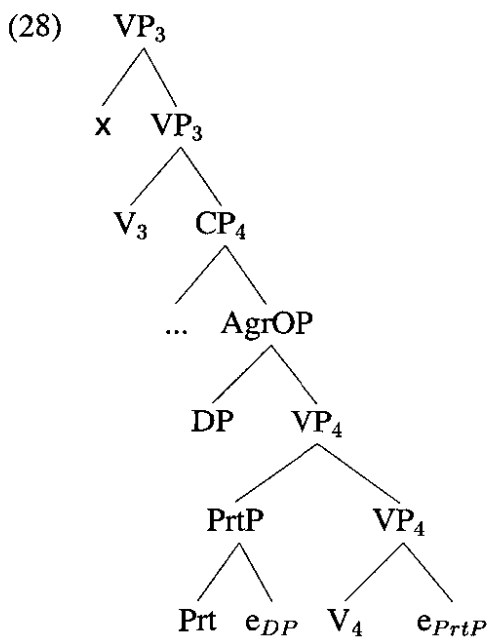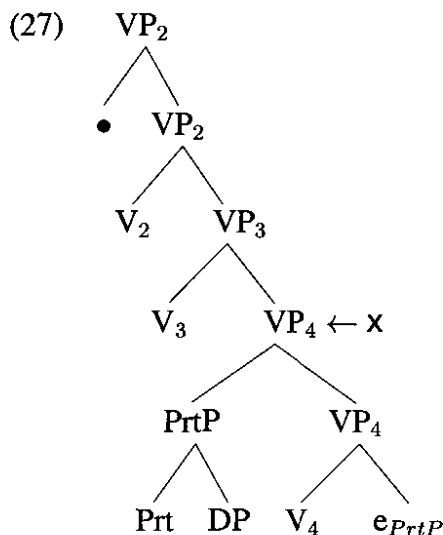
268

Supposed we start with the base structure V Prt DP, e.g., "take apart the radio+ACC" in (26a)[12] and move the PrtP (in their terminology VM-phrase) to the Spec,VP$_4$ as in (26b). We now have two basic options: if we assume (as opposed to K&Sz), that a V can select a VP complement (or the infinitival functional projection thereof), we get a structure like (26c).

(26) a. VP$_4$

```
        VP4
        / \
      V4   PrtP
           / \
         Prt  DP
```

b.

```
            VP4
           /    \
        PrtP     VP4
        / \      / \
      Prt DP   V4  e_PrtP
```

c.

```
                    VP3
                  /     \
               VP4       VP3
              /    \     / \
           PrtP    VP4  V3  e_VP4
           / \     / \
         Prt DP  V4  e_PrtP
```

Here we see a structure akin to R&T, and we observe that DP *can* leave later (it is not too deeply embedded) and the DP *must* leave at some time (because of checking features). (It is not clear what prevents it from leaving too early and becoming embedded in the verb-cluster, which is not permitted in Hungarian, as opposed to West Flemish and Swiss German, for example.) However, we are now faced with a different problem, namely what prevents the inverted cluster from unraveling? In K&Sz, where the V selects an infinitival CP complement[13] the inverted cluster becomes too deeply embedded in a specifier and cannot be undone; in (26c) nothing prevents the *category* VP$_4$ from moving up to the specifier of a V$_2$, as in (27) [next page] for example, eventually generating the ungrammatical order 1+5+4+2+3.

Similarly, what *stops* the roll-up? Roll-up in R&T is stopped at a particular point, i.e., when the next specifier is filled. In K&S it was stopped at any point where the verbs were in separate CPs and hence couldn't be moved together due to the "Move Only One Constituent" condition. But in (27) there is not enough structure to stop roll-up. Suppose we do allow the V to select a full CP, as K&Sz do; then we have a structure like (28). But here if the CP moves up to × (Spec,VP3) to create the inversion, the DP is now embedded to deeply in terms of *categories* to be extracted. Hence we are faced with a contradiction: we can't freeze verb complexes without freezing the DP.

(27) VP$_2$

```
        VP_2
       /  \
      •    VP_2
          /  \
        V_2   VP_3
             /   \
           V_3    VP_4 ← x
                 /    \
               PrtP    VP_4
              /  \     /   \
            Prt  DP  V_4   e_PrtP
```

(28) VP$_3$

```
        VP_3
       /  \
      x    VP_3
          /  \
        V_3   CP_4
             /   \
           ...   AgrOP
                /    \
              DP     VP_4
                    /    \
                  PrtP    VP_4
                 /  \    /   \
               Prt e_DP V_4  e_PrtP
```

This hopefully reveals at least the tip of the iceberg regarding the extraction problem by comparing the technical details of these two systems: when invoking iterated RM the placement of arguments requires either arbitrarily deep extraction from a complex specifier or some system of local extraction; trying to use the R&T system (albeit modified) in Hungarian seems unlikely to succeed. In the extended version of this article we discuss the problems with the reverse approach, applying a K&Sz-like system to Malagasy, as well as looking at the K&Sz system in more detail.

Endnotes

1. The analyses of R&T and Pearson are similar but differ in certain aspects. We return to this below.

2. A possible example of extraction from a merged specifier could be sentences like those in Corver (1990) [his (152d), p.241]:

(i) [Hoeveel dagen]$_i$ is deze melk [[e$_i$ langer] houdbar dan die
    how-many days   is this   milk $e.c.$ longer keepable than that
    yogurt]?
    yogurt

There is some evidence that this structure is incorrect, and this only involves extraction of a simple specifier; cf. recent work by Corver discussed in Thiersch (*in prep.b*).

3. The criteria distinguishing the two are not absolute; e.g., due to the constructions he is investigating, Kayne doesn't iterate as extensively as K&Sz; on the other hand, Dutch also has RM I, and Hungarian possibly as well [(i)b is K&Sz's (19), p.39]:

(i) a. [e$_j$ geslagen]$_i$ heeft hij [de hond]$_j$ e$_i$
         beaten   has he the dog

     b. [Mutogatni e$_i$ e$_j$]$_k$ fogja   akarni   [a játékot]$_i$ [a
         show-INF   will-3SG want-INF the toy-ACC the
         gyerekeknek]$_j$ e$_k$
         children-DAT
         'He will want to show the toys to the children'

Note however that *mutogatni* in (28b) is in a preverbal position normally occupied by the VM (verbal marker, like Dutch and German particles), and there is some controversy as to whether the structure shown (K&Sz's) is correct or whether it is a case of head-movement. Hungarian lacks, as opposed to Dutch and German, the ability to displace the VP as a whole:

(ii) a. Den   Kindern die   Spielzeuge zeigen will   er nicht.
         the-DAT children the-ACC toys   show wants he not

     b. * [$_{VP}$ Mutogatni a játékot a gyerekeknek] nem fogja akarni
           show   the toys the children   not will want

Hence although it has leftward scrambling/displacement of DPs, it lacks at least one of the operations to qualify for RM-I.

4. Note this is not deterministic for some "oblique" voices. See discussion in Al-Gerull (2004).

5. Pearson notes that the voice markers are really like Case, rather than being θ-related, and are idiosyncratically governed by the verb. For example, one has to learn that *write* "governs" the dative (as opposed to IE accusative) [his (24)]:

(i) a. Nanoratra   ny taratasy ny mpianatra
         Pst-NomP.write Det letter   Det student
         'The student was writing the letter'

271

b. Nosoratan'ny   mpianatra ny taratasy
Pst-DatP.write-Det student   Det letter
'The student wrote the letter'

6. We ignore certain problems with Circumlocutional voice; again, cf. Al-Gerull (2004).

7. Pearson, following Paul (1999) analyses them as clefts involving an empty operator [his Ch.3 (122b)]:

(i) $[_{PredP}$ Wh/Foc$_i$] $[_{WhP}$ Op$_i$ *no* $[_{PivP}$ t$_i$ $[_{TP}$ V ... t$_i$ ...]]]$_i$.

8. We take this somewhat controversial assumption as a given for the purposes of this article, as our interest lies in the problems associated with extraction from specifiers. NB: *Pre*-verbal adverbs can sometimes occur in variable positions with different meaning due to scope: e.g., "*Tsy mbola* corresponds to 'not yet', while *mbola tsy* means 'still not'." Rackowski (1998), p.10. They can also appear *after* the external argument ("Nachfeld").

9. Their translation of (13a) seems a bit odd semantically. The example (13b), their (21b), seems more natural, but doesn't illustrate the relevant word order. I assume substituting *instony* for *foana* in (13a) in the ensuing discussion will not affect the grammaticality, since the order of the adverbials in the Cinque heirarchy is preserved.

10. L. Travis, p.c. Note the movement of DP to AgrOP is presumably obligatory, so it would have to be extracted anyway as in the problematical example discussed below.

11. "Regarding moveable constituents, we assume that the *complement of X, or the specifier of the *complement of X, may move to [Spec,XP]. *Complement is the transitive closure of the complement relation. In other words, only full specifiers and full complements on their own projection line can extract; parts of specifiers cannot." [p.38]

12. Not the surface order, of course. Note that their structure, $[_{VP}$ take $[_{PP}$ apart] $[_{DP}$ radio]], is impossible given their assumptions, as it is trinary branching. I've plausibly assumed that the DP is an argument of the Prt.

13. Ignoring the yet more complicated structure suggested in their Chap.6.

References

Ala-Gerull, Hanitry Ny (2004) "Das zirkustanzielle Voice" Lecture and hand-out, Feb. 10, 2004, ZAS, Berlin.

Cinque, Gugliemo (1999) *Adverbs and functional heads; a cross linguistic perspective*. NY: Oxford U. Press

Corver, Norbert (1990) *The syntax of left branch extractions*. Ph.D. Diss., U.Tilburg

Epstein, Samuel D., Erich M. Groat, Ruriko Kawashima & Hisatsugu Kitahara (1998) *A derivational approach to syntactic relations*. NY: Oxford Univ. Press

Guilfoyle, Eithne, Henrietta Hung & Lisa Travis (1992) "Spec of IP and spec of VP: two subjects in Austronesian languages" *Natural Language and Liniguitic Theory* 10: p.375–414.

272

Kayne, Richard (1994) *The antisymmetry of syntax.* Cambridge, MA: MIT Press

—— (1998) "Overt vs. Covert Movements." *Syntax* 1:128-191.

Koopman, Hilda (1996) "The Spec Head configuration" in Garnet, E., & Lee, F., (eds.) *Syntax at sunset* (UCLA working papers in syntax and semantics 1, Dept. of Ling., UCLA) Reprinted in Koopman (2000) *The syntax of specifiers and heads* London: Routledge.

Koopman, Hilda & Anna Szabocsi (2000) *Verbal complexes.* Current studies in linguistics 34. Cambridge, MA: MIT Press

Müller, Gereon (2002) "Two types of remnant movement" in *Dimensions of Movement. From Features to Remnants,* Artemis Alexiadou, Elena Aganstopoulou, Sjef Barbiers & Hans-Martin Gärtner (eds). (Proceedings of the 1999 Potsdam Workshop on Remnant Movement.) Amsterdam: John Benjamins.

Paul, Ileana (1999) *Malagasy clause structure.* Ph.D. diss., McGill Univ.

Pearson, Matt (2001) *The clause structure of Malagasy: a minimalist approach.* Ph.D. diss. Dept. of Linguistics, UCLA

Rackowski, Andrea (1998) "Malagasy Adverbs" in I. Paul (ed.) *The structure of Malagasy,* v.II. UCLA Working Papers in Linguistics.

Rackowski, Andrea & Lisa Travis (2000) "V-initial languages: X or XP movement and adverbial placement" In *The syntax of verb-initial languages,* Andrew Carnie & Ethne Guilefoyle (eds.), Oxford UP, p.117-141.

Sabel, Joachim (2003) "*Wh*-questions and extraction asymmetries" in *Proceedings of the 8th Conference of the Austronesian Formal Linguistics Association (AFLA VIII),* Norvin Richards & Andrea Rackowski (eds.), MIT Working Papers in Linguistics 44, Dept. of Linguistics, MIT, p.304–319.

Thiersch, Craig (2002) "Remnant Movement: Interim Report 1", DFG Projekt STA 519/1-1.

—— (*in preparationa*) "Koopman & Szabolsci's *Verbal complexes*; a critical evaluation" Ms. U.Potsdam

—— (*in preparationb*) "Extraction from complex specifiers and the structure of degree phrases" Ms. U.Potsdam

Travis, Lisa (2004) "VP-, D^0 movement languages: Malagasy in wonderland" AFLA 2004 lecture and handout; see also this volume.

Craig Thiersch U. Tilburg, U .Potsdam
Wilgenstraat 45
NL-5000 LE Tilburg
C.L.THIERSCH@UVT.NL