Asia-Pacific Linguistics

Open Access

College of Asia and the Pacific
The Australian National University

Argument realisations and related constructions in Austronesian languages

Papers from 12-ICAL, Volume 2

I Wayan Arka and N. L. K. Mas Indrawati (eds.)

A-PL 013 / SAL 002

This volume contains twenty-two papers describing and discussing the salient features of argument realisations in Austronesian languages as manifested in the nominal or verbal domain, or in both. The Austronesian languages featuring in this volume are typologically and geographically diverse, from those with rich morphology, as in Taiwan, to those that are highly isolating, as in Flores. The papers also reflect diversity in approaches and theoretical frameworks. This volume should be of interest to Austronesianists and typologists.

Asia-Pacific Linguistics

SAL: Studies on Austronesian Languages

Published by Asia-Pacific Linguistics
College of Asia and the Pacific
The Australian National University
Canberra ACT 2600
Australia

First published: 2014

URL: [to be completed]

National Library of Australia Cataloguing-in-Publication entry:

Title: Argument realisations and related constructions in Austronesian languages : papers from 12-ICAL. Volume 2 / edited by I. Wayan Arka and N. L. K. Mas Indrawati.

ISBN: 9781922185129 (ebook)

Series: Asia-Pacific linguistics 013 / Studies on Austronesian languages 002

Subjects: Austronesian languages --Congresses.
Dewey Number: 499.2
Other Authors/Contributors:
Arka, I Wayan, editor.
Indrawati, Ni Luh Ketut Mas, editor.
Australian National University. Department of Linguistics. Asia-Pacific Linguistics

Cover illustration: courtesy of Vida Mastrika
Typeset by I Wayan Arka and Vida Mastrika
.

Preface and Acknowledgements

The 12th International Conference on Austronesian Linguistics was held in Denpasar-Bali in July 2012. The organisers publish a series of compilations of papers based on specific topics, and the present volume is one of the planned four volumes containing papers that describe and discuss argument realisations and related constructions in Austronesian langauges. All papers have been peer-reviewed and revised before publication.

The editors would like to thank our colleagues who happily acted as external referees and some of the contributors who also acted as internal reviewers: Ketut Artawa, Bob Blust, Resty Cena, Bill Davies, Lilian Huang, Marian Klamer, Amy Pei-jung Lee, Paul Li, Hsiu-chuan Liao, Åshild Næsh, Hiroki Nomoto, Ika Nurhayani, I Wayan Pastika, Scott Paauw, Elizabeth Pearce, Sonja Riesberg and Matt Shibatani. We also thank Paul Sidwell, Bryce Kositz, Andrew Morrison and Vida Mastrika for their editorial help and technical assistance in preparing the manuscript for publication.

About the Editors

I WAYAN ARKA is a Fellow at the Department of Linguistics, School of Culture, History and Language, College of Asia and the Pacific ANU and a senior Lecturer at Udayana University Bali (English Department and Graduate Program in Linguistics). His interests are in descriptive, theoretical and typological aspects of Austronesian and Papuan languages of Indonesia. His publications include, among others, books and journal articles on Balinese, Indonesian, Rongga, and Marori. He is currently working on on Marori as part of the ARC-funded project on the Languages of Southern New Guinea (2011-2015) and on his Humboldt project on the languages of Flores (2012-2014).

NI LUH KETUT MAS INDRAWATI is a Senior Lecturer and Head of the Department of English Faculty of Letters, Udayana University. She obtained her PhD in 2012 in Linguistic Typology from Udayana University and her MA in Applied Linguistics in 1994 from Sydney University. Her research interests are in linguistic typology, sociolinguistics and applied linguistics. Her publications include studies on Indonesian reduplication (2008), the role of women in Balinese language maintenance (2008), and Balinese serial verb constructions (2012).

Contents

1 *A comparative look at nominalizations in Austronesian*

WILLIAM A. FOLEY

1 Introduction

Nominalizations are a major site of morphological derivational processes in many languages. Within theoretical linguistics they have been a major focus of analytical contention, in fact their proper analysis being the topic of a major paper by Chomsky (1970), which in turn spawned lexicalist formal theories and indirectly the rivalry between unificational parallel architecture lexicalist theories and procedural serial derivational non-lexicalist theories which so dominates the theoretical landscape of linguistics today (Culicover and Jackendoff 2005; Jackendoff 2007, 2011). Within the more parochial confines of Austronesian linguistics, they play the central role in certain accounts of how the typologically unusual voice systems of Philippine and Formosan languages have evolved (Tuuk 1971 [1864-7]; Capell 1964; Starosta,Pawley and Reid 1982; Kaufman 2009). Yet surprisingly, other than Moyse-Faurie's (2007) discussion of Oceanic languages, little comparative attention has been paid to the syntax of nominalizations within Austronesian. This paper seeks to remedy that gap in a small way by a comparative look at nominalizations in Tolai, of the Oceanic subgroup and spoken in New Britain, Mori Bawah, of the Bungku-Tolaki subgroup of eastern Sulawesi, and Puyuma of Taiwan, argued by Ross (2009) to be a primary first order subgroup of Austronesian. Obviously, with three languages we cannot hope to arrive at a complete and definitive statement of the syntax of nominalizations across Austronesian, but, as these three languages belong to different subgroups, are widely scattered geographically and exhibit significant typological diversity, such an approach should provide us at least with a preliminary understanding of the typological similarities and differences in nominalizations across Austronesian. None of these three languages were included in Moyse-Faurie's survey, so a comparison of our results with hers will further enrich our understanding of the parameters of variation.

Fundamentally, nominalizations across the Austronesian languages fall into two types, what I will call here D-nominalizations and M-nominalizations (see Moyse-Faurie 2007). D-nominalizations are formed by simply taking a finite verb with or without its arguments and placing it as a complement of a D head in a DP, as in these Trukese examples (Goodenough and Sugita 1980):

(1) (a) *wú-se* *wúkkún* *suupwa*
1SG-don't smoke cigarettes
'I don't smoke cigarettes'

Asia-Pacific Linguistics, 2014.

(b) *meyi* *ngngaw* ***ewe*** *wúkkún* *suupwa*
COP bad **D**.INVIS smoke cigarettes
'smoking cigarettes is bad'

While M-nominalizations are done morphologically (hence the M-) by adding an overt derivational affix of nominalization to the verb (with perhaps other morphological effects such as reduplication), as in Tagalog (Schachter and Otanes 1972):

(2) (a) *nag-salita* *si* *Juan* *sa* *klase*
AV-speak NOM.PRP John OBL class
'John spoke in class'

(b) ***pag-****sa-salita* *ni* *Juan* *sa* *klase*
NOMZ-RED-speak GEN.PRP John OBL class
'John's speaking in class'

The analyses proposed in this paper will be cast in the framework of Lexical-Functional Grammar (Bresnan 2001), and, in particular, in its head sharing analysis of the mixed category structure of many nominalizations in the languages of the world. First, we need, following the important insight of Chomsky (1970), to distinguish between nominalizations which are formed by non-productive, clearly lexically stipulated formal correspondences and those derived by productive morphological rule based processes. In the latter case, the derivational rules and their inputs will still be stipulated in the lexicon, but not their outputs, as they are productively generated by the rules, while in the former, the derived nominal themselves also need to be stated in the lexicon. This contrast is easy to illustrate in English with the nominals derived from a verb like *destroy*, i.e. *destruction* versus *destroying*. Consider the following well known examples:

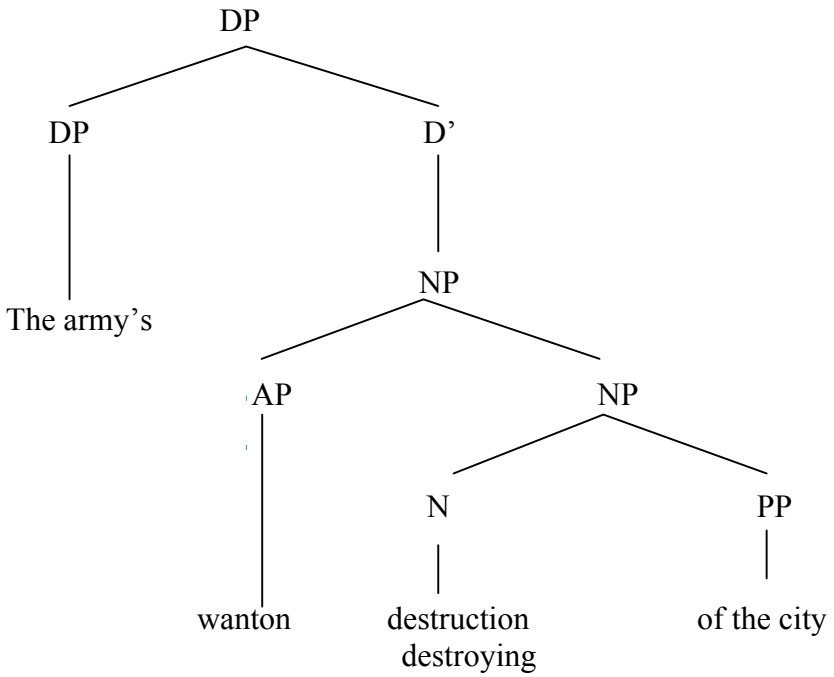
(3) (a) the army's destruction of the city
(b) the army's destroying of the city

The former is a lexically stipulated, morphologically unproductively derived noun from *destroy*, while the latter is productively derived from the same verb by a lexical derivational rule via the nominalizing affix *–ing*. In both of these cases, however, the derived words are clearly nouns, as demonstrated by the fact that they both collocate with modifying adjectives:

(4) (a) the army's wanton destruction of the city
(b) the army's wanton destroying of the city

and comprise the heads of their respective NPs (Bresnan 2001:290-291):

(5)



This is a straightforward NP structure of the type we would find with basic underived nouns, as in (6):

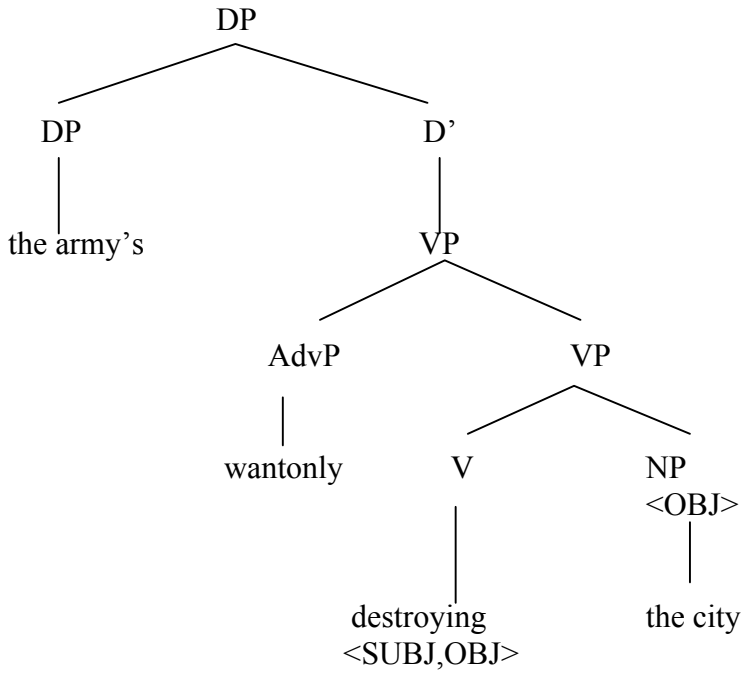
(6) the teacher's long lost book of linear equations

However, (4b) contrasts with a different structure where the nominalized verb's complement does not appear in a PP, the choice for complements of noun heads, but as a bare DP like the complement direct objects of verbs; note that this structure is not available for the lexically stipulated derived noun, *destruction*, nor can it be modified by adjectives, but requires an adverbial modifier instead:

(7) (a) the army's destroying the city
(b) *the army's destruction the city
(c) *the army's wanton destroying the city
(d) the army's wantonly destroying the city

Such nominalizations are the classic mixed nominalization constructions (Lefebvre and Muysken 1988; Malouf 2000); the syntax of these structures is NP-like, taking a possessive specifier within the dominating DP, but VP-like in that the complement of the nominalized verb is realized as a direct object (and with overt accusative case if pronominal). This indicates that the verb does not entirely cede its lexical properties in the process of nominalization; in particular, its subcategorization for an object grammatical relation remains, and, due to the constraints of completeness and coherence, must be specified. It is this which allows syntactic structures for *destroying* which are prohibited for *destruction*; these verbal property retaining nominalizations in English are restricted to the former type. As generally in Lexical-Functional Grammar, the modules of constituent structure and functional structure, in this case, subcategorization for grammatical relations, must be distinguished. This is handled in these English gerund nominalization constructions by embedding a VP directly under the DP (Bresnan 2001: 292):

(8)



The base verb *destroy* is, of course, transitive, and by completeness and coherence requires both its subject and object functions to be satisfied. The object function is satisfied straightforwardly by being a complement, a phrase structure sister, of the V node, but the subject function needs to be satisfied by a rule which states: identify the subject function of a verb derived into a gerund by *–ing* with the POSS specifier function of the dominating DP (Bresnan 2001:294):

(9) V (gerund) ⇒ (↑ POSS) = (↑ SUBJ)

It is not always the case that the difference between lexically stipulated nominalizations and productively derived ones is indicated formally, as with *destruction* and *destroy*. Consider *painting*:

(10) (a) Leonardo's painting of the Mona Lisa
(b) Leonardo's painting the Mona Lisa

Here the formal markings are the same. But note a crucial ambiguity in (10a), not present in (10b); *painting* in (10a) can refer to the actual physical object, framed and now hanging in the Louvre or the event of the painting itself, but, excluding the appositional reading, (10b) only has the latter reading. This specialized meaning of (10a) in referring to the resulting object from the event is clearly a lexical stipulation, and the item must be listed as such in the lexicon, just like *destruction*. Thus, while non-productive forms always correspond to lexically stipulated items, productive forms may or may not. This pattern also correlates with a contrast noticed by Grimshaw (1990). She notes a difference between result nominals, which have no argument structure, i.e no semantically required nominal participants, and event nominals, which may. The simple nominalization *Leonardo's painting* illustrates this well. In the result reading it denotes the object, e.g. the framed work hanging in the Louvre, which comes into existence as a result of the event; clearly no participants are semantically entailed by such a word, *a painting*. But in the event reading (or related manner readings, i.e, Leonardo's style of painting), clearly Leonardo as the agent carrying out the event is entailed. A similar contrast obtains with

destroy. In *the army's wanton destruction*, the only available reading is the result one; the destruction by/of the army is the result of the event occurring, and no participants need be entailed, e.g. *the destruction was extensive*, but such a structure is never possible for *destroying: *the army's wanton destroying.* Rather this as an event nominalization has argument structure and requires an affected participant to be realized as a complement: *the wanton destroying of the city*. The striping away of argument structure and hence functional structure with derived nominals like the result reading of *painting* is a major conduit for the reanalysis of nominals derived from verbs into basic nouns. Indeed, for English, it is doubtful that many naïve native speakers are aware of the derivational histories of *painting, dwelling, building,* etc.

The varieties of nominalization constructions across languages are actually more diverse than the English types we have surveyed thus far, and, in particular, we need to consider head sharing structures (Bresnan 1997; Bresnan and Mugane 2006). These are superficially similar to English gerund nominalizations, but contrast in that both their nominal and verbal properties are projected from a single head word. These are widespread crosslinguistically, as we shall see in examples from Italian (Zucchi 1993), Gikũyũ (Bresnan and Mugane 2006) and Yimas (Foley 1991). As Bresnan (1997) demonstrates, such structures are syntactically coherent and compartmentalized: the mixed structure can be divided into two coherent and consistent constituents, one corresponding to a VP and the other to an NP, with the former embedded into the latter, although both projected from a single shared head:

(11) (a) Italian (Zucchi 1997)

Il	*suo*	*continuo*	*eseguire*	*la*	*canzone*	*impeccabilimente*
the.M	3SG.POSS.M	continual.M	perform.NOMZ	the.F	song.F	impeccably

'his/her continual performing the song impeccably'

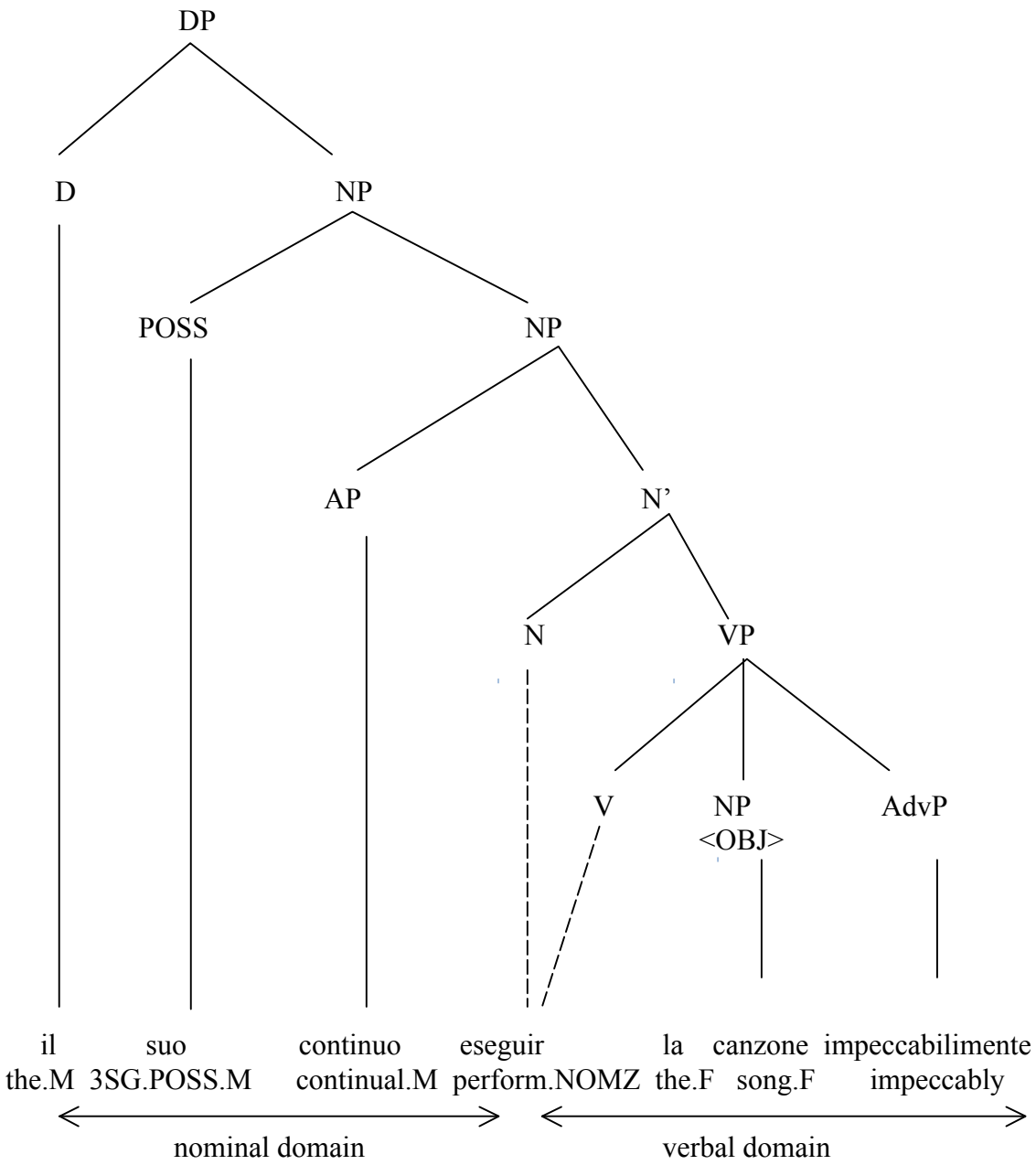
(b) Gikũyũ (Bresnan and Mugane 2006)

mũ-thĩĩnj-ir-i	*a-ndũ*	*mbũri*	*wega*	*ũyũ*
cl.1-slaughter-APPL-NOMZ	cl.2-people	cl.10.goat	well	cl.1.this

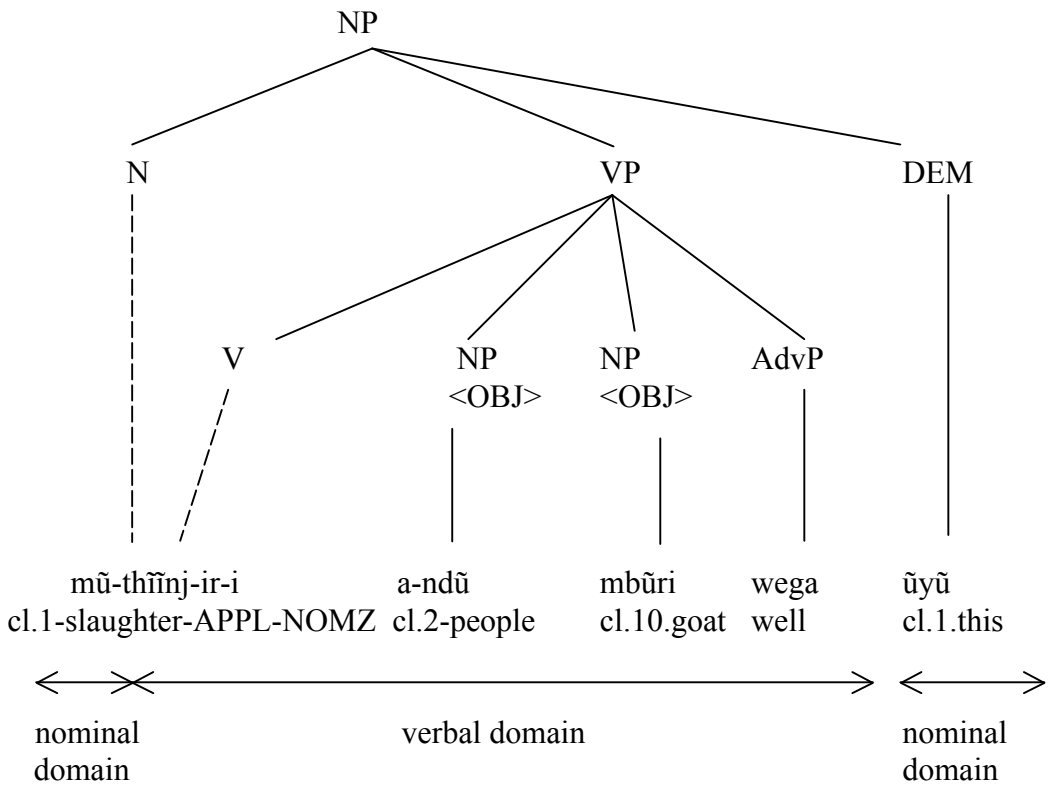
'this goat slaughterer for people well' 'this one who slaughters goats well for people'

Note in both of these languages a mixed syntactic structure consisting of nominal and verbal properties, but delimited to coherent properties of the phrase. In the Italian example the nominal markers of the gender-specified determiner, possessor and adjective precede the head, while the verbal properties of taking a direct object complement and an adverbial modifier follow it. In Gikũyũ the head takes the distinctly nominal class marking prefix, but is suffixed with the exclusively verbal applicative suffix *–ir* and followed by two subcategorized object grammatical functions and a modifying adverb. The corresponding phrase structures for (11) are in (12), with dashes indicating the shared head arrangement (Bresnan 1997):

(12) (a) Italian


DP
D
NP
POSS
NP
AP
N'
N
VP
V
NP
<OBJ>
AdvP
il suo continuo eseguir la canzone impeccabilimente
the.M 3SG.POSS.M continual.M perform.NOMZ the.F song.F impeccably
nominal domain
verbal domain

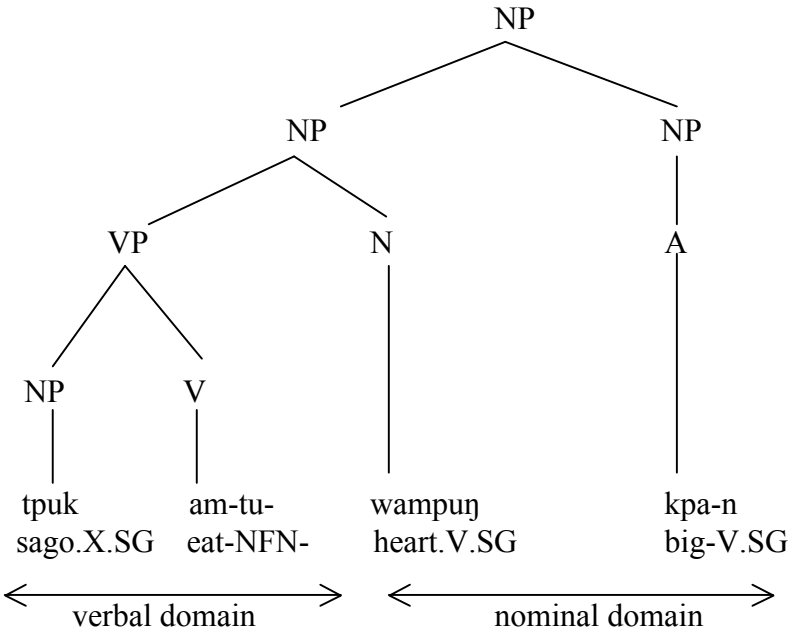
(b) Gikũyũ



It is important to realize that not all nominalizations exhibit the mixed structure with shared heads. Besides the English gerund in which we find a discrete and intact VP, albeit with a head verb deficient in some verbal properties like tense, embedded in a DP, other languages form nominalizations with both overt nominal and verbal heads, so there is no head sharing. Yimas (Foley 1991) is an example of such a language:

(13) (a) *tpuk* *am-tu-wampuŋ* *kpa-n*
sago.X.SG eat-NFN-heart.V.SG big-V.SG
'a big craving to eat sago'

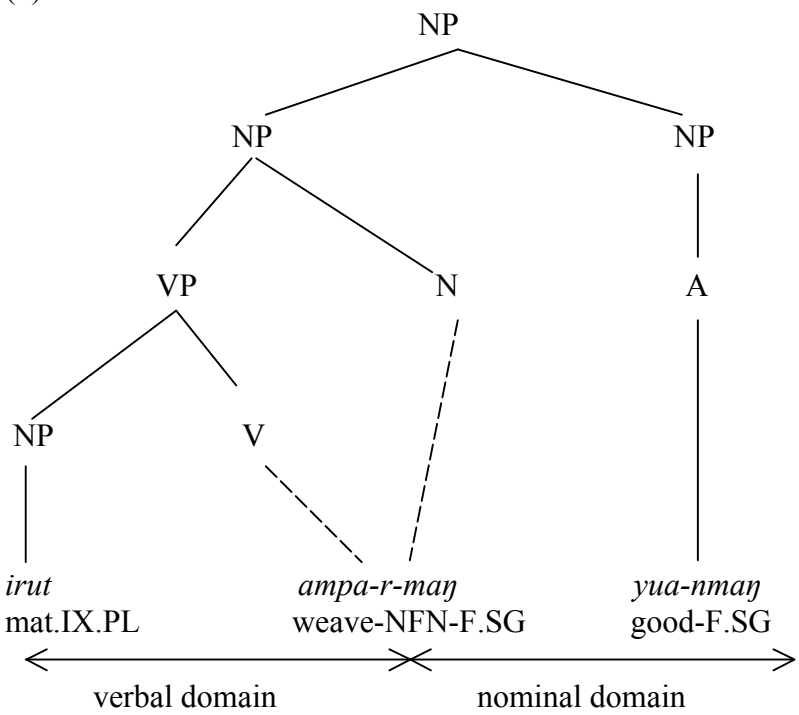
(b)



Note that both the VP and NP have their own heads, *am-tu* eat-NFN and *wampuŋ*'heart, care' respectively, which project their own distinct endocentric phrases; there is no shared head. However, in most cases of nominalizations in Yimas, a lexical noun like *wampuŋ*'heart, care' is replaced as head by a suffix, but one which carries all grammatical features of nouns like gender and number, so these now look similar to the shared head nominalization of (12b):

(14) (a) *irut* *ampa-r-maŋ* *yua-nmaŋ*
mat.IX.PL weave-NFN-F.SG good-F.SG
'a good mat-weaver'

(b)



With this theoretical background, we are now ready to turn to the individual analyses of the three Austronesian languages.

2 Tolai

The data for Tolai are drawn from the grammars of Bley (1912), Mosel (1984) and Rinderknecht (1987), the dictionaries of Lanyon-Orgill (1960) and Wright (1964), and the text collections of Kleintitschen (1924), Meier (1909) and Mosel (1977). Tolai belongs to the class of languages known as discourse-configurational (Kiss 1995); the word order of full DP clausal constituents is not strictly fixed by structural principles, in this case the assignment of grammatical relations, subject and object, but rather is determined by pragmatic discourse information structure categories, topic and focus, here, following LFG tradition called discourse functions (Bresnan 2001; King 1995). The grammatical relations in Tolai are always defined though a mix of structural determination and agreement: the object is the complement sister of its verb, either an anaphoric agreement accusative pronominal, if the full object DP is in a discourse function position, or the in situ DP. Object DPs in such discourse function positions are functionally identified through the accusative pronominal occupying the complement position as the sister of the verb. Subject full DPs, on the other hand, are always determined by functional identification with the overt nominative pronominals, regardless of their positioning; whether these nominative pronominals are functioning as grammatical or anaphoric agreement (Bresnan and Mchombo 1987) remains to be established by further research.

(15) (a)

	S		V	O	
a	*tutana*	*i*	*kita*	*ra*	*bul*
D	man	3SG.NOM	hit	D	child

'the man hit the child'

(b)

		V	O		S
i	*ga*	*kita*	*ra*	*davai*	*ra*	*vavina*
3SG.NOM	REM	hit	D	wood	D	woman

'the woman chopped the firewood'

(c)

	O				S	
a	*bala*	*na*	*boroi*	*go*	*ra*	*maleo*
D	stomach	C	pig	DEM	D	eel

	V		
i	*ra*	*pa*	***ia***
3SG.NOM	rob	PERF	3SG.ACC

'the pig's guts were snatched by this eel'

(d)

		O			S	
nam	*bula*	*ka-dor*	*vavagui*	*ka-dor*	*umana*	*bul-mur*
DEM	also	POSS-1DL.IN	animal	POSS-1DL.IN	PL	child-back

		V	
diat	*a*	*rapu*	***ia***
3PL.NOM	IRR	hit	3SG.ACC

'our animals will also be hit by our descendants'

(e)

			V		S		O		
diat	*iat*	*dia*	*vatang*	***ia***	*ra*	*umana*	*Iapan*	*a*	*take*
3PL	self	3Pl.NOM	call	3SG.ACC	D	PL	Japan	D	sake

'they themselves, the Japanese called it sake'

Prototypical focal elements, the information question words, can appear in situ or in a discourse function position at the beginning of the sentence; in the latter case they are indistinguishable from cleft constructions, so widespread among Austronesian languages for information questions:

(16) (a)

to	*ia*	**(i)*	*ga*	*tova*	*u*
D	who	3SG.NOM	REM	teach	2SG.ACC

'who taught you?'

(b)

u	*tar*	*gire*	*to*	*ia*
2SG.NOM	PERF	see.TR	D	who

'who have you seen?'

(c)

u	*papait*	*ra*	*ava*
2SG.NOM	do.RED	D	what

'what are you doing?'

(d)

ava	*u*	*papait*	***ia***
what	2SG.NOM	do.RED	3SG.ACC

'what are you doing?'

(e)

amur	*tabar*	*kom-amur*	*tarai*	*ma*	*ra*	*ava*
2DL.NOM	give	POSS-2DL.GEN	people	with	D	what

'what did you two give your people?'

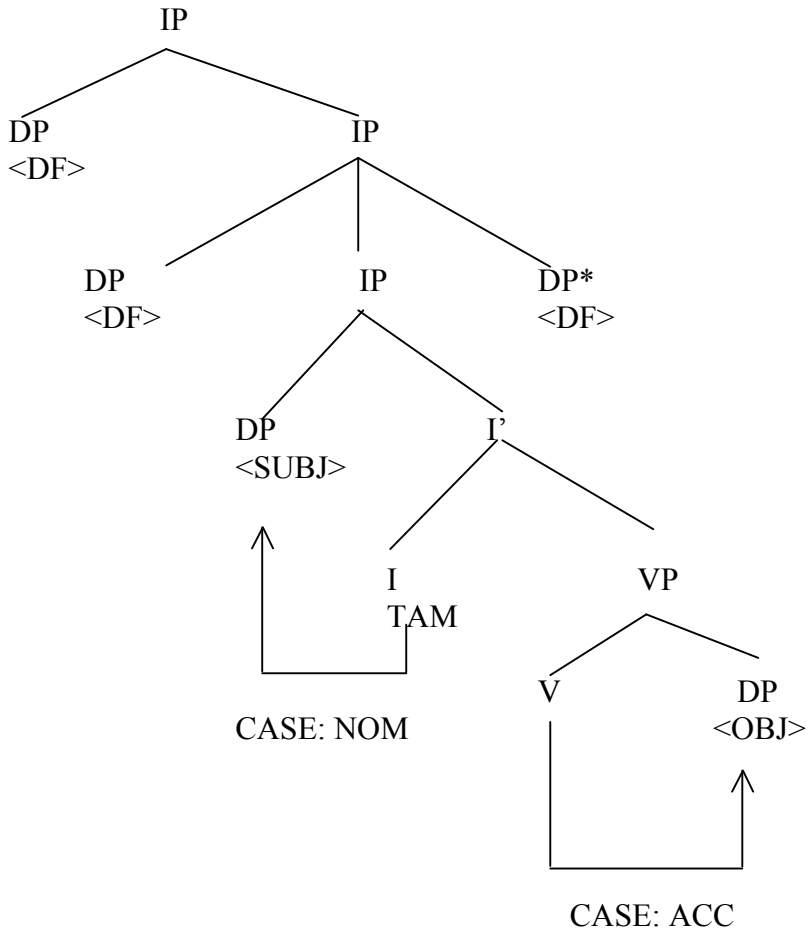
(f)

ava	*nam*	*ra*	*vavina*	*i*	*garau*	*ta-**na***
what	DEM	D	woman	3SG.NOM	anxious	P-3SG.GEN

'what is that woman anxious about?'

These clause patterns suggest a phrase structure like the following:

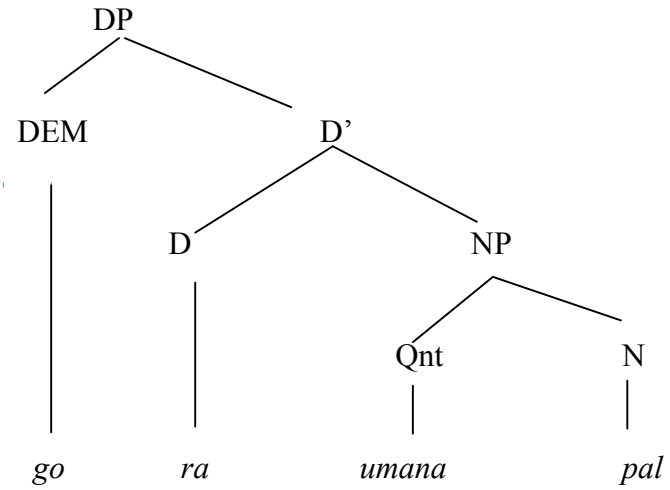
(17)



The subject and object grammatical relations are filled for a subject by the nominative pronominal to which any overt DP links by functional identification by agreement and for an object by being the complement as a phrase structure sister of the lexical verb or by anaphoric agreement to an accusative pronominal in that position.

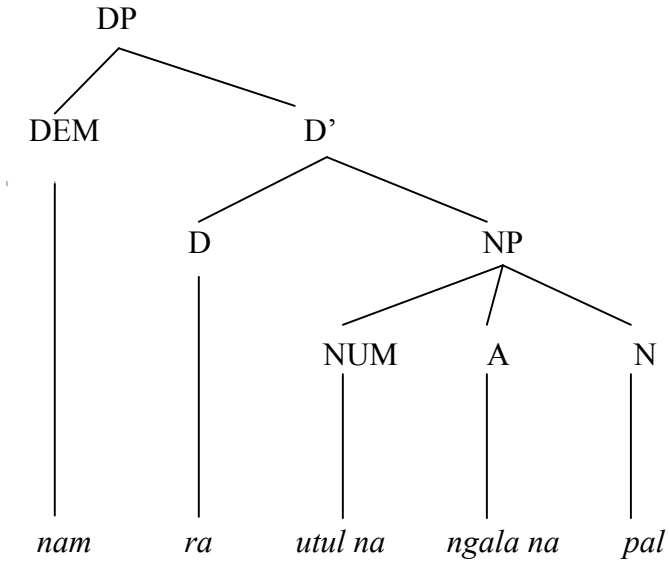
The structure of the DP parallels in many respects that of the IP, and the NP that of the VP. The determiner *a* ~ *ra* has complex conditioning factors which are not completely understood, and demonstratives function as the specifier of the DP, hence:

(18) *go* *ra* *umana* *pal*
DEM D PL house
'these houses'



Modifiers of the head noun require a connective particle *na* (formally realized as *a* after a word-final obstruent); this is not dissimilar to the function of *–ng/na* in Tagalog (Foley 1979):

(19) *nam ra utul* ***na*** *ngala* ***na*** *pal*
DEM D three C big C house
'those three big houses'



Nouns modifying nouns which specify a characteristic of the head noun are also linked by *na,*rather like the use of *of* in English:

(20) (a) *ra pal* ***na*** *kunai*
D house C grass
'a grass hut/ a hut of grass'

(b) *ra mapi* ***na*** *davai*
D leaf C tree
'a tree leaf'

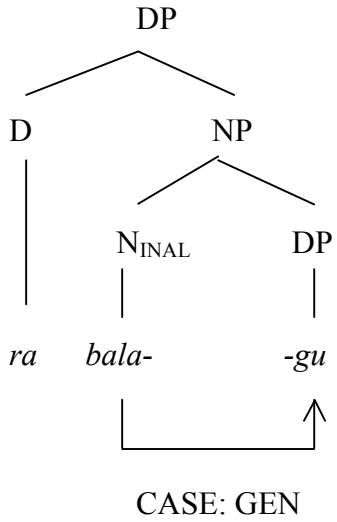
Tolai nouns differ from those of more familiar languages like English and exhibit a typical Oceanic pattern in that they divide into two classes: one which has argument structure and could be labeled transitive nouns, as they require an overtly expressed complement, and another class which lacks argument structure. The former are traditionally known as inalienable nouns and the latter, alienable, and in Tolai, again like Oceanic languages generally, the alienable nouns themselves divide into two classes as to whether they take dominant or subordinate possession. The class of inalienable nouns is comprised of the usual suspects: body parts, kin terms, and other types of relations. Inalienable nouns take possessors directly suffixed with singular pronominal DPs or with the genitive case marker *i* otherwise:

(21) (a) *ra bala-gu*
D belly-1SG.GEN
'my belly'

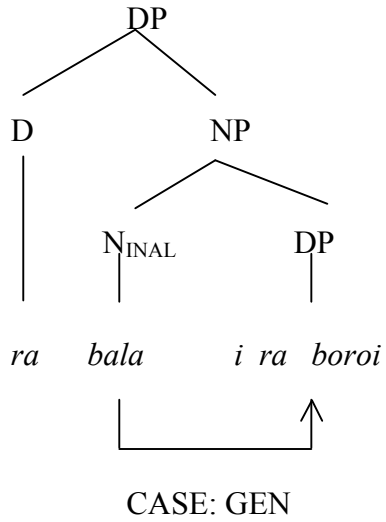
(b) *ra bala i ra boroi*
D belly GEN D pig
'the pig's belly'

We can treat *bala* 'belly' somewhat like a transitive verb, as having argument structure with an obligatory complement, a [__POSS] grammatical function, to which it must assign genitive case; failure to do so will violate completeness and produce an ungrammatical string:

(22) (a)



(b)



In some Oceanic languages the parallelism between complement taking inalienable nouns and transitive verbs is even more apparent in that the genitive and accusative pronominals are closely related or even homophonous (Rotuman, Gapapaiwa, 'Ala'ala (Donohue, personal communication). Outside of Oceanic such homophony is found in Sou Amana Teru of central Maluku. Even in Tolai the relationship is striking in that non-singular accusative pronouns often occur with an overt case marker *i*, exactly the same formative when they are used as possessors:

(23) (a) *i* *ga* *pulu* ***i*** *avet*
3SG.NOM REM wrap GEN/ACC 1PL.EX
'it enveloped us'

(b) *dia* *ga* *vong* ***i*** *avet*
3PL.NOM REM lie GEN/ACC 1PL.Ex
'they lied to us'

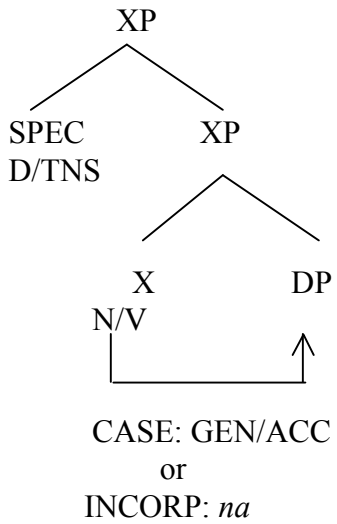
Furthermore, generic complements of both inalienable nouns and transitive verbs are incorporated and linked to their heads by the connective *na*:

(24) (a) *ra* *bala* *na* *boroi*
D belly C pig
'pig guts'

(b) *i* *rovai* *na* *boroi*
3SG.NOM hunt C pig
'he pig-hunts'

So the phrase structure for NPs and VPs with inalienable nouns and transitive verbs respectively as heads (and their dominating DPs and IPs) are closely parallel:

(25)



Alienable nouns have no argument structure and hence no obligatory complements, and they fall into two classes depending on whether the possessor has a dominant, controlling, agent-like relationship to the possessum, marked by the possessive classifier *ka-,* versus a subordinate, determined, patient-like relationship, indicated by *a-* ~ *ra-*. The possessive classifiers are directly followed by the possessor DPs if pronominal suffixes; full DPs are extraposed behind the possessum NP:

(26) (a) (i) *kau-gu* *tutana*
POSS-1SG.GEN man
'my husband'

(ii) *ra* *pal* *ka* *i* *ra* *umana* *tutana*
D house POSS GEN D PL man
'the house of the men'

(b) (i) *a-na* *vudu*
POSS-3SG.GEN banana
'his banana' (to be eaten)

(ii) *ra* *mal* *a* *i* *ra* *tutana*
D loincloth POSS GEN D man
'the man's loin cloth'

These possessive classifiers play a major role in the syntax of nominalizations, in a way that will be familiar to specialists in Polynesian languages. Nominalizations with possessive classifiers are of the M-nominalization type, and Tolai typically uses an infix -*in*- with its allomorphs to nominalize verbs. When an agentive transitive or intransitive verb (the language is structurally stative-active in alignment) is nominalized, its subject's nominative case pronominal is replaced by genitive, with a co-occurring dominant, agent-like possessive classifier *ka*-:

(27) (a) (i) *iau* *ga* *pot*
1SG.NOM REM arrive
'I arrived'

(ii) ***kau-****gu* *p-in-ot*
POSS-1SG.GEN arrive.[NOMZ]
'my arrival'

(b) (i) *a* *tutana* *i* *ga* *pot*
D man 3SG.NOM REM arrive
'the men arrived'

(ii) *ra* *p-in-ot* ***ka*** *i* *ra* *tutana*
D arrive.[NOMZ] POSS GEN D man
'the arrival of the man'

(c) (i) *a* *tutana* *i* *kul* *(r)a* *boroi*
D man 3SG.NOM buy.TR D pig
'the man bought a pig'

(ii) *ra* *k-un-ukul* *na* *boroi* ***ka*** *i* *ra* *tutana*
D buy.[NOMZ] C pig POSS GEN D man
'the man's buying of the pig'

However, if the verb is a stative, non-agentive one, the possessive classifier can be *a-*:

(28) (a) *a* *umana* *Iapan* *dia* *ga* *kankan*
D PL Japan 3PL.NOM REM be.angry
'The Japanese were angry'

(b) *i* *ga* *al* *pa* *nam* ***ra****-na* *kankan*
3SG.NOM REM attract PERF DEM POSS-3SG.GEN be.angry
'he drew his anger (upon himself)'

And, surprisingly, the subordinate possessive classifier is used in some cases where the verb is agentive; what conditions this is unclear, although perhaps the nominalized verb has now been reanalyzed as an alienable noun similar to pieces of clothing, which are treated as subordinately possessed in Tolai:

(29) (a) *ina* *me*
1SG.NOM.IRR chew.TR
'I will chew it (mostly betelnut)'

(b) *ina* *vung* *boko* ***ra****-gu* *m-in-amai*
1SG.NOM.IRR put yet POSS-1SG.GEN chew.ITR.[NOMZ]
'I will still produce my betelnut chewing stuff

For objects of transitive verbs in nominalizations there are a number of options. Normally, objects are incorporated as part of a process of detransitivization that is widespread for nominalized verbs. Incorporated objects of verbs, nominalized or not, are typically generic, obligatorily so in the Traditional Tolai of the texts of Kleintitschen (1924) and Meier (1909):

(30) (a)

a	*tutana*	*i*	*ga*	*rove*	*ra*	*boroi*
D	man	3SG.NOM	REM	hunt.TR	D	pig

'the man hunted a/the pig'

(b)

a	*tutana*	*i*	*ga*	*rovai*	*na*	*boroi*
D	man	3SG.NOM	REM	hunt.ITR	C	pig

'the man hunted for pig(s)'

(c)

ra	*ni-rovai*	*na*	*boroi*	*ka*	*i*	*ra*	*tutana*
D	NOMZ-hunt.ITR	C	pig	POSS	GEN	D	man

'the man's hunting of pig(s)'

In Traditional Tolai texts incorporated objects are always bare nouns (N_o), as befits generics, but in the version of Modern Tolai represented by Mosel's (1977) texts, incorporated objects may be full NPs (but never DPs). This is closer to the type of noun incorporation we find in Polynesian languages and permits nominalizations of verbs with specific objects:

(31) (a)

ra	*m-in-omo*	*na*	*ongor*	*na*	*sipirit*
D	drink.ITR.[NOMZ]	C	strong	C	liquor

'the drinking of strong liquor'

(b)

i	*pait*	*ka-na*	*t-in-ata*
3SG.NOM	do	POSS-3SG.GEN	speak.ITR.[NOMZ]

na	*var-va-dovot*	*ure*	*ra*	*var-takun*
C	NOMZ-CAUS-true	about	D	NOMZ-accuse

pire	*August*	*To*	*Ima*
near	August	D	Ima

'he made his speech for finding out the truth about the accusation against August To Ima'

In Traditional Tolai this option is not available, but there are a number of ways to express non-generic objects in nominalizations. First, rather like English *the city's destruction* versus *the army's destruction of the city,* if the agent is not present, the object can be expressed by a genitive case marked NP with the dominant possessive classifier *ka-*, even though the semantics here is clearly not agentive:

(32) (a)

da	*tibe*	*diat*	*ma*	*ra*	*vudu*
IPR.NOM.IRR	divide.TR	3PL.ACC	with	D	banana

'the bananas will be divided among them'

(b)

i	*vung*	*mut*	*vue*	*ka-dir*	*t-in-iba*
3SG.NOM	put	all	down	POSS-3DL.GEN	divide.ITR.[NOMZ]

'he finished off their allotments'

An alternative to this is to demote the object noun to a prepositional phrase:

(33) (a)

diat	*ga*	*tataue*	*ra*	*pap*	*ma*	*ra*	*okin*
3PL.NOM	REM	speak.TR	D	dog	with	D	wallaby

'they spoke about the dog and the wallaby'

(b)

ra	*t-in-ata*	*ure*	*ra*	*pap*	*ma*	*ra*	*okin*
D	speak.ITR.[NOMZ]	about	D	dog	with	D	wallaby

'a story about a dog and a wallaby'

tataue speak.TR in (33a) is apparently a contraction of *tata ure* 'speak about'. Yet another possibility is for the object NP to be realized as a complement to the nominalized verb, like an inalienable possessor, but this seems very lexically restricted to a few verb roots and these might be better analyzed as lexicalized inalienably possessed nouns:

(34) (a)

iau	*ga*	*tataue*	*ra*	*tubuan*
1SG.NOM	REM	speak.TR	D	spirit

'I spoke about the spirit'

(b)

iau	*nukure*	*ra*	*t-in-ata*	*i*	*nam*	*ra*
1SG.NOM	know.TR	D	speak.ITR.[NOMZ]	GEN	DEM	D

tubuan
spirit
'I know the story of that spirit'

But the most striking option is to present the object as a complement DP of the nominalized verb within a **VP**, in other words to employ a D-nominalization in place of the expected M-nominalization. This is most obvious with pronominals, which overtly show accusative and not genitive case:

(35) (a)

pa	*u*	*ngo*	*ra*	*va-bilak*	*diat*
NEG	2SG.NOM	stop	D	CAUS-bad	3PL.ACC

'you didn't stop hurting them'

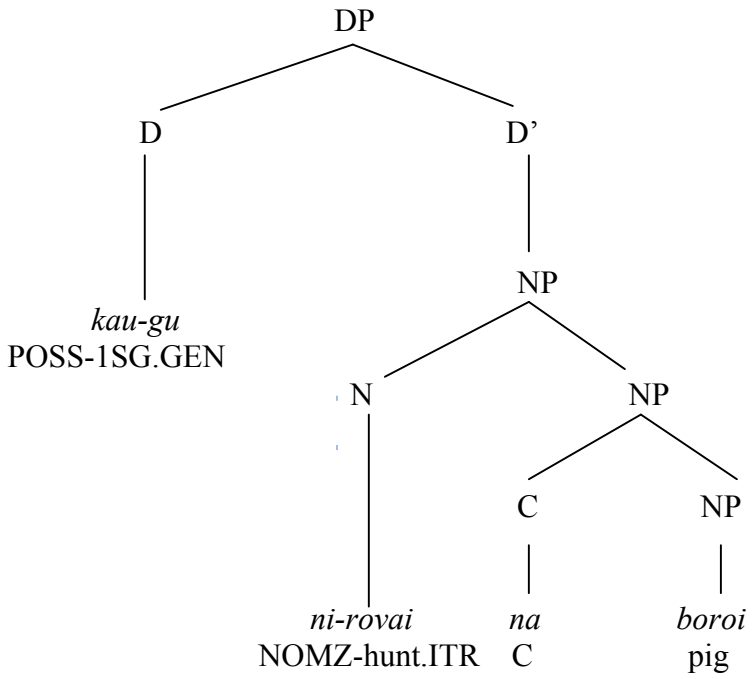
(b)

i	*ga*	*va-utul*	*ra*	*rapu*	*ia*
3SG.NOM	REM	CAUS-three	D	hit.TR	3SG.ACC

'he hit him three times' = 'the hitting him was thrice'

In the previous types of M-nominalizations, the relevant syntactic structure is straightforwardly that of a NP, and the nominalized verb behaves as a canonical noun, having ceded all of its verbal properties in the process of nominalization. Hence these are syntactically like the English nominalizations of (5) (remember nouns as well as verbs can incorporated their complements, see (24)):

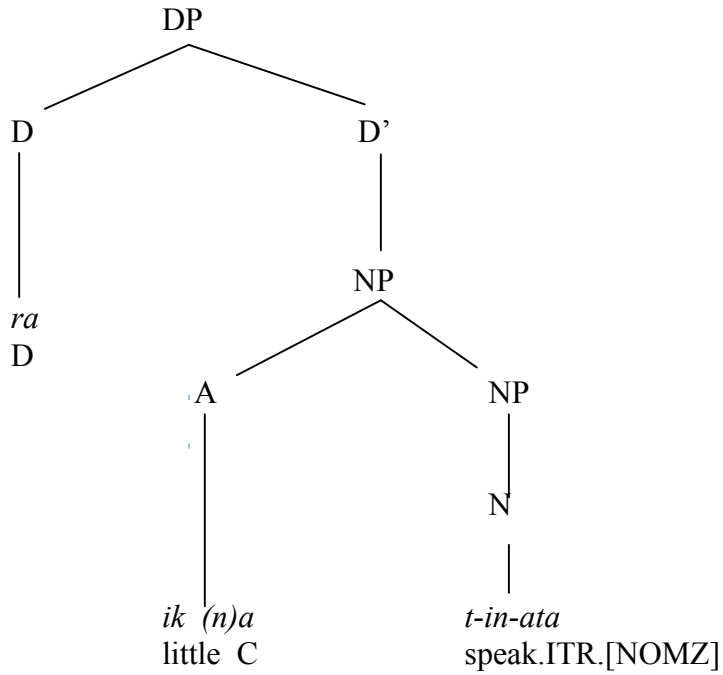
(36)



'my pig-hunting'

The nominal properties of such nominalizations is so robust as to permit adjectival modification, although the possibility that these have been reanalyzed as regular nouns along the lines of English *painting* cannot be ruled out:

(37)

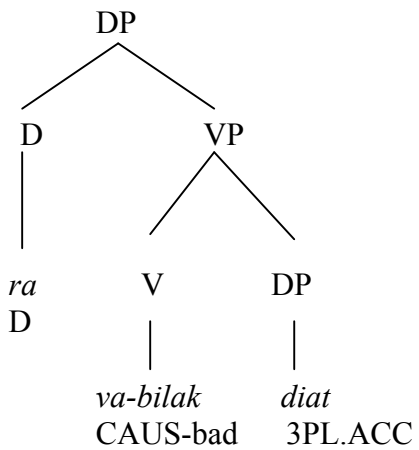


'a little speech'

But the D-nominalization examples of (35) cannot be analyzed in this fashion. They are like the English gerunds of (8) in that the nominalized verb does not cede its subcategorization for an object complement and hence the requirement to assign accusative case to such complements. All of the examples I have of such constructions in Tolai are

subjectless, unlike English gerunds, so only the subcategorized object grammatical relation is carried over into the nominalized form:

(38)



Note further differences between the nominalizations of (36) and (38). There is no overt nominalization morphology in (38), as the nominalization is being simply accomplished by being the complement of a D functional head, and the verb is fully transitive and requires the assignment of accusative case to its object complement. But in (36) the verb is intransitive, with the erstwhile object an obligatorily incorporated adjunct marked by *na*. In fact, nominalizations of the type represented by (36) always require the verbs to occur in their intransitive forms. This is clearly due to the fact that they are nouns, alienable ones at that, which lack argument structure and hence cannot co-occur with overt complements:

(40) (a) *a tutana i ga mom**e** **ra** tava*
D man 3SG.NOM REM dring.TR D water
'the man drank the water'

(b) *a tutana i ga mom**o** **na** tava*
D man 3SG.NOM REM drink.ITR C water
'the man drank water'

(c) *ra m-in-om**o** nat ava ka i ra tutana*
D drink.ITR.[NOMZ] C water POSS GEN D man
'the man's drinking of the water'

(d) **ra m-in-om**e** ra tava ka i ra tutana*
D drink.TR.[NOMZ] D water POSS GEN D man

A particularly striking fact about all M-nominalizations of verbs in Tolai is that only intransitive verbs can be the input to the morphological derivational process of nominalization. This is so noticeable that earlier researchers of the language (Bley 1912; Rinderknecht 1987) posited a number of nominalization morphemes, but all of these save *ni-* ~ *-in-* are actually intransitivizing verbal formatives (and remember that even *ni-* ~ *-in-* can only be added to intransitive verb forms):

(41)

	V_{TR}	V_{ITR}	NOMZ
formed by reduplication:			
‘throw’	bir	birbir	birbir
‘shout’	bor	bobor	bobor
‘retaliate’	bali	babali	babali
‘enchant’	kom	komkom	komkom
‘buy’	kul	kukul	kukul
‘think’	nuk	nuknuk	nuknuk
‘do’	pait	papait	papait
‘work’	palum	papalum	papalum
‘give’	tabar	tababar	tababar
‘write’	tumu	tutumu	tutumu
‘blow’	vu	vuvu	vuvu
formed with the reciprocal prefix *var-*:			
‘obstruct’	bat	var-bat	var-bat
‘marry’	ben	var-ben	var-ben
‘squeeze’	bing	var-bing	var-bing
‘see’	bobo	var-boboi	var-boboi
‘abuse’	gi	var-gi	var-gi
‘bite’	karat(e)	var-karat	var-karat
‘catch’	kinim	var-kinim	var-kinim
‘hit’	kita	var-kita	var-kita
‘ask’	tir	var-tir	var-tir
‘teach’	tovo	var-tovo	var-tovo
formed by conversion, ie no overt formative:			
‘rain’		bata	bata
‘flood’		lubu	lubu
‘discharge (pus)’		ben	ben
‘wave’		bobol	bobol
‘clear’		kapa	kapa
‘angry’		kankan	kankan
‘dream’		ririvon	ririvon
‘be in pain’		mog	mog
‘ulcer(ate)’		manua	manua
‘pray’		lotu	lotu
‘bleed’		gap	gap
‘sin’		pekato	pekato (< Latin)

The overwhelming homophony between intransitivizing and nominalizing morphemes suggests that these are not nominalizing morphemes at all, merely intransitivizing ones. In line with the intransitive constraint we saw with *ni-* ~ *-in-,* all verbs must be intransitivized before they undergo the process of nominalization. What distinguishes the examples in (41) from previous ones is that nominalization in these cases is not marked morphologically by *–in-*; it is simply done by conversion. What distinguishes these cases from (38), the likely historical source of zero nominalization in Tolai, is that the examples

of (41) are truly derived nouns and occur in structures like (36). Further evidence for this is that many are lexicalized nominalizations with quite idiosyncratic meanings: *bat* 'guard against', *var-bat* 'defence', but also 'jealousy, ill will'; *palum* 'work, touch', *var-palum* 'shaking of hands'; *puk* 'soak', *pukpuk* 'crocodile'; *bolo* 'cross', *bolo* 'a crossbeam'; *dolom* 'swallow', *dolom* 'a glutton'; as indeed are some forms with *ni-* ~ *-in-*: *mat* 'die', *m-in-at* 'a corpse'; *dok* 'pay', *ni-dok* 'marriage presentations'; *igir* 'cook in coconut milk', *ni-igir* 'a portion of vegetables cooked in coconut milk'. This is, of course, completely parallel to lexicalized nominalizations in English like *building, my favorite show, his stay,* etc.

3 Mori Bawah

Mori Bawah is a member of the Bungku-Tolaki sub-group (Mead 1998) of central and southeastern Sulawesi; all data here are drawn from Esser's (2011 [1927-1933]) comprehensive grammar, a monograph on verb morphology by Barsel (1994) and a set of papers by Mead (1999, 2005, 2008). Its basic syntactic structure is rather similar to other verb initial languages of Indonesia, Malaysia, the Philippines and Taiwan (see, for example Kroeger's (1993) analysis of Tagalog), and indeed the language bears quite a close resemblance to Chamorro (Chung 1998; Cooreman 1987) in many respects. As Esser (2011 [1927-1933]) points out, the language is VSO in unmarked word order, although it is not all that common to have two overt full NPs for both grammatical relations. VOS is also acceptable, provided there would be no ambiguity. Mori Bawah also signals grammatical relations by pronominal verb agreement, by absolutive enclitics for objects and by ergative proclitics for subject (the latter have been analyzed as nominative by earlier researchers in Bungku-Tolaki languages, eg Mead (2005, 2006), but I regard them as ergative; the grounds for this disagreement are too complex to go into here). There is no gender or number marking for nouns, and if both subject and object are overt third person NPs with animate or human referents, fixed order of VSO is necessary to clarify the assignment of grammatical relations:

(42) *i='akala=o* *i* *Re'a* *i* *Bange*
3SG.ERG=deceive=3SG.ABS PN Turtle PN Monkey
'Turtle deceived Monkey'
*'Monkey deceived Turtle'

Nouns freely occur before the verb in discourse functions like topic or focus:

(43) (a) *i* *Rawu andio* *i=sampaowea=o* *i* *Luka*
PN Blind DEM 3SG.ERG=carry=3SG.ABS PN Lame
'Blind carried Lame on his shoulders'

(b) *i* *Re'a* *langkai=o=mo* *punti-no,*
PN Turtle BIG=3SG.ABS=PERF banana-3SG.GEN

i *Bange* *pingko=o=mo* *punti-no*
PN Monkey finish=3SG.ABS=PERF banana-3SG.GEN
'Turtle's bananas were already big, Monkey's bananas were already finished'

Focal question words also occur sentence initially, although not necessarily in cleft constructions like in Philippine languages, and they can also appear in situ as in Tolai, again unlike Philippine languages:

(44) (a) *isua* *u=hawe=o* *ama-mu*
where 2SG.ERG=meet=3SG.ABS father-2SG.GEN
'where did you meet your father?'

(b) *isema* *u='aiwa-ako*
who 2SG=come-APPL
'who did you come for?'

(c) *ko* *wee-ako=no* *isema*
3SG.FUT give-APPL=3SG.ABS who
'who will you give it to?

Unlike Philippine-Formosan languages, Mori Bawah has no obligatory overt inflection for tense, mood or aspect. But there are two aspectual clitics, *mo* PERF and *po* IMPERF. These are second position clitics, which attach to the first constituent of the clause, commonly a verb, but not necessarily:

(45) (a) *me-lako='ira=mo* *m-pom-paho*
PL-go=3PL.ABS=PERF PL-ANTI-plant
'they have gone out to plant'

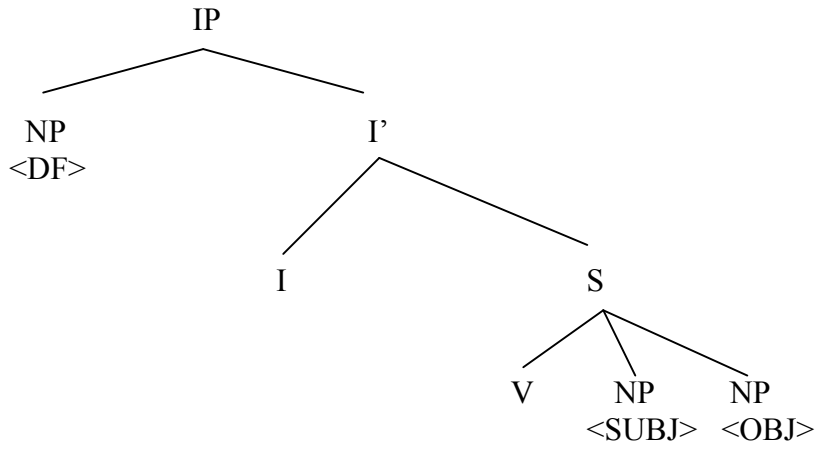
(b) *mondee-no=mo* *anu* *n-in-ee-no*
whatever-3SG.GEN=PERF REL say.[PASS]-3SG.GEN
'regardless of what he said'

(c) *bou-mu=mo* *koa* *ku='ala=o*
fish-2SG.GEN=PERF just 1SG.ERG=get=3SG.ABS
'I got your fish then'

(d) *rua* *wongi=po* *ka* *do=me-'amba* *me-lako*
two night=IMPERF and 3PL.ERG=PL=then PL=go
'two nights more and then they'll go'

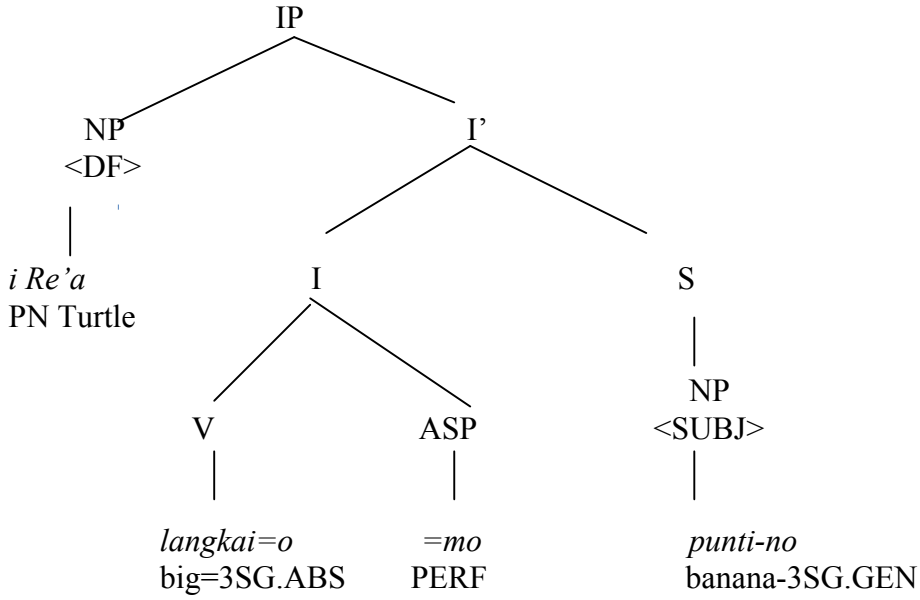
In fact, these clitics occupy the I functional head slot, and attach to whatever proceeds them, either a verb in the I position, or if lacking, the constituent in a discourse function in the specifier of IP position. These facts suggest a basic clause constituent structure for Mori Bawah as follows (compare Kroeger's (1993) analysis of Tagalog):

(46)

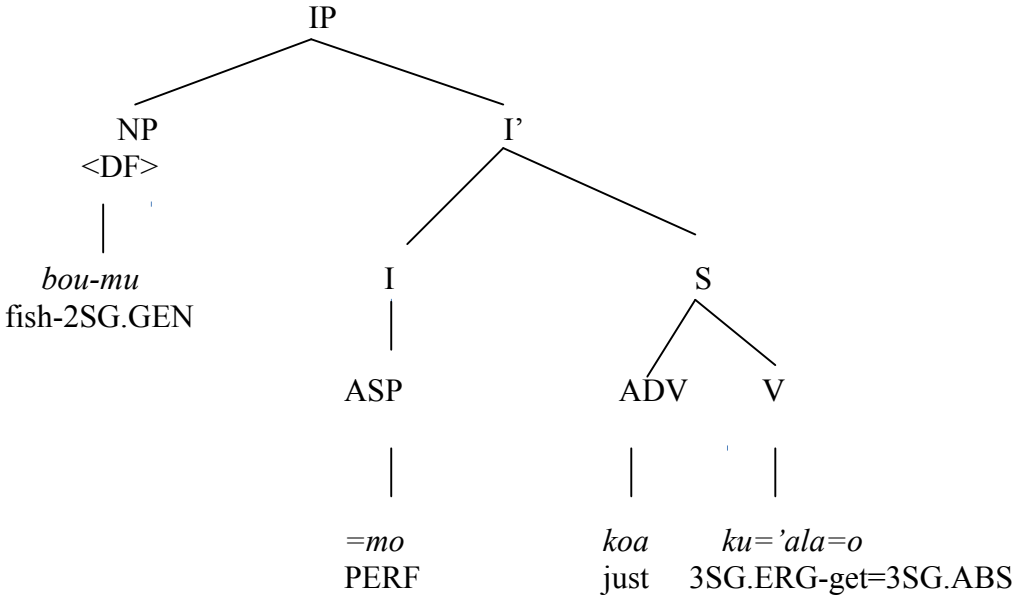


As both the verb and the aspectual clitics are the bearers of I features, albeit covert for the former, one or both must occupy the I head position:

(47) (a) = (43b)



(b) = (45c)



While by no means comparable with Philippine languages, Mori Bawah has quite a rich system of morphology, and indeed voice oppositions. There is a primary contrast between undergoer voice (UV), which is unmarked and signaled exclusively by pronominal agreement, and actor voice (AV), which is marked by *–um-* ~ *m(V)* ~ Ø. In addition to these, there are both passive and antipassive voices. The details are too complex to go into here, but a summary explication is necessary. The unmarked voice formally, syntactically and discoursally is UV, as in Philippine languages. Of course, UV and AV can only fully contrast with transitive verbs. If the object NP is definite and highlighted, UV is obligatory. This is indicated by absolutive pronominal agreement, which is obligatory for the object, and normally ergative pronominal agreement for the subject:

(48) (a) *i=kaa=ira* *i* *Ngeo*
3SG.ERG=eat=3PL.ABS PN Cat
'Cat ate them up'

(b) *bou-mu=mo* *koa* *ku='ala=o*
fish-2SG.GEN=PERF just 1SG.ERG=get=3SG.ABS
'I got your fish then'

The unmarked status of the UV shows the expected correlation with ergative alignment patterns in the syntax; for example:

(49) (a) *molai=o=mo* *i* *Bange,*
flee=3SG.ABS=PERF PN Monkey
i=poboi=o=mo *i* *Bibiundi*
3SG.ERG=call=3SG.ABS=PERF PN Wild.Duck

'Monkey took off fast, Wild Duck called him'
*'Monkey took off fast, he called Wild Duck'

(b) *tutulu-no* *i* *Bonti*
story-3SG.GEN PN Wild.Pig
i='oluako=no *mia*
3SG.ERG=invite=3SG.ABS person

'the story of Wild Pig: a person invited him'
*'the story of the Wild Pig: he invited the person'

AV is marked by *–um-* and its allomorphs and, as the marked option, has a more restricted distribution; its use is similar to what has been dubbed agent focus in Mayan languages (Aissen 1999; Coon 2009). Verbs in AV with *–um-* may not co-occur with ergative pronominal clitics, but absolutive clitics remain to mark objects if the verb is transitive (or subjects if intransitive, see (52b)). AV transitive verbs are not intranstivized nor antipassivized; while lacking the ergative proclitics, they remain fully transitive (examples (50b), (53b), (54)). Compare the following UV and AV counterparts:

(50) (a) UV: *i=andu=o* *ana-ku* *i* *Ali*
3SG.ERG=massage=3SG.ABS child-1SG.GEN PN Ali
'Ali massaged my child'

(b) AV: *i* *Ali* *um-andu=o* *ana-ku*
PN Ali AV-massage=3SG.ABS child-1SG.GEN
'Ali massaged my child'

AV is obligatory in a number of constructions; for example, 1. in relative clauses when relativizing on subjects; indeed examples could be analyzed as obligatory usage of AV in agentive cleft constructions as is typical in Mayan languages:

(51) (a) *hina=o* *mia* *[(anu)* *l-um-ako* *uwoi]*
exist-3SG.ABS person REL go.[AV] water
'there was someone who went to the water'

(b) *saa* *[me-wela* *m-po(N)-rako* *mia]*
crocodile AV-regularly AV-ANTI-grasp person
'a crocodile which regularly takes people'

2. in XCOMPs with controlled subjects:

(52) *me-lulu* *i* *Oloe* *[Ø-buta=o* *apali]*
AV-leap PN Sun AV-extract=3SG.ABS banyan.tree
'Sun leapt to extract a banyan tree'

3. in a verb series with shared subjects:

(53) (a) *mansa=mo* *l-um-ako* *um-ala=o*
at.once=PERF go.[AV] AV-get=3SG.ABS
'right then he went getting it

(b) *l-um-ako* *ira* *me-lempa* *r-um-apati-o*
go.[AV] 3PL.ABS AV-walk follow.edge.[AV]-3SG.ABS
koro-no
river-3SG.POSS

4. with future pronouns:

(54) *aku* *n-um-ahu=ko*
1SG.FUT cook.[AV]=2SG.ABS
'I'm going to cook you' originally 'it will be me who will cook you'

AV and UV voices do not affect the core grammatical relations, subject and object, but the other two voices do. In keeping with crosslinguistic generalizations, passive demotes the subject of a transitive verb to adjunct status, while antipassive does the same to the object. As long as the verb is not a derived ditransitive applied from, the complexities of which I cannot deal with here, the results of both passive and antipassive are intransitive verbs. Passives can only be derived from UV verbs, i.e., they never co-occur with *–um-* or its allomorphs. Passive verbs are marked by the infix *–in-*; their agentive adjuncts can only be indicated by a genitive agreement suffix to the passive verb or if in future tense, the future pronouns:

(55) (a) *ta* *p-in-epate*
3SG.FUT kill.[PASS]
'he will be killed'

(b) *p-in-aho-mu=mo* *ke* *pae* *arau*
plant.[PASS]-2SG.GEN-PERF Q rice that
'was that rice planted by you?'

Passives in keeping with typical Western Austronesian practice are used to relativize objects (56a), although not obligatorily (56b):

(56) (a) *punti* *(anu)* *p-in-aho-do*
banana REL plant.[PASS]-3PL.GEN
'bananas planted by them'

(b) *mia* *anu* *nahi* *ku=to'ori=o*
person REL NEG 1SG.ERG=know=3SG.ABS
'someone that I don't know'

The antipassive is marked by *poN-* (the *N* is realized as a homorganic nasal before a voiceless obstruent; otherwise it is dropped). The object is realized as an adjunct, but surprisingly has no overt oblique case marking. The double agreement for subject and object of the corresponding erstwhile transitive verb now appears with single agreement, absolutive or ergative, depending on the surrounding syntactic environment, but always for the subject of the derived intransitive verb. The antipassive is employed when the object is indefinite or partitive, or discoursally backgrounded for some reason:

(57) (a) *i=pepate='ira* *ana-no*
3SG.ERG-kill=3PL.ABS child-3SG.GEN
'she killed her children'

(b) *m-pom-pepate='ira* *manu*
PL-ANTI-kill-3PL.ABS chicken
'they killed a chicken'

Antipassives are found in both UV and AV, but are obviously more common in the latter, as the semantics of UV and antipassive are somewhat in conflict, although antipassives in UV do occur:

(58) (a) *omue=mo* *koa* *po-wala* *bangka* *atuu*
2SG=PERF just ANTI-chop boat that
'you just go ahead and chop such a boat'

(b)
(i) UV
kuro *atuu* *do=m-pokoli-ako=no* *wuku* *uho*
pot that 3PL.ERG=PL-put-APPL=3SG.ABS shell periwinkle
'in that pot they put the periwinkle shells'

(ii) UV + ANTI
kuro atuu do=m-pom-pokoli-ako=no
pot that 3PL.ERG=PL-ANTI-put-APPL=3SG.ABS
wuku uho
shell periwinkle
'in that pot they put periwinkle shells'

The absolutive agreement in (58b) is for the applied object *kuro atuu* 'that pot'. Note the meaning difference in the definiteness of the theme argument between (58bi) and (58bii). Compare AV verbs with antipassives:

(59) (a) *m-po-nako='ira* *manu* *nana'ote*
PL-ANTI-steal-3PL.ABS chicken children
'the children steal chickens'

(b) *o* *haki* *m-(p)om-pepate* *hadio* *mia*
TOP disease AV-ANTI-kill many person
'the disease killed many people'

Surprisingly, provided the verb is ditransitive, for instance, as a derived applied verb, both passive and antipassive can apply to the same verb:

(60) *p-in-o-'isa-ako=aku* *inisa*
ANTI.[PASS]-pestle-APPL-1SG.ABS pestle.rice
'rice was pestle for me'

and adjunct NPs resulting from antipassive can still head relative clauses:

(61) *inahu* *anu* *i=p-in-om-pahoari-ako* *bonde*
greens REL 3SG.ERG=ANTI.[PASS]=plant.in-APPL garden
'the greens which were planted in the garden'

The antipassive is presumably resorted to here because *inahu* 'greens' are a partitive collective.

The structure of the NP in Mori Bawah is not complicated. As the language is left-headed, only proper name markers or the topic marker precede the head noun. Nouns are suffixed by the genitive pronominal suffixes for possessors, and any full NP possessor must be adjuncts, in apposition to those. There is no inalienable-alienable contrast. Possessive constructions can also be nested:

(63) *lere-mami* *i* *pu'u-no* *torukuno*
field-1PL.EX.GEN LOC base-3SG.GEN mountain
our field at the base of the mountain'

Relative clauses follow all adjunct phrases, both genitive and prepositional phrase adjuncts, and relativization is by gapping of the relativized noun:

(64) (a) *inia-do* *To* *Molungkuni* *anu* *kodei*
village-3PL.GEN people Molungkuni REL small
'the small village of the To Molungkuni'

(b) *mokole* *andio* *[anu* *Ø-mokole-i='ira*
noble this REL AV-rule-APPL-3PL.ABS
mia *i* *inia* *andio*
person LOC village this
'this noble, who ruled over the people of this village'

Like Tolai, nominalizations in Mori Bawah are typically intransitive. Furthermore, nominalizations are mostly straightforward NPs: there is one marginal type of mixed category nominalizations and no D-nominalizations like the Tolai examples of (38). In any case, the latter is impossible in Mori Bawah because the language lacks the functional category D altogether. Nominalizations are formed in two ways in Mori Bawah: by no overt nominalization morphology, simply an apparent verb form in what is normally a nominal syntactic position (this could either be analyzed as M-nominalization by conversion, i.e. by a zero morpheme, or D-nominalization by a zero D element, but given the fact that there are no D elements in the language at all, the latter seems rather extravagant), or by the suffix *–a*, originating from a collapse of PAN **-ən* and **-an.* The former usually denotes an event, result or occasionally patient nominalization, while the latter typically a locative or temporal nominalization:

(65) (a) (i) *da* *lako-no*
still go-3SG.GEN
'still on his outward trip'

(ii) *piso* *pon-sumbele*
knife ANTI-slaughter
'knife for slaughtering'

(b) (i) *lako-a-do*
go-NOMZ-3PL.GEN
'their destination'

(ii) *balo* *po-walu-a*
bamboo ANTI-cook.rice-NOMZ
'bamboo for cooking rice in'

Some verbs can occur with or without *–a* with no difference in meaning (66a), while in other cases there is a clear semantic contrast (66b):

(66) (a) *na-hina* *tenangi-(a)-no*
NEG-exist defeat-(NOMZ)-3SG.GEN
'he can't be defeated'

(b) (i) *hina=o=mo* *pe'wali-a-do*
exist=3SG.ABS=PERF wage.war-NOMZ-3PL.GEN
To *Isareli* *ka* *To* *Pelisiti*
people Israel and people Philistines
'there came a time of war between the Israelites and Philistines'

(ii) *pe'wali-do* *To* *Romaka* *To* *Iahudi*
wage.war-3PL.GEN people Rome and people Jew
nahi komba *para* *lahi* *tehine*
NEG by.no.means endure exceed long.time
'the war between the Romans and Jews didn't last long'

Furthermore, there is a contrast between nominalizations with the passive infix *–in-* and those with *–a* along the lines of realized versus unrealized events, a striking parallel with Tagalog *–an* nominalizations with and without *–in-* and what we shall see in Puyuma below (there are even vestiges of this contrast in Tolai as well):

(67) (a) *m-in-ama* — chew.betelnut.[PASS] — 'that which is chewed' = 'betelnut'
mama-a — chew.betelnut=NOMZ — 'what which is to be chewed' = 'ingredients for a betelnut mix'

(b) *k-in-aa* — eat.[PASS] — 'that which is eaten' = ' 'food'
kaa-nga — eat-NOMZ — 'that which is to be eaten' = 'foodstuffs'

The syntactic structure of Mori Bawah nominalizations is that of NPs, and nominalized verbs are normally intransitive, i.e. subcategorize for no complements, although there is evidence that one type does subcategorize for an object complement. The nominalization of intransitive verbs is straightforward; the genitive pronominal suffix, if present, always refers to the subject of the verb, which otherwise would be realized as an ergative or absolutive clitic, depending on syntactic context:

(68) (a) *pe-booli-no* *i* *Tanggasi*
ITR-call-3SG.GEN PN Tarsier
'Tarsier's call'

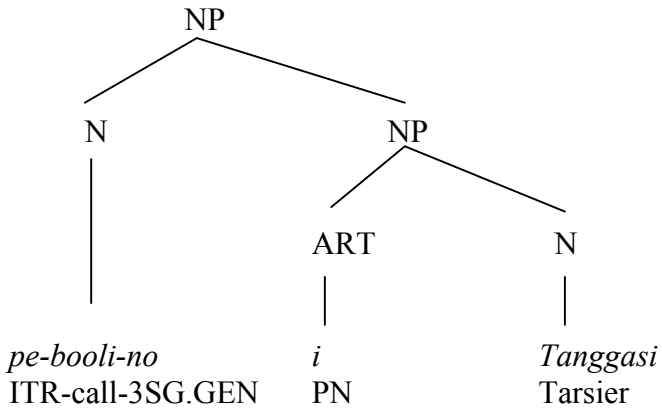
(b) *pe-kuu-ku*
ITR-dive-1SG.GEN
'my dive'

(c) *asa-mbali-no* *koa* *pe-kempa-ku*
one-side-3SG.GEN just ITR-walk-1SG.GEN
'I can only walk on one side'

(d) *me'aa-no*
have.holes-3SG.GEN
'the hollow portion of it'

(e) *na=m(o)=i* *hina* *tuwu-do*
NEG=PERF=3SG.ERG exist live-3PL.GEN
'there were no longer any of them living'

(69) = (68a)



Transitive verb roots are rather more complex. They can occur either in their bare root form (with or without the overt nominalizer *–a*) or with the antipassive prefix *poN-*, which overtly derives an intransitive verb from the transitive root. Not all transitive roots can be nominalized in their bare root from. It seems fairly lexically stipulative, although verbs of cutting and separation as a class seem to favor it, a group of highly canonical transitive verbs, with strongly affected patientive object complements. But in keeping with the deep ergativity of Mori Bawah, when bare root transitive verbs are nominalized, it is their object complements, and not their subjects, which is indicated by the genitive pronominal suffix:

(70) (a) *nahi* *komba* *ondalo* *keke-no*
NEG by.no.means deep dig-3SG.GEN
'not deeply dug' literally 'not at all deep of its being dug'

(b) *nahi* *moiko* *nahu-no* *inahu* *andio*
NEG good cook-3SG.GEN greens this
`the cooking of these vegetables is not good'

(c) *moroso* *koa* *ke* *tanggo-no*
tight just Q bind-3SG.GEN
'are its bindings tight?'

(d) *pudu-no*
break-3SG.GEN
'the broken off piece'

It is not possible at all to express the agent with these bare root nominalizations of transitive verbs; that requires prefixation of *poN-* ANTI. This pattern is very similar to what we found in Tolai in (32), although Tolai is underlyingly a stative-active language in contrast to Mori Bawah's deep ergativity. But it should be noted that even robustly accusative English sometimes exhibits the same pattern in nominalizations: *the army's destruction* only has one reading, with the army as object of destroying, although perhaps more commonly such nominalizations are labile: *the man's cooking*, with the subject reading favored, but object reading possible, or *the man's questioning*, with either reading equally plausible. Genitive pronominal suffixes coreferential to the object complement are obligatory in bare root nominalizations of transitive verbs, but not intransitive verbs:

(71) (a) *petao-a*
marry-NOMZ
'marriage'

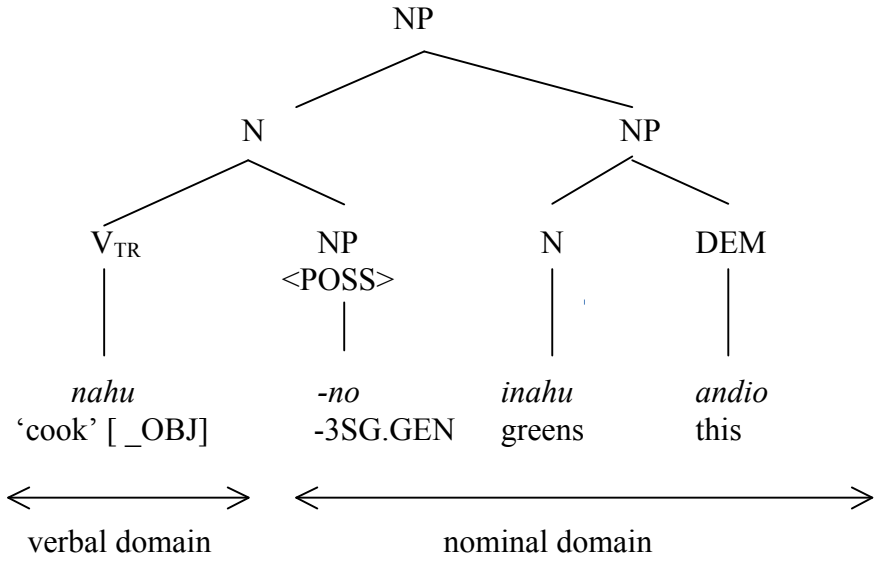
(b) *pe-sala-a*
ITR-travel-NOMZ
'road'

(c) *tanggo-*(no)*
bind-3SG.GEN
'its binding'

(d) *nahu-a-*(no)*
cook-NOMZ-3SG.GEN
'its place of being cooked'

This strongly suggests that the verb root, albeit nominalized, still subcategorizes for an object complement. This seems to be a candidate for a mixed category, shared head nominalization so a structure along the lines of (72) seems correct:

(72)



V_{TR} NOMZ ⇒ (↑ POSS) = (↑ OBJ)

If a transitive verb with both an overt subject and object is nominalized, then intransitivization via the antipassive is obligatory, removing the object from complement function and realizing it as an adjunct. Unlike English and many other languages, there is no marking, prepositional or otherwise (prepositions in Mori Bawah only mark locative notions), to mark its adjunct status, merely the fact that it cannot be indicated on the verb by absolutive pronominal agreement. The only agreement possible is the genitive pronominal suffixes, but they indicate the subject:

(73) (a) *po-doa-no* *i* *Nggasi* *buaea*
ANTI-count-3SG.GEN PN Tarsier crocodile
'Tarsier's crocodile counting'

(b) *po-nahu-do* *inahu* *andio*
ANTI-cook-3PL.GEN greens this
'their cooking of these vegetables'

(c) *pon-to'ori-a-mu* *lele* *atuu*
ANTI-know-NOMZ-2SG.GEN news that
'your knowing of that news'

(d) *po-'ala-mami* *uwoi*
ANTI-carry-1PL.EX.GEN water
'our carrying of water'

Like nominalizations of transitive verbs and again in pointed contrast to bare root nominalizations of transitive verbs, genitive pronominal agreement is not obligatory with antipassivized nominalizations:

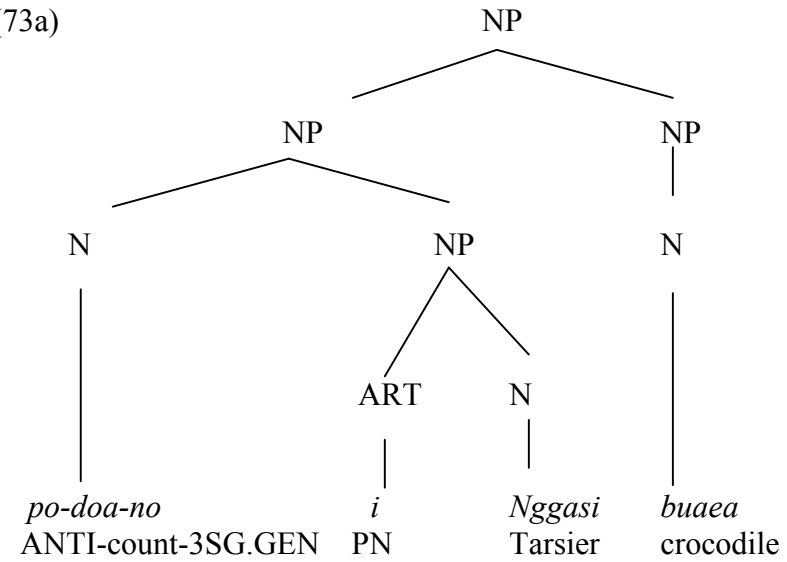
(74) (a) *po-wutu* *asa* *m-pali*
ANTI-bind one C-loop
'binding wrapped one time'

(b) *pom-pa-ka* *boe*
ANTI-CAUS-eat pig
'pig fodder'

(c) *sangka* *m-po-nahu*
utensil C-ANTI-cook
'utensil for cooking with'

This indicates that the antipassivized nominalized verb actually subcategorizes for no arguments like nominalized intransitive verb roots:

(75) = (73a)



Finally, Mori Bawah also has nominalizations in the passive voice; these, of course, will be intransitive by definition, ignoring the complications raised by derived applied ditransitive verbs, which are beyond our purview here. Passive nominalizations are most likely to be headless (or sometimes headed) relative clauses, as they typically denote objects rather than events or results (actually objects created as a result of the typical actions denoted by the verb):

(76) (a) *onae* *koa* *p-in-ewowolo-no*
3SG just think.[PASS]=3SG.GEN
'that was her only thought'

(b) *in-isa-mu* *koa* *ntu'u* *ke* *in-isa* *andio*
PASS-pestle-2SG.GEN just really Q PASS-pestle this
'Is this pestled (rice) really what was pestled by you?

Note that, as with passives generally, these passive nominalizations have genitive pronominals to mark the agent, but again they are not obligatory (see example of *in-isa* PASS-pestle in (76b)). The subjects of passive nominalizations can never be marked overtly by pronominals. Many of these passive nominalizations have been reanalyzed as nouns in their own right: *inahu* 'greens, vegetables' <-*in*- + *nahu* 'cook', *m-in-ama*

'betelnut' <-*in*- + *mama* 'chew betelnut', *inomba* 'dried meats' <-*in*- + *'omba* 'salt and dry meat in the sun', as well as *inisa* 'pestled rice' in (75b).

4 Puyuma

Puyuma is typologically much more similar to Mori Bawah than it is to Tolai; all my data here are drawn from Teng (2008) and Teng and Ross (2010). The syntactic structure of Puyuma clauses is quite similar to that of Mori Bawah, with verb initial order and second position clitics marking aspect. One notable difference is the much richer voice system; in keeping with the typical Philippine-Formosan pattern, it has an actor voice and three undergoer voices:

(77) (a) *tr-em-akaw=Ø* *dra* *paisu* *i* *Isaw*
steal.[AV]-3.ABS IDF.ERG money PRP.ABS Isaw
'Isaw stole money'

(b) *tu=trakaw-aw=Ø* *na* *paisu* *kan* *Isaw*
3.ERG=steal-UV1-3.ABS DF.ABS money PRP.ERG Isaw
'Isaw stole the money'

(c) *tu=trakaw-ay=ku* *dra* *paisu* *kan* *Isaw*
3.ERG=steal-UV2-1SG.ABS IDF.ERG money PRP.ERG Isaw
'Isaw stole money from me'

(d) *tu=trakaw-anay=Ø* *i* *tinataw* *kan* *Isaw*
3.ERG-steal-UV3-3.ABS PRP.ABS his.mother PRP.ERG Isaw
'Isaw stole money for his mother'

As (77) illustrates, all UV voice verbs show agreement for the actor via a set of ergative proclitics; these are not used in AV, as the actor in such cases is already indicated by the absolutive enclitics, which is zero for third person. The set of absolutive and ergative clitics is:

Table 4.1: Puyuma Pronominal Clitics

		ABS	ERG
	1	=ku	ku=
SG	2	=yu	nu=
	3	Ø	tu=
	1 IN	=ta	ta=
	1 EX	=mi	mi=
PL	2	=mu	mu=
	3	Ø	tu=

Note that except for second singular and third person, the absolutive enclitics and the ergative proclitics are homophonous. The absolutive forms are used for subjects and the ergative ones elsewhere. There is in addition a set of determiners that mark case, in a typically Philippine-Formosan way, but along the same binary dimension of absolutive versus ergative:

Table 4.2: Puyuma Determiners

	[+proper]		[-proper]	
	[+SG]	[-SG]	[+DEF]	[-DEF]
ABS	i	na	na	a
ERG	kan	kana	kana	dra

Note the collapse of plural proper nouns with definite common nouns. This makes sense semantically as a set of identifiable people named by proper nouns are indeed definite, but not unique, as they would be in the singular. Perhaps a better analysis would be:

Table 4.3: Alternative Analysis of Puyuma Determiners

	Unique	[+DEF]	[-DEF]
ABS	i	na	a
ERG	kan	kana	dra

This would also account for why unique DPs cannot be modified by numerals nor co-occur with relative clauses (Teng 2008:52). Like Mori Bawah and Philippine-Formosan languages generally, in Puyuma AV is obligatory when relativizing the actor of a transitive verb, i.e. it only allows relativization of absolutive DPs:

(78) *kirtrebung=ta* *dra* *ma'idrang-an*
meet=1PL.IN.ABS IDF.ERG old-NOMZ
[dra mag-sangal dra basak]
IDF.ERG AV-carry IDF-ERG sack
'we meet old people who carry packages on their shoulders'

The structure of NPs in Puyuma is extremely interesting. They typically occur as complements of the case marked determiners. But rather than the determiners being restricted to the beginning of the DP, internal members are also case marked with Ds, and word order among these internal constituents is free, although a D must occur phrase initially:

(79) *dra* *matrina* *dra* *utreutrem* *dra* *suan*
IDF.ERG big IDF.ERG black IDF.ERG dog
'a big black dog'

But when a DP is definite, subsequent members can be linked by the connective *na,* homophonous with the definite absolutive determiner, in lieu of the expected case marked determiner:

(80) *kana* *dru* *na* *kur-dikes=driya* *dra* *kiakarunan*
DF.ERG those C get-hold=IMPERF IDF.ERG job
'those who got hold of a job'

In (80) if case agreement held across the DP we would expect *kana* in place of *na*.

Possessors are indicated in Puyuma by the ergative forms, either as proclitics or determiners:

(81) (a) *tu=tiyal* *kana* *unan*
3.ERG=belly DF.ERG snake
'the snake's belly'

(b) *tu=ngalrad* *kana* *ma-a-rengay*
3.ERG=name DF.ERG AV-RED-talk
'the name of the one talking'

Possessed nouns also take case marked determiners like other nouns, but in this case special portmanteau free form pronouns marking both the possessor and the case of the DP are employed:

(82) (a) *nantu* *ngalrad*
DF.ABS:3.ERG name
'his/her/its name'

(b) *kanta* *ruma'*
DF.ERG:1PL.IN.ERG house
'our house'

Nominalizations in Puyuma are of the M-nominalization type, and the morphology of nominalizations is quite complex, much more so than in Tolai and Mori Bawah, but that does not concern us here, as we are only considering the syntactic structures of nominalizations in this paper. Most nominalizations employ a suffix *–an,* typically in combination with other affixes or derivational processes to indicate aspectual and other types of semantic contrasts. In this they are quite verb-like, and in fact there is some correlation between the morphological properties of verb roots and their nominal derivatives, as we find also in Tagalog. Also, as with Tolai and Mori Bawah (not to mention in passing again Tagalog), there is a clear aspect-mood contrast between nominalizations with and without the perfective affix *ni-* ~ *–in-*: *ni-ekan(-an)* PERF-eat-NOMZ 'food already eaten', *a-akan-an* RED-eat-NOMZ 'food to be eaten'; *ni-rengay-an* PERF-say-NOMZ 'things said', *ra-rengay-an* RED-say-NOMZ 'things to be said'; *s-in-enay-an* sing.[PERF]-NOMZ 'songs sung', *sa-senay-an* RED-sing-NOMZ 'songs to be sung', *s-in-a-sa-senay* RED.[PERF]-RED-sing 'songs often sung'.

Three criteria are crucial in determining the syntactic structures of nominalizations in Puyuma. Firstly, free pronominal possessors as in (82) can only take NPs as complements. Secondly, only verbs can assign absolutive case to their complements; complements of nouns, if any, must be in ergative case, as with possessors. And thirdly, Puyuma has distinct negators for nominal and verbal predicates, *ameli* and *adri* respectively:

(83) (a) *adri/*ameli* *saygu=Ø*
NEG/ NEG able-3.ABS
'he/she can't'

(b) *ameli/*adria* *suan*
NEG/ NEGIDF.ABS dog
'it's not a dog'

On the basis of these traits we could analyze the following as a cleft sentence with a nominalization as predicate:

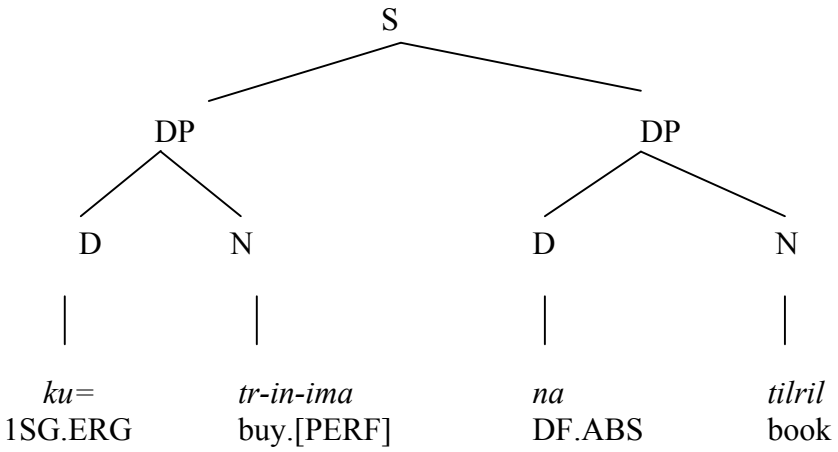
(84) *ku=tr-in-ima*
1SG.ERG=buy.[PERF]
nanku
DF.ABS:1SG.ERG

tr-in-ima
buy.[PERF]

na
DF.ABS

tilril
book

'the book was my buying' = 'I bought the book'

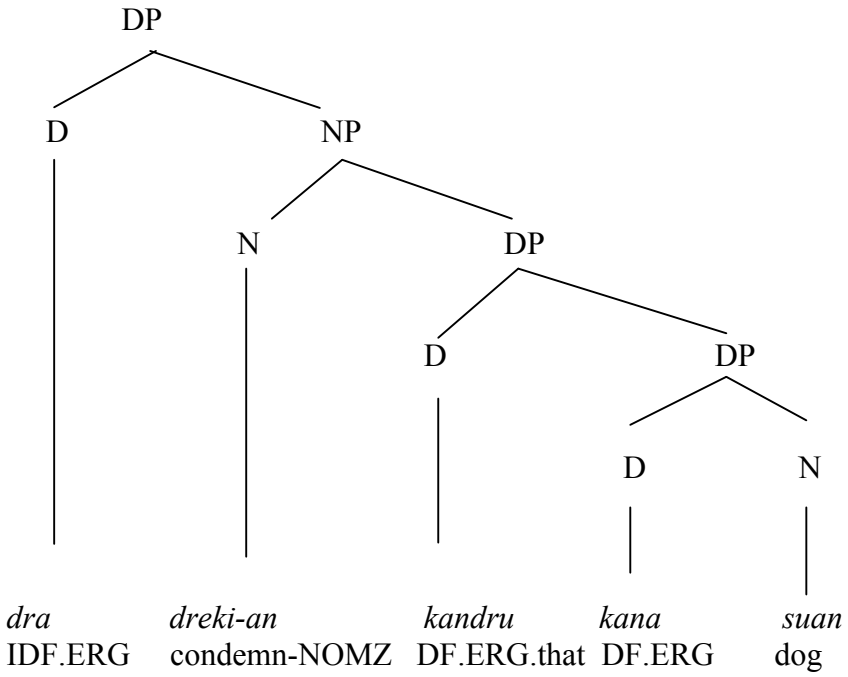


Such nominalizations are also negated with *ameli,* not *adri*:

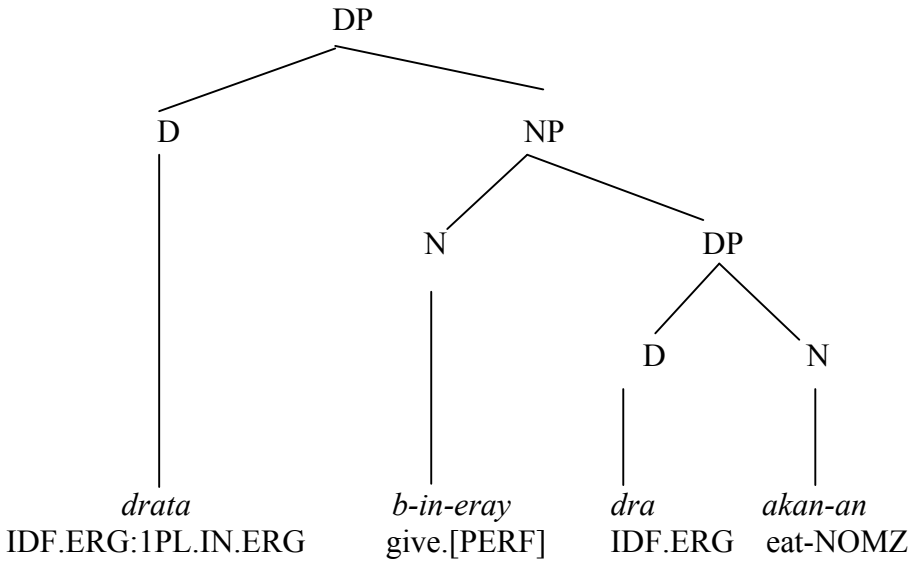
(85) *ameli nantu* *ni-ladra-ladram*
NEG DF.ABS:3.ERG PERF-RED-know
'not something they know'

Furthermore, erstwhile complements of transitive verbs when nominalized like this appear as adjuncts in ergative case, just as overt possessors do:

(86) *dra* *dreki-an* *kandru* *kana* *suan*
IDF.ERG condemn-NOMZ DF.ERG.that DF.ERG dog
'any condemnation of that dog'

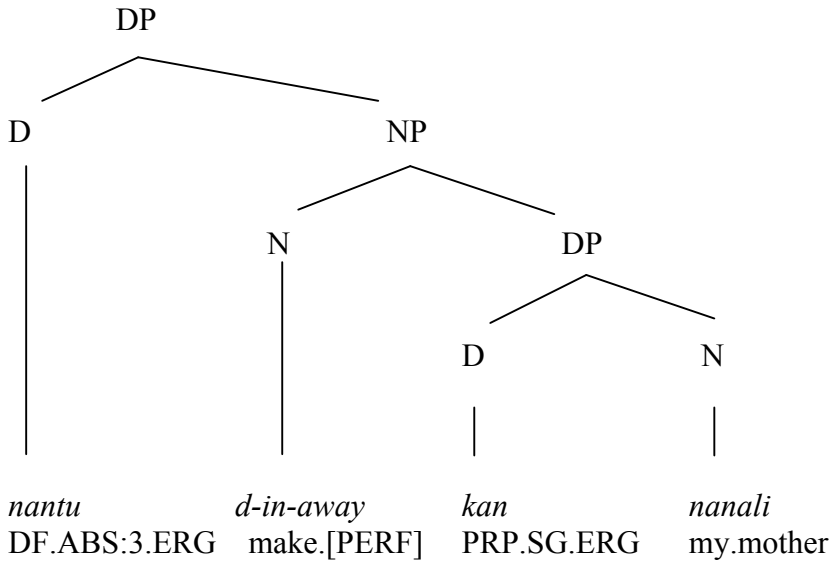


(87) *drata* *b-in-eray* *dra* *akan-an*
IDF.ERG:1PL.IN.ERG give.[PERF] IDF.ERG eat-NOMZ
'our giving of food' = 'whatever food we have given'



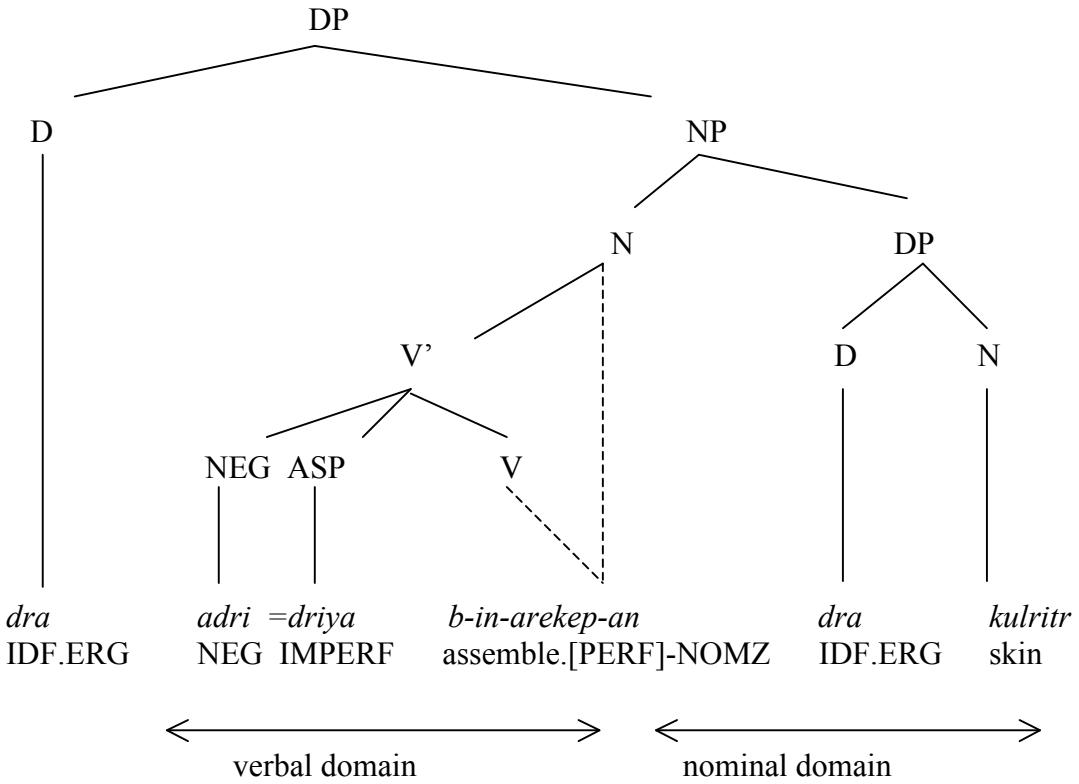
Furthermore, these adjuncts of nominalized verbs are not restricted to objects of transitive verbs, as subjects can appear there as well:

(88) *nantu* *d-in-away* *kan* *nanali*
DF.ABS:3.ERG make.[PERF] PRP.SG.ERG my.mother
'those makings (clothes) by my mother'



While this fully nominal M-nominalization type is the most common kind in Puyuma, there are other types, albeit rarer, as well. One is a mixed category, shared head type of M-nominalization in which a V' is embedded under an NP and where the V' and NP share the same head. This is clearly diagnosed by the use of the verbal negator *adri* instead of nominal *ameli*:

(89) *dra* *adri=driya* *b-in-arekep-an* *dra* *kulritr*
IDF.ERG NEG=IMPERF assemble.[PERF]-NOMZ IDF.ERG skin
'that which has not yet been assembled with a skin'



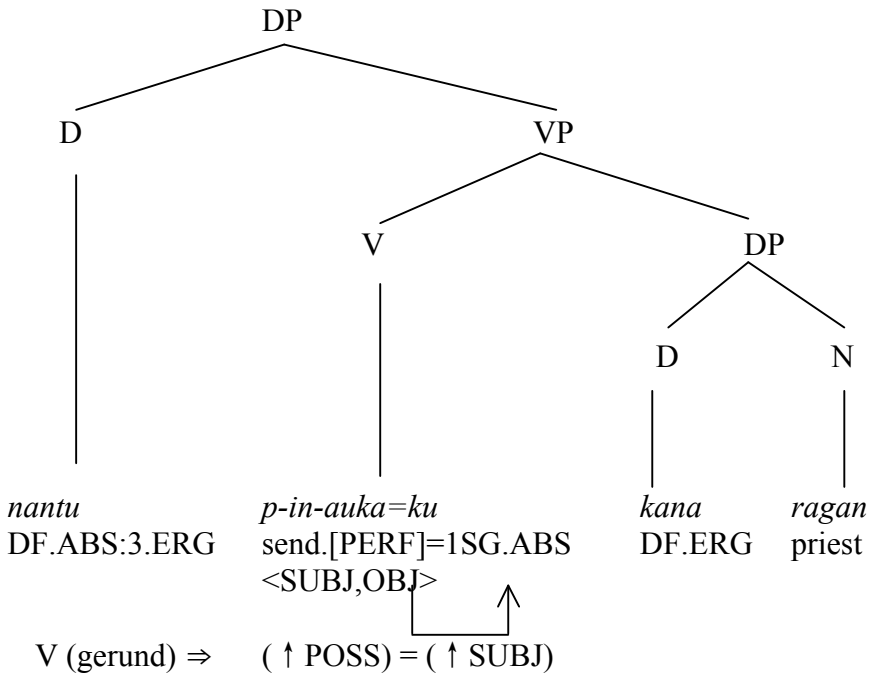
This structure provides an explanation for an anomaly noted by Teng (2008:143), namely, the lack of an ergative possessive pronominal preceding the nominalized verb. The explanation is obvious: such a pronominal would be inside the verbal projection between the modifiers and its head, and because this is strictly a verbal domain, it will not license a nominal modifier like a possessive pronominal. Note that a structure like (89) casts doubt on Teng and Ross's (2010) claim that Puyuma has not undergone the reanalysis of nominalizations to verbs, one of the innovations that they use to define Nuclear Austronesian, and from which they therefore exclude Puyuma. The analysis of (89) clearly indicates that *b-in-arekep-an* assemble.[PERF]-NOMZ is synchronically a verb in Puyuma, whatever its earlier status.

More problematic for Teng and Ross' (2010) claim is the following example, which illustrates a type of D-nominalization in Puyuma:

(90) *nantu* *p-in-auka=ku* *kana* *ragan*
DF.ABS:3.ERG send.[PERF]=1SG.ABS DF.ERG priest
'the priest's sending of me'

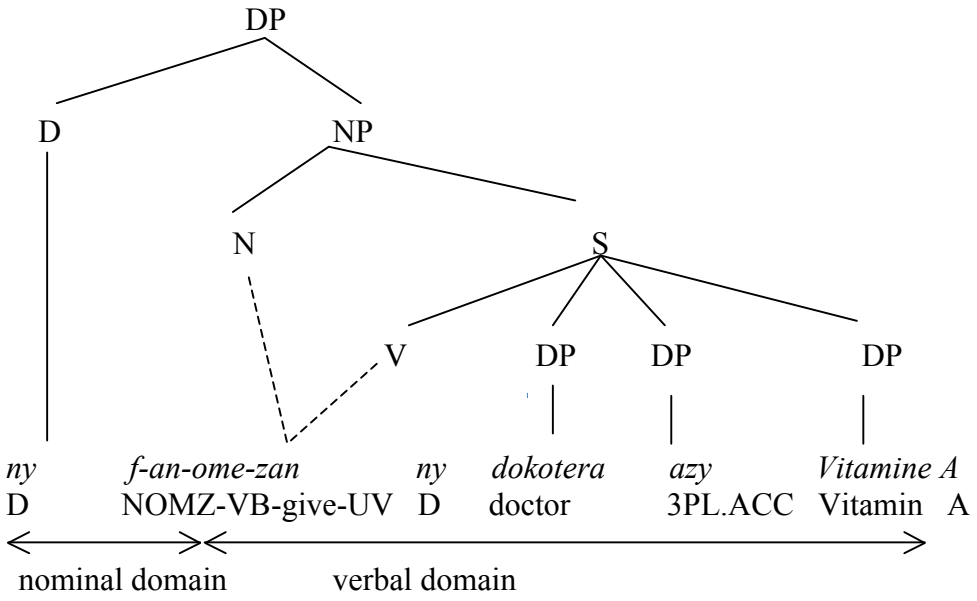
The crucial point to note here is the absolutive enclitic on the nominalized verb. Only verbs can assign absolutive case, and transitive verbs such as *pauka* 'send' assign it to their object complements. In order for (90) to be possible, *pauka* 'send' must preserve its verbal status and its ability as a transitive verb to assign absolutive case to its object complement. The structure of (90) is parallel to that of English gerunds and Tolai D-nominalizations in (38), i.e. a VP dominated by a DP:

(91)



D-nominalizations are only possible if the verb or verb phrase being nominalized is already verbal. In the light of (91), Teng and Ross' (2010) claim that Puyuma has not reanalyzed nominalizations as verbs sees problematic. Although it is arguable whether they appear in main clauses (although given (91), the cleft analysis of (84) now looks decidedly shaky), they certainly have other crucial and **unique** syntactic properties of verbs. Other Western Austronesian languages clearly extend nominalizations to permit the embedding of full S constituents, as in Malagasy (Ntelitheos 2010):

(92) *ny* *f-an-ome-zan'* *ny* *dokotera* *azy* *Vitamine A*
D NOMZ-VB-give-UV2 D doctor 3PL.ACC Vitamin A
'the doctor's giving of Vitamin A to them'



5 Comparative Notes

Moyse-Faurie (2007) is the only previous attempt at a comparative look at nominalizations in Austronesian, and she confined herself to Oceanic languages. Still, some striking parallels between her data and the analyses provided here are strong confirmation of these. Moyse-Faurie (2007) established three basic types of nominalizations in Oceanic languages, all of which are paralleled by the three languages discussed here. An interesting comparison to Mori Bawah is provided by the New Caledonian language Xârâcùù. In this language verbs are M-nominalized by a set of prefixes to the verbal root. If only a single argument occurs with the nominalized verb, it is marked by the genitive particle:

(93) (a) *kèè-xwèrii* *rè* *nâ*
NOMZ-want GEN 1SG
'my desire'

(b) *kèè-xwèrii* *rè* *mwâ*
NOMZ-want GEN house
'desire for a house'

But if both the subject and object arguments of a nominalized transitive verb are present, one occurs in the genitive case and the other as an adjunct, either unmarked in the case of the object, just as in Mori Bawah, or marked by the agentive particle *ngê,* if the subject. We can determine that the object is now an adjunct because when appearing as a pronoun, it is in the independent form, not the accusative one (94b):

(94) (a) *kèè-xwèrii* *rè* *nâ* *mwâ*
NOMZ-want GEN 1SG house
'my desire for a house'

(b) *kèè-xâpârî* *rè* *nâ* *gè/•ro*
NOMZ-see GEN 1SG 2SG/2SG.ACC
'my seeing of you'

(c) (i) *kèè-söömè* *rè* *panèè-rè* *nie*
NOMZ-kill GEN father-3SG.GEN 3SG
'his father's killing of him'

(ii) *kèè-söömè* *rèè* *panèè-rè*
NOMZ-kill 3SG.GEN father-3SG.GEN
'his killing of his father'

(d) *kèè-toanôô* *rèè* *ngê* *gè*
NOMZ-find 3SG.GEN AGT 2SG
'your finding of him'

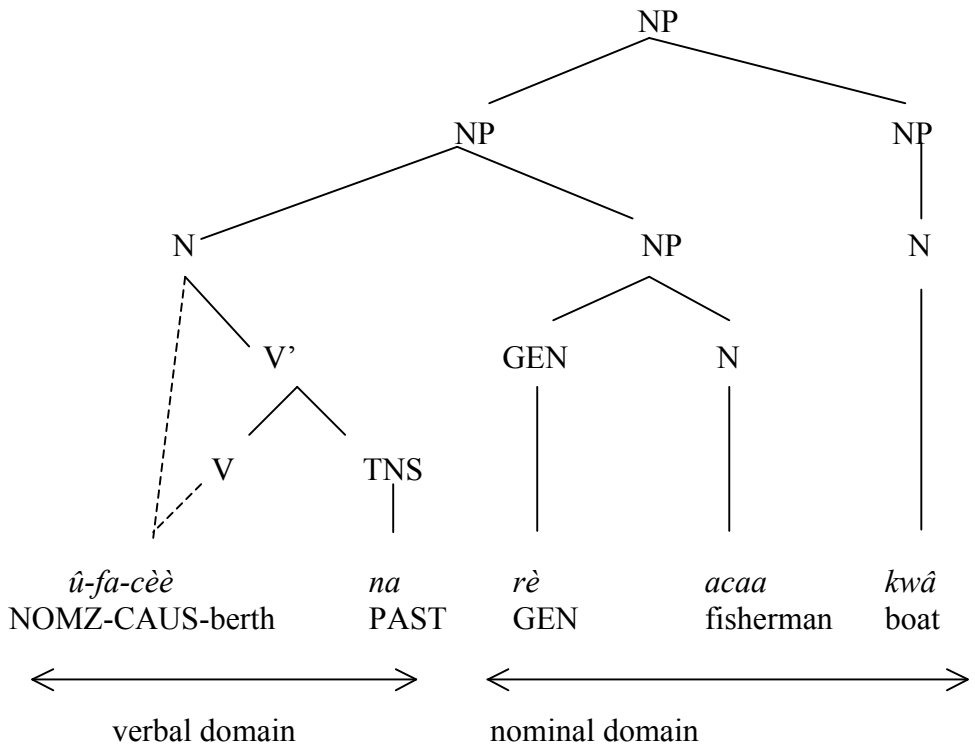
Moyse-Faurie points out that in another New Caledonian language Nêlêmwa, M-nominalizations like (94d) are the norm; the object is realized as an accusative pronominal suffix on the verb and the subject as an adjunct with an oblique or genitive marker (Bril 2002):

(95) (a) *u-no-mwimwi-na* *nai* *hla*
NOMZ-see-know-1SG.ACC OBL 3PL
'their manner of recognizing me'

(b) *u-axi-yo* *i* *na*
NOMZ-see.TR-2SG.ACC GEN 1SG
'my sight of you'

These structures are very similar to the nominalizations of Mori Bawah in (74). However, unlike Mori Bawah, in which such nominalizations are unproblematic NPs, in Xârâcùù and Nêlêmwa, they are mixed categories, because in the former, nominalized verbs can co-occur with tense marking particles, while in the latter, as (95) demonstrates, nominalizations of transitive verb roots continue to apply accusative case to their object complements, and both of these are properties of verbs, not nouns. Here is a Xârâcùù example of a tensed M-nominalization:

(96) *ù-fa-cèè* *na* *rè* *acaa* *kwâ*
NOMZ-CAUS-berth PAST GEN fisherman boat
'place of the fisherman's berthing of the boat'



Note that unlike Mori Bawah, there is no formal derivation which intransitivizes the verb, although through the process of nominalization it loses the ability to assign case, particularly accusative case to its object complement, which thereby must be realized as an adjunct. In Nêlêmwa, however, this is not the case, because object complements of nominalized verbs continue to be accusative pronominal suffixes; subjects, on the other hand, are marked by oblique cases.

Polynesian languages have particularly interesting nominalization constructions, and the data here is mainly from East Uvean, again drawn from Moyse-Faurie (2007). A notable point of contrast between Polynesian and Xârâcùù is that, while nominalizations can be marked for aspect, as in this East Futuna example:

(97) *le koi no-nofo a Malia mo lonaa tinana*
D PROG RED-stay POSS Malai and 3SG.GEN mother
'Malia's staying with her mother'

Inflection for tense is prohibited. The D-nominalizations of East Uvean are similar to those of Tolai, except that again there is no explicit derivational process of intransitivization. As with Tolai, there is a split between unaccusative and unergative intransitive verbs in nominalization; the former mark the genitive adjunct DP with the subordinate possessive classifier *o* and the latter with the dominant classifier *a*:

(98) (a) unaccusative
(i) *'e kokolo Ø te 'ua*
NPAST noise ABS D rain
'the rain makes noise'

(ii) *te kokolo **o** te 'ua*
D noise POSS D rain
'the noise of the rain'

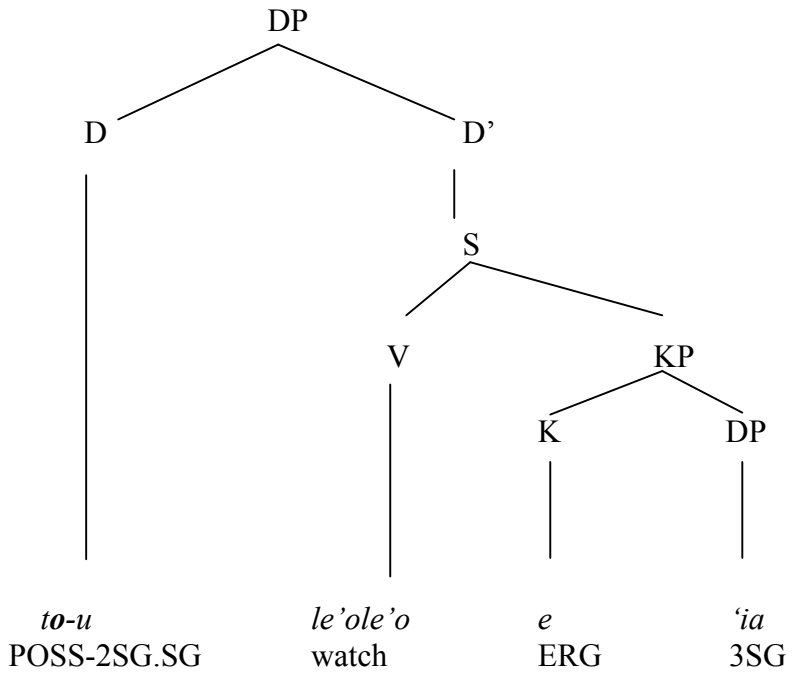
(b) unergative
(i) *'e lele Ø te hosi*
NPAST run ABS D horse
'the horse is running'

(ii) *te lele **a** te hosi*
D run POSS D horse
'the running of the horse'

With canonical transitive verbs, which ergatively case mark their subjects and absolutively their objects, the possibilities depend on whether the DPs are pronominal or not. If both are pronominal, either subject or object can appear as genitive, but the other argument retains its case, ergative or absolutive, as appropriate. In order for this to be possible, particularly the object complement to remain in absolutive case, the verb must retain its case marking properties and subcategorization frame. Hence, these East Uvean nominalizations are again D-nominalizations, essentially DPs dominating S constituents, not too unlike the English gerunds of (8) or Tolai nominalizations of (38) or Puyuma nominalization of (91), but dominating S rather VP:

(99) (a) *t**a**-na le'ole'o ia koe*
POSS-3SG.GEN watch ABS 2SG
'his watching of you'

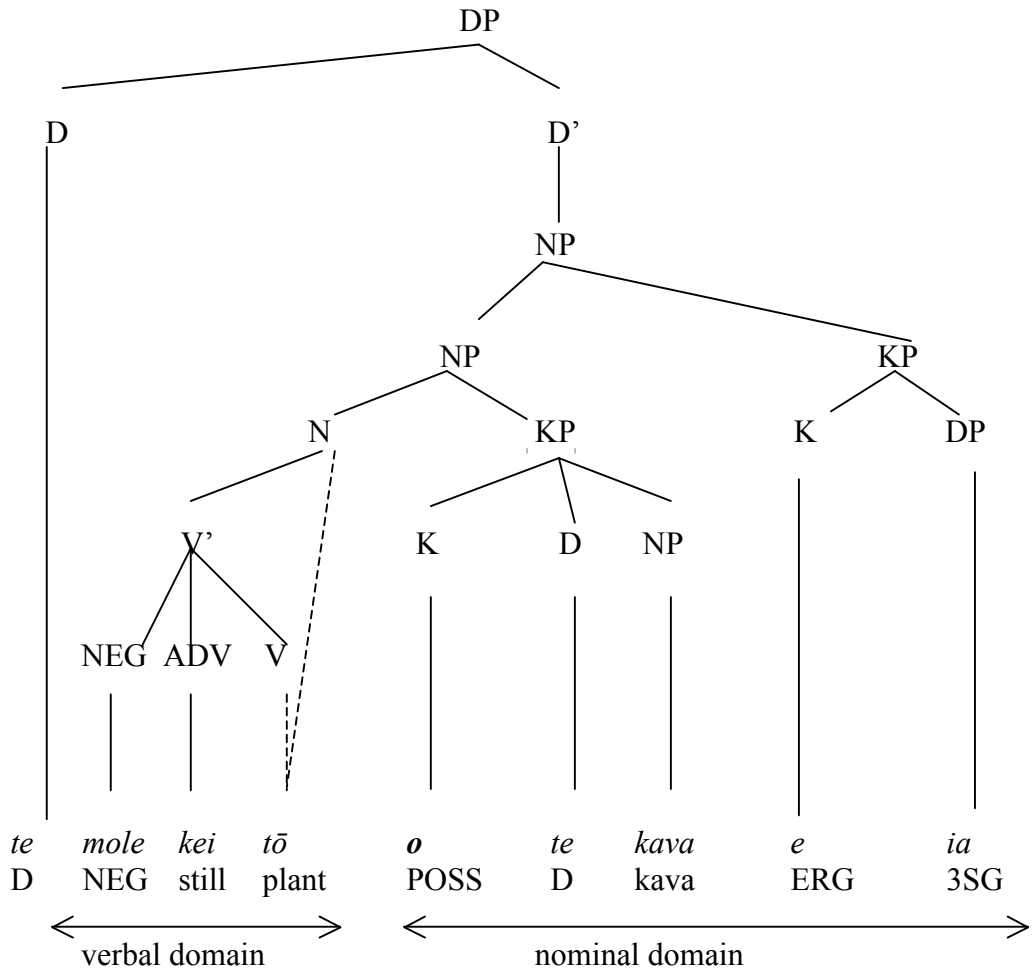
(b) *t**o**-u le'ole'o e 'ia*
POSS-2SG.SG watch ERG 3SG
'your being watched by him'



V (NOMZ) ⇒ (↑ POSS) = (↑ GF)

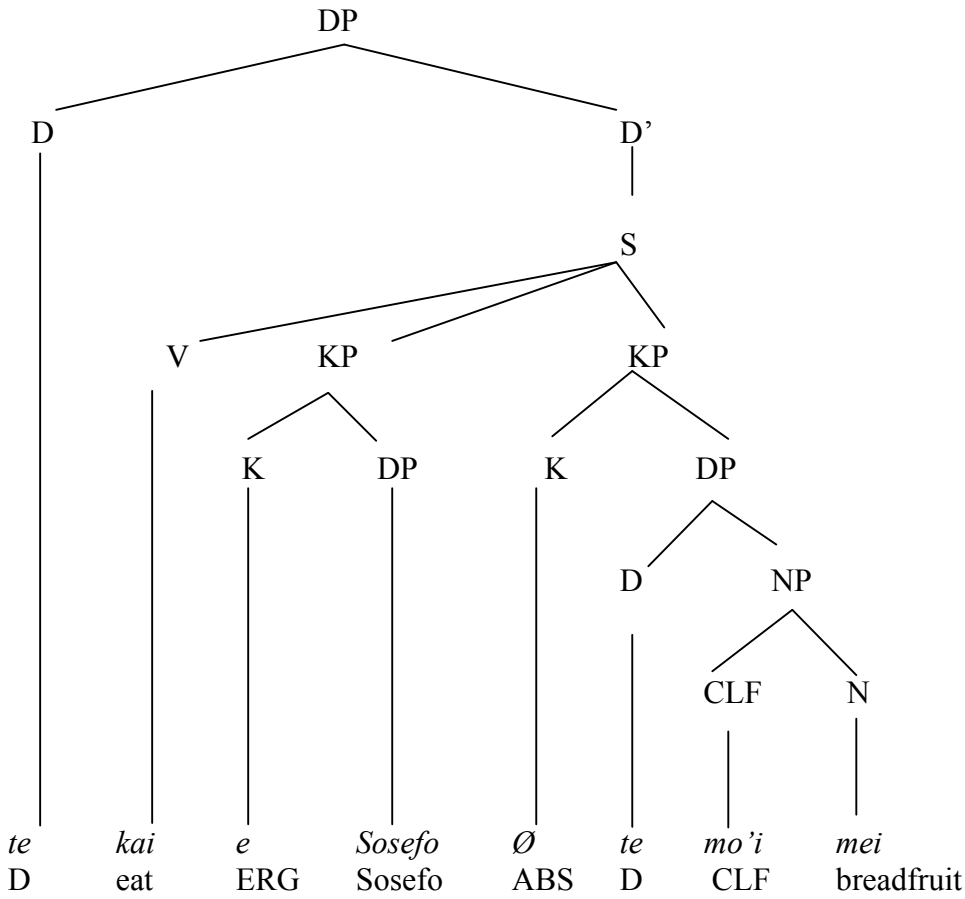
If one DP is pronominal and the other a full nominal, the result depends on argument structure. If the object is the pronominal, then it must occur as the possessive D with the subordinate classifier *o* in the nominalization, which again has the structure in (99). If the subject argument is the pronominal then it (1) either occurs as the possessive D with dominant classifier *a*, and the object has absolutive case, so again the resulting structure is that exemplified in (99), or (2) it remains in ergative case and the object DP assumes genitive case marked with the subordinate classifier *o*. But in this setup, the usual free word order of the subject and object DPs becomes fixed, so that the object in genitive case must immediately follow the nominalized verb, strongly suggesting the structure in question now contains an NP, and is then a mixed category construction:

(100) *te mole kei tō* ***o*** *te kava e 'ia*
D NEG still plant POSS D kava ERG 3SG
'his no longer planting of kava'



If both DPs are full DPs, the subject must remain in the ergative case, but the object can occur in genitive case in *o,* again immediately following the verb, so again in an NP in a mixed category like the structure illustrated in (100); or the object can remain in absolutive case, but also in a fixed constituent structure, so it must follow the ergatively marked subject DP. I have no explanation why the order of constituents is fixed in this S constituent dominated by a DP; there appears to be no compelling syntactic reason why it should be.

(101) *te kai e Sosefo Ø te mo'i mei*
D eat ERG Sosefo ABS D CLF breadfruit
'Sosefo's eating of the breadfruit'



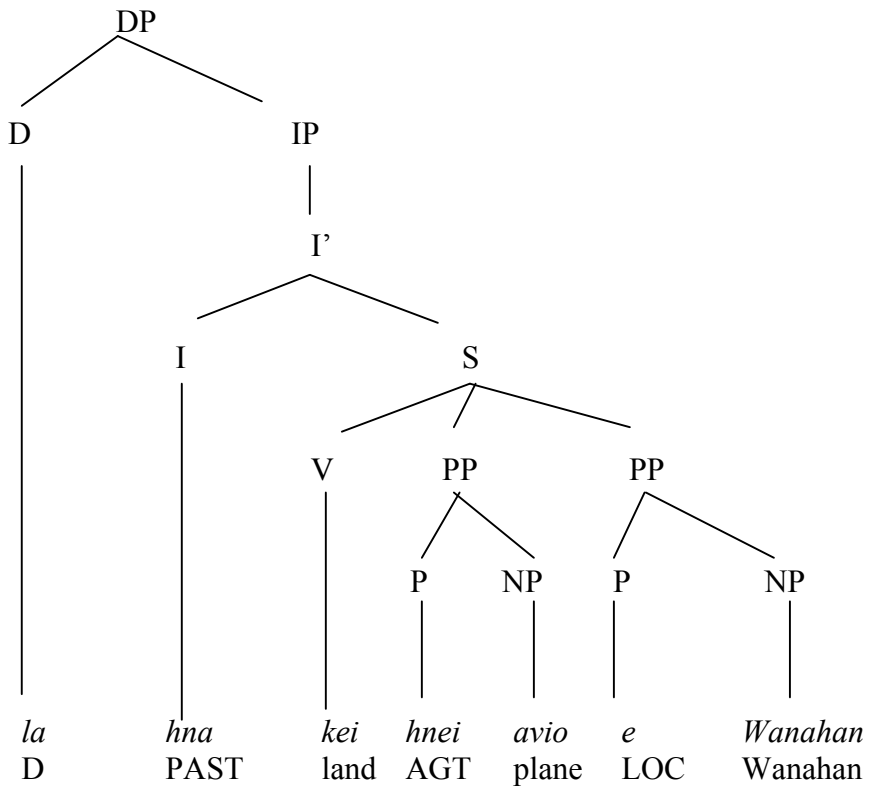
Interestingly, Chung (1973:665) claims that this second option is not available in closely related Samoan; only the pattern with the genitive object, i.e. the structure of (100) is said to be possible in that language. It is worth pointing out that the structures of (99) and (101) are similar to what is found in Tolai in examples like (38), the difference being the nature of the constituent dominated by the DP, S in East Uvean versus VP in Tolai.

6 Conclusion

In surveying the range of nominalization structures we have investigated in this paper, I have demonstrated that they fall into two major types: 1. syntactic nominalization by determiners, D-nominalizations or 2. nominalization by morphological derivation, M-nominalization, although, as in the case of some Tolai nominalizations, there may not be any overt morphological formative to indicate this. Mixed category, shared head nominalizations are found in both types. The first is the more straightforward of the two, and are in general rather like the English gerund of (8). They are formed by a functional head D taking as its complement a verb or its projections: VP (Tolai, example (38)), S (East Uvean, example (101)). Potentially the complement of the D functional head may be a full IP, i.e. a verb fully inflected for tense and with its subcategorized DPs case marked as in main clauses; it may be that the headless relative clauses so common in Philippine-Formosan languages should be analyzed in this way, but to go further into this possibility would take us too far afield right now. But perhaps a clearer example is provided by Drehu of the Loyalty Islands (Moyse-Faurie 2007):

(102) (a) *hna kei hnei avio e Wanahan*
PAST land AGT plane LOC Wanahan
'the plane landed at Wanahan'

(b) *la hna kei hnei avio e Wanahan*
D PAST land AGT plane LOC Wanahan
'the landing (place) of planes at Wanahan' = 'the airport at Wanahan'

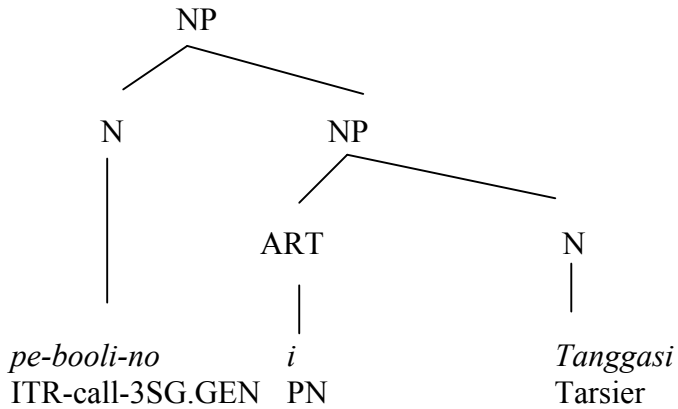


Of course, the option of nominalizations via a complement of the functional head D is only possible in languages with an overt category D, and consequently it is lost in languages like Mori Bawah that lack D elements. Still the pervasiveness of nominalization constructions by D across the Austronesian world is no doubt due to the ancient pedigree and widespread distribution of the functional category D in these languages, and this is a striking point of contrast between the Austronesian language family and the other language families of Southeast Asia and the Pacific. Further, there can be little doubt that this widespread feature of nominalization simply via a functional head D is a major source of the much remarked about categorical indeterminacy of the major lexical classes noun and verb in these languages, and a feature they share with other Pacific Rim language families in which this distinction has been a major point of contention like Salishan and Wakashan.

The second type of nominalization is formed by the lexical category N dominating a verb (Mori Bawah, example (75)), or, again, any of its projections, V' (Xârâcùù, example (96); Puyuma, example (89); East Uvean, example (100)), VP (Nêlêmwa, example (95) and (105) below) or S (Malagasy, example (92)). These parallel in part the English structures of (5). While I have no examples among Austronesian languages, there is no reason to exclude the possibility of a full IP being nominalized in this manner, i.e., a formally morphologically derived nominalized verb with full tense inflection plus all its subcategorized arguments retaining the case marking they would receive in a main clause. Crucially, if during the lexical nominalization process the verb cedes all of its verbal

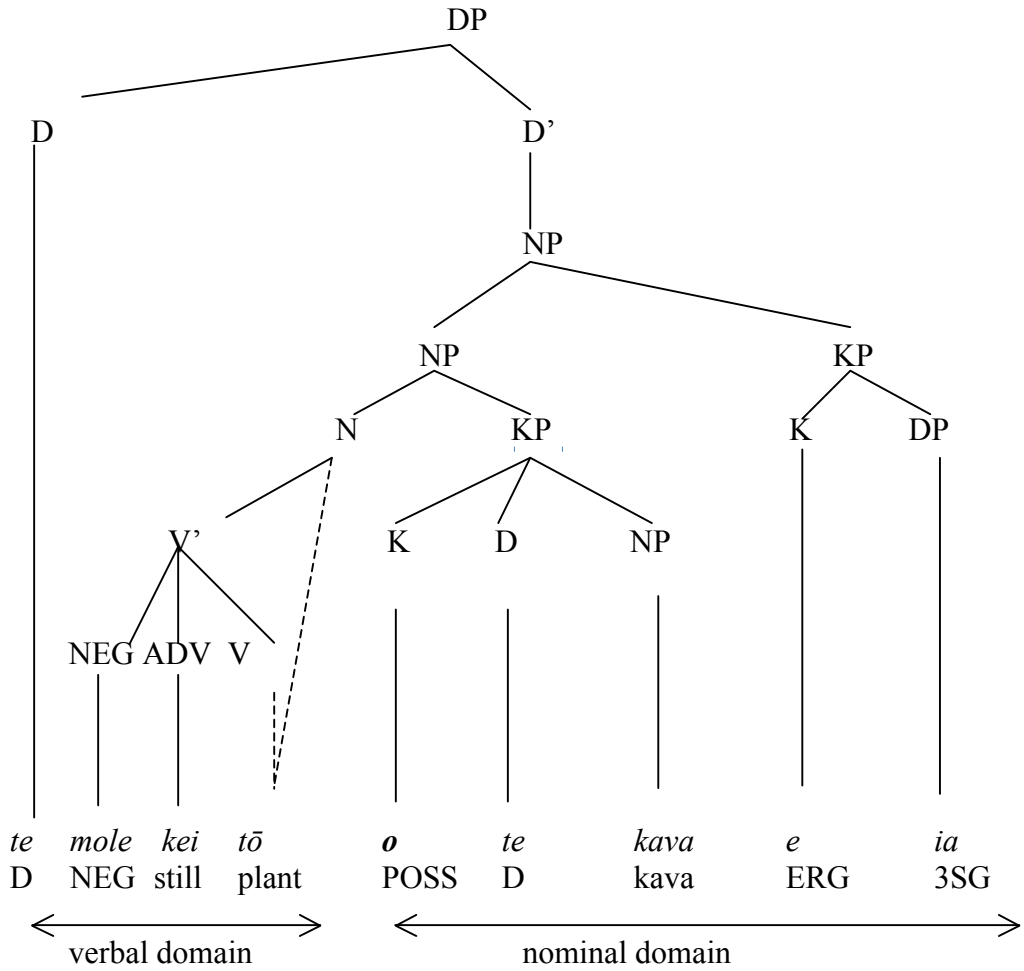
properties, then the head of the nominalization is now lexically and syntactically only a noun, and the syntactic structure of the projected phrase will be an unproblematic and straightforward NP, as in Mori Bawah:

(103) = (69)

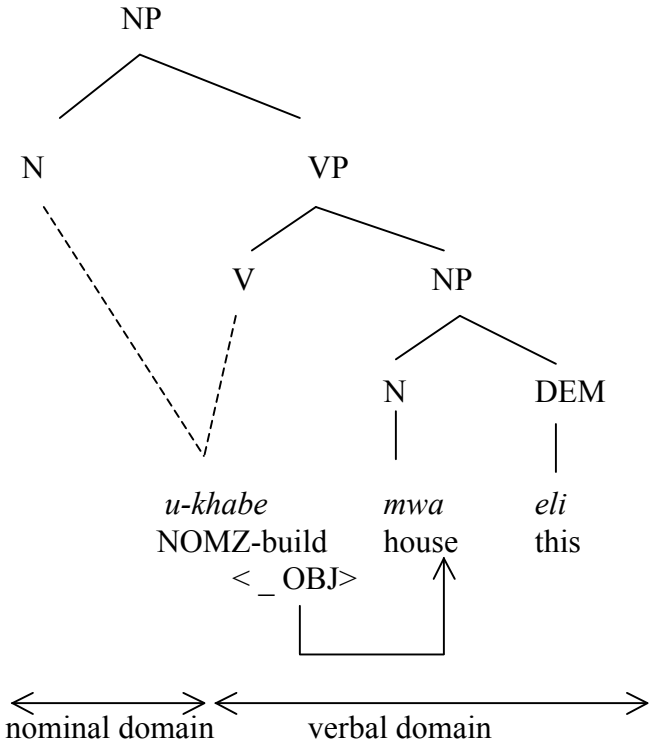


But if the verb retains some of its morphosyntactic properties and/or further undergoes syntactic processes of concatenation with its modifiers, specifiers or complements, as permitted by those retained properties, a phrasal category headed by the verb has now been projected. When the N lexical head of the nominalization now dominates the verb and any associated projections, a mixed category necessarily results, with the verbal properties within the maximal projection of the V head and the nominal properties with the projection of the higher lexical head N:

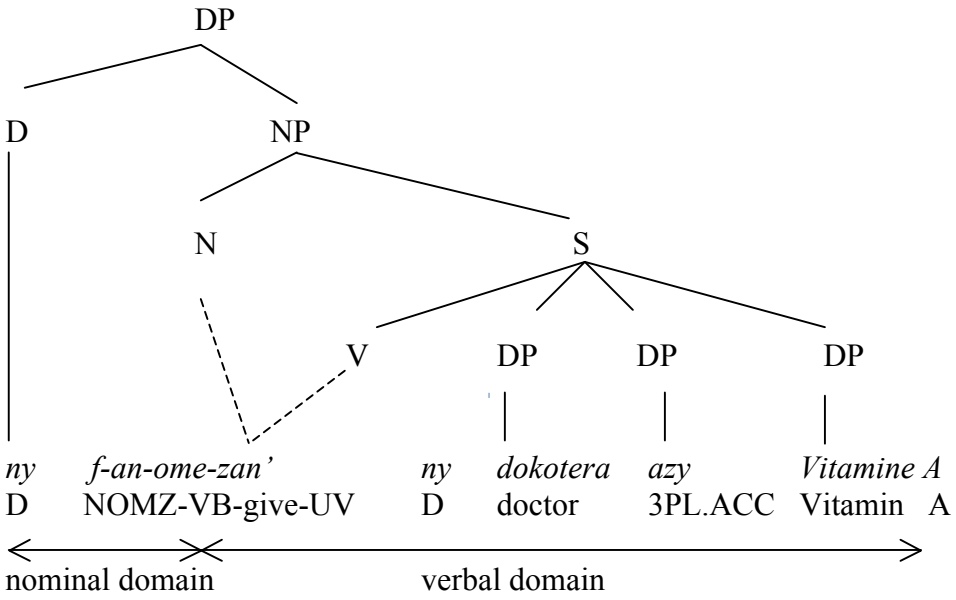
(104) = (100) East Uvean



(105) Nêlêmwa (Bril 2002)



(106) = (92) Malagasy



The syntax of nominalization constructions provides a window into the basic nature of the fundamental lexical classes of human language, verb and noun. We are accustomed to think of them as sharp, all or nothing groupings, but the evidence here suggests that they are gradient (Aarts 2007). Mixed category nominalizations in the languages investigated here show that verbs can preserve their ability to select for tense/aspect and other modifiers like negation, but cede their ability to assign case (Puyuma, East Uvean, Xârâcùù) or they may cede tense/aspect marking, but retain the ability to assign case to object complements (Mori Bawah, Nêlêmwa) or indeed assign case as if they were in main clauses (Malagasy). Further work will define the parameters of variation in the ceding or preserving of various

grammatical properties and in the process greatly enrich our typology of the basic building blocks of human language. Given their number, widespread geographical distribution, typological diversity, and relatively well known historical development, the Austronesian languages afford a vital laboratory in which to study these phenomena.

References

Aarts, B. 2007. *Syntactic Gradience.* Oxford: Oxford University Press.

Aissen, J. 1999. Agent focus and inverse in Tzotzil. *Language* 75:451-485.

Barsel, L. 1994. *The Verbal Morphology of Mori, Sulawesi.* Canberra: Pacific Linguistics.

Bley, B. 1912. *Praktisches Handbuch zur Erlernung der Nordgazellen-Sprache.* Münster: Westfälische Vereinsdruckerei.

Bresnan. J. 1997. Mixed categories as head-sharing constructions. *Proceedings of the LFG 97 Conference, University of California, San Diego.* Online: CSLI Publications: http://www-csli.stanford.edu/publications

Bresnan, J. 2001. *Lexical-Functional Syntax.* Oxford: Basil Blackwell.

Bresnan, J. and Mchombo, S. 1987. Topic, pronoun and agreement in Chichewa. *Language* 63:741-782.

Bresnan, J. and Mugane, J. 2006. Agentive nominalizations in Gĩkũyũ and the theory of mixed categories. In Butt, M., Dalrymple, M. and King, T. eds., *Intelligent Linguistic Architectures: Variations on Themes by Ronald M. Kaplan,* 201-234. Stanford: CSLI Publications.

Bril, I. 2002. *Le nêlêmwa (Nouvelle-Calédonie): Analyse syntaxique et sémantique.* Paris-Louvain: Peeters.

Capell, A. 1964. Verbal systems in Philippine languages. *Philippine Journal of Science* 93:231-249.

Chomsky, N. 1970. Remarks on nominalization. In Jacobs, R. and Rosenbaum, P. , eds., *Readings in English Transformational Grammar,* 184-221. Waltham, MA: Ginn and Company.

Chung, S. 1998. *The Design of Agreement: Evidence from Chamorro.* Chicago: University of Chicago Press.

Coon, J. 2009. Comments on Austronesian nominalism: A Mayan perspective. *Theoretical Linguistics* 35:73-93.

Cooreman, A. 1987. *Transitivity and Discourse Continuity in Chamorro Narratives.* Berlin: Mouton-de Gruyter.

Culicover, P. and Jackendoff, R. 2005. *Simpler Syntax.* Oxford: Oxford University Press.

Esser, S. 2012 [1927-1933]. *Esser's Grammar of Mori.* Translated by David Mead. Dallas: Summer Institute of Linguistics e-book 27.

Foley, W. 1979. Toward a universal typology of the noun phrase. *Studies in Language* 4: 171-199.

Foley, W.1991. *The Yimas Language of New Guinea.* Stanford: Stanford University Press.

Goodenough, W. and Sugita, H. 1980. *Trukese-English Dictionary.* Philadelphia: American Philosophical Society.

Grimshaw, J. 1990. *Argument Structure.* Cambridge, MA: MIT Press.

Jackendoff, R. 2007. *Language, Consciousness, Culture: Essays on Mental Structure.* Cambridge, MA: MIT Press.

Jackendoff, R. 2011. What is the human language faculty. Two views. *Language* 87:586-624.

Kaufman, D. 2009. Austronesian nominalism and its consequences: A Tagalog case study. *Theoretical Linguistics* 35:1-49.

King, T. 1995. *Configuring Topic and Focus in Russian.* Stanford: CSLI Publications.

Kiss, K. 1995. *Discourse Configurational Languages.* Oxford: Oxford University Press.

Kleintitschen, A. 1924. *Mythen und Erzahlungen eines Melanesierstamme aus Paparatava, Neupommern, Südsee.* St. Gabriel-Mödling: Anthropos Institut.

Kroeger, P. 1993. *Phrase Structure and Grammatical Relations in Tagalog.* Stanford: CSLI Publications.

Lanyon-Orgill, P. 1960. *A Dictionary of the Raluana Language (New Britain, Southwest Pacific).* Victoria, B.C.: Published by the author.

Lefebvre, C. and Muysken, P. 1988. *Mixed Categories: Nominalizations in Quechua.* Dordrecht: Kluwer.

Malouf, R. 2000. *Mixed Categories in the Hierarchical Lexicon.* Stanford: CSLI Publications.

Mead, D. 1998. *Proto-Bungku-Tolaki: Reconstruction of its phonology and Aspects of its Syntax.* Unpublished PhD dissertation, Rice University.

Mead, D. 1999. Active, passive and antipassive in Bungku-Tolaki languages. *NUSA: Linguistic Study of Indonesian and other Languages of Indonesia* 45:113-145.

Mead, David. 2005. Mori Bawah. In Himmelmann, N. and Adelaar, A., eds., *The Austronesian Languages of Asia and Madagascar,* 883-708. London: Routledge.

Mead, D. 2008. When to use a genitive pronoun in Mori Bawah (Sulawesi, Indonesia). *Studies in Philippine Languages and Cultures*17:137-178.

Meier, J. 1909. *Mythen und Erzählungen der Küstenbewohner der Gazelle-Halbinsel. (Neu-Pommern).* Münster: Aschendorffschen.

Mosel, U. 1979. Tolai Texts. *Kivung* 10. Port Moresby: Linguistic Society of Papua New Guinea.

Mosel, U. 1984. *Tolai Syntax and its Historical Development.* Canberra: Pacific Linguistics.

Moyse-Faurie, C. 2007. Les formes nominalisées du verbe dans quelques langues océaniennes. *Faits de Langues* 30:97-116.

Rinderknecht, P. 1987. *Nomen und Verb im Melanesischen Tuna (Tolai).* Frankfurt: Peter Lang.

Ross, M. 2009. Proto-Austronesian verbal morphology: A reappraisal. In Adelaar, A. and Pawley, A., eds., *Austronesian Historical Linguistics and Culture History: A Festschrift for Robert Blust,* 295-326. Canberra: Pacific Linguistics.

Schachter, P. and Otanes, F. 1972. *Tagalog Reference Grammar.* Berkeley: University of California Press.

Starosta, S., Pawley, A. and Reid, L. 1982. The evolution of focus in Austronesian. In Halim, A., Carrington, L., and Wurm, S., eds., *Papers from the Third International Conference on Austronesian Linguistics I: Tracking the Travellers,* 145-170. Canberra: Pacific Linguistics.

Teng, S. 2008. *A Reference Grammar of Puyuma, an Austronesian Language of Taiwan.* Canberra: Pacific Linguistics.

Teng, S. and Ross, M. 2010. Is Puyuma a primary branch of Austronesian? A reply to Sagart. *Oceanic Linguistics* 49:543-558.

Tuuk, H. van der. 1971 [1864-1867]. *A Grammar of Toba Batak.* The Hague: Nijhoff.

Wright, L. 1964. *A Kuanua Dictionary.* Rabaul, PNG: Methodist Mission Press.

Zucchi, A. 1993. *The Language of Propositions and Events.* Dordrecht: Kluwer.

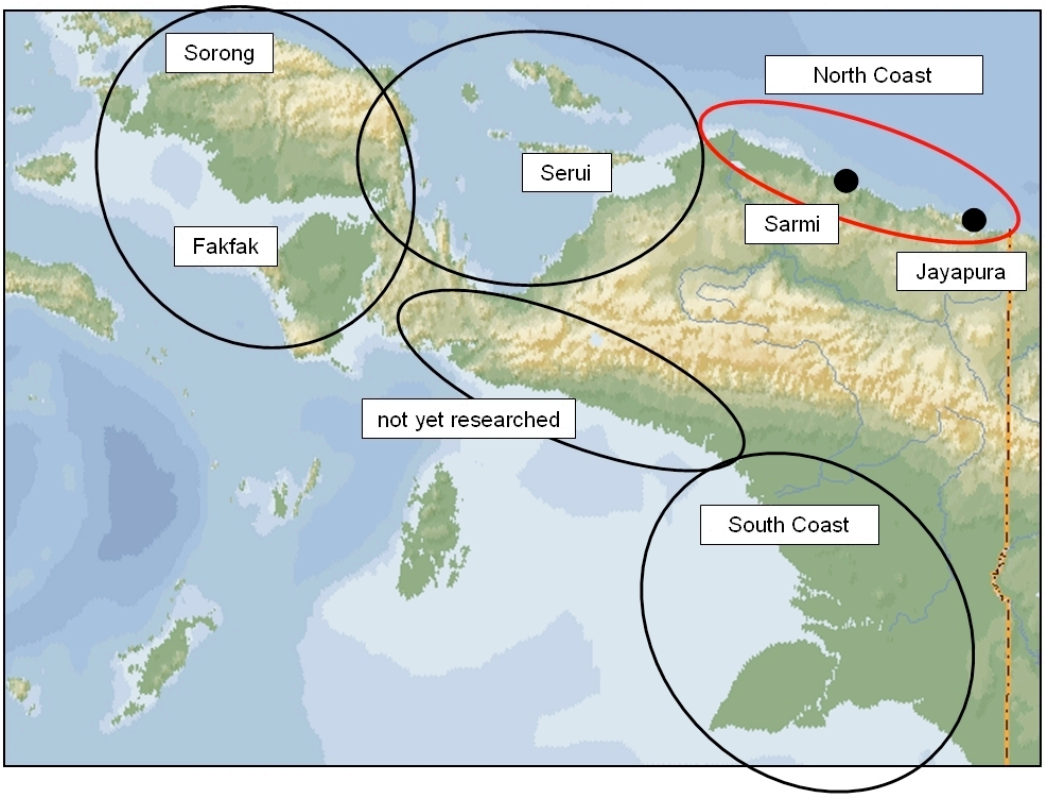
2 *Demonstratives in Papuan Malay*

LENICE HARMS AND ANGELA KLUGE

This paper discusses the demonstratives in Papuan Malay as spoken along the north-east coast of Papua on the island of New Guinea.

"Papuan Malay" refers to the Malay varieties spoken in coastal Papua. So far five varieties of Papuan Malay have been identified. The description of the Papuan Malay demonstratives is based on recordings of narratives and spontaneous conversations between Papuan Malay speakers in the Sarmi and Jayapura areas, both of which are located on Papua's north-east coast.

Map 1: Papuan Malay varieties (based on Donohue in press, and Kim et al. 2007)



Demonstratives are deictic expressions that orient the hearers, and focus their "attention on entities in the situation surrounding the interlocutors" or "in the speech situation" (Diessel 1999:93–94). Papuan Malay has a two-term demonstrative system: proximal ***ini*** 'D.PROX' and distal ***itu*** 'D.DIST'.

Asia-Pacific Linguistics, 2014.

The syntax of the demonstratives is discussed in §1 with the different functions that Papuan Malay demonstratives have in §2. The main points of this paper are summarized in §3, followed by a list of references.

1 Syntax

The Papuan Malay two-term demonstrative system is distance-oriented (Diessel 2006): proximal ***ini*** 'D.PROX' and distal ***itu*** 'D.DIST', together with their reduced fast-speech forms ***ni*** 'D.PROX' and ***tu*** 'D.DIST'.[1] Proximal ***ini*** 'D.PROX' indicates proximity between the deictic center and a referent, while distal ***itu*** 'D.DIST' signals distance between them – in spatial and in non-spatial terms.

The Papuan Malay demonstratives have the following distributional properties:

1. Co-occurrence with noun phrases (adnominal uses): N/NP DEM
2. Substitution for noun phrases (pronominal uses)
3. Co-occurrence with verbs or adverbs (adverbial uses)
4. Demonstrative stacking: (N) DEM DEM

Adnominally used demonstratives occur in post-head position at the right periphery of the noun phrase. That is, all noun phrase constituents occur to the left of the demonstrative, which has scope over the entire noun phrase as illustrated in (1). Demonstratives can also modify constituents other than nouns, namely pronouns as in (2), interrogatives as in (3), or locatives as in (4).

Adnominal uses

(1) *baru* *ana* *kecil* *satu* *ini* *de* *tra* *gambar*
and.then [[[[child] be.small] one] D.PROX] 3SG NEG draw

ana ***murit*** ***satu*** ***ni*** *de* *tra* *gambar*
[[[[child] pupil] one] D.PROX] 3SG NEG draw
'but then ***this particular small child***, he doesn't draw', ***this particular pupil-child***, he doesn't draw'

(2) ***ko*** ***itu*** *manusia* *yang* *tra* *taw* *bicara* *temang*
2SG D.DIST human.being REL NEG know speak friend
'***you*** (EMPH) are a human being who doesn't know how to talk (badly about) friends'

(3) *ana* *laki-laki* *ini* *de* *mo* *ke* ***mana*** ***ni***
child man D.PROX 3SG want to where D.PROX
'this boy, ***where*** (EMPH) does he want to (go)?'

---

[1] The short demonstrative forms are fast-speech phenomena that for the most part are conditioned by their phonological environment. They fulfill the same syntactic functions as the long demonstrative forms. Moreover, they are also employed in the same domains of use with two exceptions, that is, the identificational uses (see §3.2.4) and the placeholder uses (see §3.2.6).

(4) *di* ***sini*** ***tu*** *ada* *orang* *swanggi* *satu*
at L.PROX D.DIST exist person nocturnal.evil.spirit one
'***here*** (EMPH) is a certain evil sorcerer'

In their pronominal uses, the demonstratives stand in place of noun phrases. They occur in all syntactic positions within the clause. In (5), a demonstrative takes the subject slot, in (6) the direct object slot, and in (7) the indirect object slot.

Pronominal uses in argument position

(5) *yo,* ***itu*** *mo* *putus*
yes D.DIST want break
'yes, ***it*** (the river) is going to get dispersed'

(6) *ko* *suka* *makang* ***ini****?*
2SG like eat D.PROX
'do you like to eat ***these*** (that is, fried bananas)?'

(7) *dong* *percaya* ***sama*** ***itu***
3PL trust to D.DIST
'they believe ***in those*** (that is, in evil spirits)'

In their adverbial uses, the demonstratives co-occur with verbs as in ***percaya tu*** 'really believe' in (8) or with adverbs as in ***skarang ini*** 'right now' in (9).

Adverbial uses

(8) *jadi* *dong* ***percaya*** ***tu*** *sama* *setan*
so 3PL trust D.DIST to evil.spirit
'so they ***really believe*** in evil spirits'

(9) ***skarang*** ***ini*** *kamu* *nakal*
now D.PROX 2PL be.mischievous
'***right now*** you're mischievous'

Papuan Malay also allows the stacking of demonstratives. In (10), for instance, short proximal ***ni*** 'D.PROX' modifies the pronominally used long proximal demonstrative.

Demonstratives stacking

(10) *ada* *segala* *macang* *tulang,* *dia* *buang* ***ini*** ***ni***
exist all variety bone 3SG discard D.PROX D.PROX
'there were all kinds of bones, he threw away ***these very (ones)***'

2 Functions

The demonstratives have a range of different functions and uses. Their spatial uses are discussed in §2.1, temporal uses in §2.2, psychological uses in §2.3, identificational uses in §2.4, textual uses in §2.5, and placeholder uses in §2.6. Unless the context of an utterance

is clear and explicit, however, the specific domain of use of the demonstrative cannot always be established without potential alternative readings.

2.1. Spatial uses

A major domain of use for the Papuan Malay demonstratives is to provide spatial orientation. Proximal ***ini*** 'D.PROX' indicates that the referent is conceived as spatially close to the speaker, whereas distal ***itu*** 'D.DIST' signals spatial distance. This function is attested for the adnominally, pronominally, and adverbially used demonstratives.

In (11), the adnominally used demonstratives modify the common noun ***ruma*** 'house'. This example is part of a conversation that took place at the speaker's house. Employing ***ini*** 'D.PROX', the speaker relates her plans to move from her current house, ***ruma ini*** 'this house', to a different house in a neighboring village. Because the new house is smaller than the older one, the speaker's husband is going to enlarge ***ruma itu*** 'that house', with ***itu*** 'D.DIST' indicating that the new house is located at some distance.

Spatial uses of the adnominally used demonstratives

(11) *ini* *kas* *tinggal,* *ana~ana* *dong* *tinggal,* *tong* *pi* *tinggal* *di*
D.PROX give stay RDP~child 3PL stay 1PL go stay at

Sawar *sana* ... ***ruma*** ***ini*** *tinggal* ... *baru* ***ruma*** ***itu***
Sawar L.DIST house D.PROX stay and.then house D.DIST

biking *besar*
make be.big
'(we'll) leave this (house) behind, the children will stay (here), (and) we'll move to Sawar over there .., (we'll) leave ***this house*** behind … and (we'll) make ***that house*** (in Sawar) bigger'

In (12), the speaker replies to a question. Employing pronominally used distal ***itu*** 'D.DIST' the speaker states that ***de ada tu*** 'she is over there'.

Spatial uses of the pronominally used demonstratives

(12) *de* *ada* ***tu,*** *de* *ada* ***tu***
3SG exist D.DIST 3SG exist D.DIST
[Reply to a question:] 'she's ***over there***, she's ***over there***'

In (13), adverbially used proximal ***ini*** 'D.PROX' signals spatial proximity. Talking about the teenagers living in the house, the speaker uses proximal ***ini*** 'D.PROX' to assert that they ***tinggal ini*** 'live here' in this house.

Spatial uses of the adverbially used demonstratives

(13) *ko* ***tinggal*** ***ini***
2SG stay D.PROX
'you ***live here***'

2.2. Temporal uses

In their temporal uses, the demonstratives signal the temporal setting of the situation or event talked about in terms of some temporal reference point. Proximal ***ini*** 'D.PROX' signals that the event is temporally close to the current speech situation while distal ***itu*** 'D.DIST' indicates temporal distance. This function is attested for the adnominally, pronominally, and adverbially used demonstratives.

In (14), adnominally used distal ***itu*** 'D.DIST' indicates that the temporal reference point is located at some distance, in this case in the future: ***Rabu tu hari Kamis itu*** 'next Wednesday, next Thursday'.

Temporal uses of the adnominally used demonstratives

(14) ***Rabu*** ***tu*** ***hari*** ***Kamis*** ***itu*** *ko* *datang* ...
Wednesday D.DIST day Thursday D.DIST 2SG come
'***next Wednesday, next Thursday***, you'll come …'

In (15) pronominally used proximal ***ini*** 'D.PROX' signals that the event is temporally close to the current speech situation: ***ini*** 'right now'.

Temporal uses of the pronominally used demonstratives

(15) *mandi* *cepat~cepat,* ***ini*** *tong* *mo* *lanjut* *lagi*
bathe RDP~be.fast D.PROX 1PL want continue again
'bathe very quickly, ***right now*** we want to continue further' (Lit. '***this*** (is when)')

In (16), adverbially used distal ***itu*** 'D.DIST' indicates that the event is temporally distant from the event described, namely in the past: ***bangung itu*** 'woke up at that time' in (16).

Temporal uses of the adverbially used demonstratives

(16) *sa* *bawa* *pulang* ... *mace* ***bangung*** ***itu*** *dia* *suda*
1SG bring go.home woman wake.up D.DIST 3SG already

snang *karna* *liat* *ada* *makangang*
feel.happy(about). because see exist food
'I brought home (the game that I had shot) … (when my) wife ***got up at that time***, already she was glad because (she) saw there was food' (Lit. '***that waking up***')

2.3. Psychological uses

In their psychological uses, the demonstratives signal the speakers' psychological involvement with the situation or event talked about (Lakoff 1974:347). Three major domains of psychological uses are attested. Emotional involvement is discussed in §2.3.1, vividness in §2.3.2, and contrast in §2.3.3.

2.3.1. Emotional involvement

Speakers employ the demonstratives to signal that they are emotionally involved with the subject matter or that the subject matter is of special interest to them (Anderson 1985:278). While proximal ***ini*** 'D.PROX' signals emotional proximity or positive attitudes,

distal ***itu*** 'D.DIST' indicates emotional distance or negative attitudes. This function is only attested for the adnominally used demonstratives.

In (17) and (18), the speakers employ the demonstratives to signal their attitudes about the referent's ***swara*** 'voice'. The utterance in (17) is part of a conversation about the young people living in the house, none of whom is present during this conversation. The speaker relates that the teenagers enjoy singing. Using direct speech, the speaker conveys her positive attitudes about the teenagers' singing: they should sing more in public because ***kamu pu swara ini*** 'these voices of yours' are good.

(17)	*maju*	*ke*	*depang*	*dang*	*menyanyi*	...	***kamu***	***pu***	***swara***	***ini***
	advance	to	front	and	sing		2PL	POSS	voice	D.PROX

bagus
be.good
'come in front and sing … ***these voices of yours*** are good'

The utterance in (18) occurred during a nightly conversation outside. When one of the teenagers laughs out loudly, the others reprimand her. Employing distal ***itu*** 'D.DIST' in ***ko pu swara tu*** 'that voice of yours', one speaker conveys her negative attitudes about this behavior.

(18)	***ko***	***pu***	***swara***	***tu***	*bahaya,*	*ko*	*stop*
	2SG	POSS	voice	D.DIST	be.dangerous	2SG	stop

'***that voice of yours*** is dangerous, stop it'

2.3.2. Vividness

The emotional involvement does not need to be as substantial as described in §2.3.1. The demonstratives are also used in more general terms to indicate that the subject matter is vivid "to the mind of the speaker" (Anderson 1985:278). To signal that an event or situation is of special interest to them, the speakers use the demonstratives adnominally or adverbially as described in §2.3.2.1, or they employ demonstrative stacking as illustrated in §2.3.2.1.

2.3.2.1. Signaling vividness through adnominally or adverbially used demonstratives

One major strategy to signal vividness is to employ the demonstratives adnominally or adverbially to modify and thereby intensify nominal constituents as in (19), or verbs as in (20).

The utterance in (19) occurred after the speaker had been provoked verbally by an older relative. In her reaction, the speaker modifies the nominal constituents ***bapa-tua*** 'uncle' and ***emosi*** 'feel.angry(.about)' with ***ini*** 'D.PROX', thereby emphasizing them. In choosing the proximal rather than the distal demonstrative to modify the constituent ***bapa-tua*** 'uncle', the speaker also signals that the referent was still near by when she made this comment.

Adnominal uses to signal vividness

(19) *adu,* ***bapa-tua*** ***ni*** *mancing* ***emosi*** ***ni***
oh.no! uncle D.PROX fish.with.rod feel.angry(.about) D.PROX
[After having been provoked:] 'oh no, ***uncle here*** is provoking ***(our) emotions*** (EMPH)'

In (20), distal ***itu*** 'D.DIST' is used adverbially to modify the verb ***lompat*** 'jump', resulting in the emphatic reading ***lompat itu*** 'really jumped'. This example again illustrates the overlapping functions of the demonstratives. The addition of the demonstrative indicates vividness, but the choice of ***itu*** 'D.DIST' also has a temporal effect. That is, the distal demonstrative signals temporal distance, namely past tense.

Adverbial uses to signal vividness

(20) *sunggu* *sa* ***lompat*** ***itu*** *dengang* *tenaga*
be.true 1SG jump D.DIST with energy
'truly, I ***really jumped*** with energy'

2.3.2.2. Signaling vividness through demonstrative stacking

Another, although less common, strategy to signal vividness is the stacking of demonstratives: the first demonstrative is always a long one, while the second is always the corresponding short one. In these constructions, the first demonstrative may be used adnominally as in (21), or pronominally as in (22). In each case, the result of the stacking is an emphatic reading of the entire noun phrase.

In (21) the second demonstrative modifies a nested noun phrase with an adnominal demonstrative such that '[[N DEM] DEM]'. The result of the stacking is an emphatic reading in the sense of 'that very N': ***ruma itu tu*** 'that very house'. In (22) the second demonstrative modifies a pronominally used first one. The result is an emphatic reading in the sense of 'this very (one)': ***ini ni*** 'these very (ones).

(21) *waktu* *kitorang* *masuk* *di* ***ruma*** ***itu*** ***tu*** ...
when 1PL enter at [[house D.DIST] D.DIST]
'when we moved into ***that very house***, …'

(22) *ada* *segala* *macang* *tulang,* *dia* *buang* ***ini*** ***ni***
exist all variety bone 3SG discard [D.PROX D.PROX]
'there were all kinds of bones, he threw away ***these very (ones)***'

2.3.3. Contrast

In their contrastive uses, the demonstratives signal contrast between a discourse referent and another entity, thereby conveying the speakers' attitudes about the subject matter. This function is attested for the adnominally and adverbially used demonstratives.

In (23), the adnominally used proximal demonstrative modifies the pronoun ***saya*** '1SG', thereby indicating an explicit contrast. The speaker compares the ill-behaved young people living in the house to himself. While they have the privilege of staying with relatives in the regional city to complete their secondary schooling, he had to stay with strangers when he was young. This contrast is indicated with ***ini*** 'D.PROX'.

Adnominal uses to signal contrast

(23) *kamu* *ana~ana* *skarang* *ini* *susa* ... ***saya*** ***ini*** *tinggal*
2PL RDP~child now D.PROX be.difficult 1SG D.PROX stay

dengang *orang*
with person
'you, the young people, nowadays are difficult … ***I, by contrast***, stayed with (other) people' (Lit. '***this I***')

In (24) the adverbially used demonstratives signal temporal contrast. In this exchange, a wife and her husband recount how a young man damaged his leg during a motorbike accident. In (24a) the wife relates that ***skarang*** 'now' the referent walks crookedly. Her husband continues the narrative in (24b) with a contrastive statement in which distal ***itu*** 'D.DIST' modifies the temporal adverb ***dulu*** 'prior', thereby signaling a temporal contrast: ***dulu itu*** 'in the past, however'. Subsequently, the wife further elaborates on the referent's condition. She concludes her account with yet another contrastive statement in (24c) in which proximal ***ini*** 'D.PROX' modifies the temporal adverb ***skarang*** 'now', again signaling a temporal contrast: ***skarang ini*** '(it's) just now'.

Adverbial uses to signal contrast

(24) a. Wife: *skarang* *ada* *jalang* *bengkok* *sedikit*
now exist walk be.crooked few
Wife: 'now he's walking a little crookedly (because of his motorbike accident)'

b. Husband: ***dulu*** ***itu*** *de* *jalang* *lurus*
be.prior D.DIST. 3SG walk be.straight
Husband: '***in the past, however***, he walked straight'

c. Wife: ... *ini* *bengkok* *ini,* *kaki* *ini,*
D.PROX be.crooked D.PROX foot D.PROX

skarang ***ini*** *baru* *ada* *baik~baik*
now D.PROX be.new exist RDP~be.good
Wife: 'this (foot) was crooked here, this foot, (it's) ***just now*** that (it got) well'

2.4. Identificational uses

The demonstratives have identificational uses when they appear in the subject slot of a nominal predicate clause. In this context, the demonstratives aid in the identification of a definite or identifiable referent encoded by the predicate. For instance, ***ini*** 'D.PROX' takes the subject slot in (25) and ***itu*** 'D.DIST' in (26). In this domain of use only the long demonstratives are attested.

(25) ***ini*** *daging* *yang* *saya* *bawa* *antar* *buat* *sodara* *dorang*
D.PROX meat REL 1SG bring deliver for sibling 3PL
'***this*** is the (wild pig) meat that I brought (and) delivered for the relatives'

(26) ***itu*** *kali Biri*
D.DIST river Biri
'***that*** is the Biri river'

2.5. Textual uses

In their textual uses, the demonstratives provide discourse orientation. Diessel (1999) discusses four pragmatic uses (see also Himmelmann 1996). The first is the basic anaphoric use, where a demonstrative (adnominal or pronominal) is co-referential with a preceding unit in the discourse. This use is discussed below in §2.5.1. In addition, there are three discourse operations specific to narratives that employ anaphoric demonstratives: Integration, Contrastive Focus, and Restaging. (Longacre and Hwang 2012). These are discussed in §2.5.2. Thirdly is the discourse deictic use, in which the pronominal demonstrative refers to a preceding utterance, rather than a noun phrase. This use is discussed in §2.5.3. The fourth use discussed by Diessel is the exophoric use, in which the speaker refers to something in the situation surrounding the interlocutors, discussed in §2.5.4. Finally, recognitional use, which activates specific shared knowledge between the speaker and audience, is discussed in §2.5.5.

2.5.1. Basic anaphora

In their basic anaphoric use, the demonstratives "are co-referential with a prior NP" and thereby "keep track of discourse participants" (Diessel 1999:93). This function is attested for the adnominally, pronominally, and adverbially used demonstratives.

In (27) the adnominally used distal demonstrative has anaphoric uses. This example is part of a narrative that describes how the speaker's ancestor first came down to the beach where he finds a ***bua mera*** 'red fruit'. At its next mention, the noun phrase ***bua mera*** 'red fruit' is marked with distal ***itu*** 'D.DIST' to signal co-reference with that particular fruit.

Adnominal anaphoric use

(27) *trus di situ ... ada* ***bua mera*** *... de pegang*
then at L.MED exist fruit be.red 3SG hold

bua mera itu *dang de jalang*
Fruit be.red D.DIST and 3SG walk
'and then there … was ***a red fruit*** … he took ***that red fruit*** and he walked (further)'

The example in (28) illustrates a pronominal anaphoric use. This example is part of a description of sagu production. After having introduced the main tool, ***penokok kayu*** 'a wooden pounder', the speaker replaces it at its next mention with proximal ***ini*** 'D.PROX'.

Pronominal anaphoric use

(28) *ada* ***penokok kayu*** *... smua orang tokok dengang* ***ini***
exist pounder wood all person tap with D.PROX
'there is ***a wooden pounder*** … all people pound (sagu) with ***this***'

The utterance in (29) is part of a narrative about a youth retreat and illustrates the adverbial anaphoric uses of proximal ***ini*** 'D.PROX'. During their journey to a retreat, the teenagers

meet an old woman who gives them advice for the retreat. The woman mentions the verb ***jalang*** 'walk' three times while advising the teenagers where to walk and how to behave. When she mentions ***jalang*** 'walk' again, she marks it with ***ini*** 'D.PROX'.

Adverbial anaphoric use

(29) *kamu* ***jalang,*** ***jalang*** *baik~baik* *saja* *kamu* *tinggal* *kamu* ***jalang***
2PL walk walk RDP~be.good just 2PL stay 2PL walk

tida *bole* *ini* *ini,* *...* *kamu* ***jalang*** ***ini*** *untuk*
NEG be.permitted D.PROX D.PROX 2PL walk D.PROX for

apa *pekerjaang* *Tuhan*
what work God
'you ***travel***, just ***travel*** well, (when) you stay (at Takar and when you) ***walk around*** (in Takar), (you) shouldn't (do) this (and) this, … you (do) ***this traveling*** for, what, God's work'

2.5.2. Anaphoric demonstratives in narratives

There are three discourse operations specific to narratives in which Papuan Malay uses demonstratives for participant reference and tracking. As Diessel attests (1999:99), "what all anaphoric demonstratives have in common is that they do not just continue the focus of attention; rather, they indicate that the antecedent is not the referent that the hearer would expect in this context (i.e. the most topical NP)". The Papuan Malay data concur with this statement in that all of the adnominal demonstratives attested in the recorded narratives occur in the context of one of the following discourse operations: integration (§2.5.2.1), contrastive prominence (§2.5.2.2), or restaging (§2.5.2.3).

2.5.2.1. Integration

In narratives, newly introduced participants may be integrated as central in the narrative by using more information than is needed for their identification (Longacre and Hwang 2012). Integration is the second mention of a participant or prop, normally in the clause immediately following its introduction. Diessel (1999:98) states, "[once] a new discourse participant has been established as topic, it is usually tracked by third person pronouns, zero anaphors, definite articles, or pronominal affixes on the verb; but when a referent is mentioned for the second time, demonstratives are often the most common tracking device." In Papuan Malay the demonstrative is frequently used in this "anaphora after first mention" (Lichtenberk 1996:387), especially when globally or locally thematic participants have been introduced.

The example in (30) illustrates a classic integration of a participant. The participants are introduced in the first clause, and in the second, a demonstrative is used to integrate the grandchild as a globally thematic participant: ***cucu ni*** 'this grandchild'. The grandchild is the main character of the story, although the grandfather also continues as a participant throughout. More research is needed to discern the factors involved in choosing this strategy over other available integration strategies of thematically prominent participants.

(30) *tete* *deng* *cucu* *dong* *dua* *pergi* *mancing* *di* *laut,*
grandfather with grandchild 3PL two go fish.with.rod at sea

cucu ***ni*** *de* *pegang* *pisang* *goreng* *dalam* *sak*
Grandchild D.PROX 3SG hold banana fry inside bag
'them two grandfather and grandchild go fishing on the sea, ***this grandchild***, he holds a fried banana in (his) bag'

2.5.2.2. Contrastive prominence

The second discourse operation which employs demonstratives is contrastive prominence, in which the demonstratives are used adnominally as a narrative technique to draw attention to a contrast between two participants. This is similar to the type of contrast discussed above in §2.3.3 but specific to the narrative context; it draws attention, or prominence, to the character in focus.

In (31), the main character's friends have already arrived home and have cooked and eaten their tree kangaroos, while he, Geri, was still attempting to carry his home. This contrast is emphasized by the use of the demonstrative two times in this example: ***dia ini*** 'this one' and ***yang tua ini*** 'this old one'.

(31) *dong* *baker* *dong* *makang, a* ***dia*** ***ini*** *blum,* ***yang*** ***tua*** ***ini***
3PL cook 3PL eat ah 3SG D.PROX not.yet REL be.old D.PROX

de *bawa~bawa* *sampe* *di* *...*
3SG RDP~carry arrive at …
'they cooked, they ate, ah, ***this one*** (has) not yet (arrived), ***this old one*** he kept on carrying until (he) arrived at …'

2.5.2.3. Restaging

Restaging, or reactivating, of participants in a narrative uses a variety of forms, all involving an increase in saliency, or coding material.[2] A participant just mentioned in previous clauses is normally tracked with minimal encoding. However, when the referent has not been mentioned for a number of clauses, further encoding is needed to aid participant tracking.

One of the strategies employed to restage a participant in Papuan Malay is a noun phrase plus demonstrative construction. The other common constructions used for restaging in Papuan Malay are a proper noun or a full noun phrase (without a demonstrative). More research is needed to determine if there are rules governing when each form is used.

In (32) the woman, although she is the main character of the story, has not been part of the action for a few clauses. She is brought back onto stage with the use of a noun phrase plus ***ini*** 'D.PROX': ***prempuang ini*** 'this woman'.

[2] See Givón (1983:17-18) for a description of topic accessibility as it relates to the quantity of coding material used to identify the topic.

(32) *trus* *temang~temang* *turung* *ke* *sana* *panggil~panggil,* *a* *mari*
then RDP~friend descend to L.DIST RDP~call ah! D.PROX

kitong *turung* *olaraga,* *akh,* *sa* *masi* *mo* *tidor,* *trus*
1PL descend do.sports ugh! 1SG still want sleep then

prempuang ***ini*** *dia* *bertriak,* ***ade*** ***prempuang*** ***ini***
woman D.PROX 3SG scream ySb woman D.PROX
'(my) friends came down there calling (me), 'ah, let us go down (to the beach to) do sports', 'ugh, I still want to sleep', then ***this woman***, she screams, ***this younger sister***'

2.5.3. Discourse deictic

The demonstratives also have discourse deictic uses. In this function, the demonstratives do not corefer with a previously mentioned noun phrase. Instead, they are co-referential with a preceding or following proposition (Diessel 1999:101). Only the pronominally used demonstratives have discourse deictic uses.

Proximal ***ini*** 'D.PROX' may refer to a preceding statement as in (33) or to a following statement as in (34). The example in (33) is part of a conversation about difficult children. Maintaining that children should be disciplined, the speaker makes a number of suggestions how to do so. Employing the short demonstrative form ***ni*** 'D.PROX', the speaker summarizes her previous statements. Thereby she creates a link to her closing statement that her interlocutor should decide for herself what to make of these suggestions. In (34), ***ini*** 'D.PROX' creates a link to the following direct quote.

Discourse deictic uses of ***ini*** 'D.PROX'

(33) ***ni*** *usul* *saja* *jadi* *kaka* *sendiri* *...*
D.PROX suggestion just so oSb be.alone
'***this*** is just a suggestion, so you ('older sibling') (have to decide for) yourself …'

(34) *pace* *de* *bilang* ***ini,*** *mace* *ko* *sendiri* *yang* *ikut* *...*
man 3SG say D.PROX wife 2SG be.alone REL follow
'(my) husband said ***this***, 'you wife yourself (should) go (with them) …''

Distal ***itu*** 'D.DIST' is only used to create a link to a preceding statement as in (35). This example is part of joke about an uneducated person who notes that ***di kalender dua blas*** 'in the calendar are twelve (moons)' while ***di langit ini cuma satu*** 'in the sky here is only one'. Distal ***itu*** 'D.DIST' summarizes this proposition about the moon and creates an overt link to the speaker's conclusion that this state of affairs is ***tipu skali*** 'very deceptive'.

Discourse deictic uses of ***itu*** 'D.DIST'

(35) *masa* *di* *kalender* *dua* *blas* *baru* *di* *langit* *ini* *cuma*
be.impossible at calendar two teens and.then at sky D.PROX just

satu *...* ***itu*** *tipu* *skali*
one D.DIST cheat very
[Joke:] '(it's) impossible, in a calendar are twelve (moons), but in the sky here is only one (moon) … ***that's*** very deceptive'

The discourse deictic uses of distal ***itu*** 'D.DIST' are very commonly extended to that of a "sentence connective" that signals "a causal link between two propositions" (Diessel 1999:125). Distal ***itu*** 'D.PROX' alone introduces a reason relation as in (36). When co-occurring with the relativizer ***yang*** 'REL' as in (37), ***itu*** 'D.DIST' introduces a result relation.

In (36), the speaker recounts a conversation with a local doctor after a motorbike accident. In employing distal ***itu*** 'D.DIST' the doctor summarizes the speaker's health concerns and creates an overt link to his explanation as to why she is in pain: ***itu hanya ko jatu kaget*** 'that's just because you're in shock'. In this context ***itu*** 'D.DIST' functions as a causal link that marks a reason relation.

Discourse deictic uses of ***itu*** 'D.DIST': Marker of a reason relation

(36) *sa bilang, tulang bahu yang pata, tulang rusuk, ooo, a*
1SG Say bone shoulder REL be.broken bone rib oh! ah!

mama ***itu*** *hanya ko jatu kaget*
mother D.DIST only 2SG fall feel.startled(.by)
'I said, '(it's my) shoulder bone that is broken, (my) ribs', (the doctor said,) 'oh! ah, Mother ***that is*** just ***because*** you're in shock''

The utterance in (37) is part of a conversation about the speaker's husband who had fallen sick after a straining journey. The speaker relates that her husband had returned home hungry. At the beginning of the next clause ***itu*** 'D.DIST' summarizes the speaker's account and in combination with the relativizer ***yang*** introduces a result relation: ***itu yang de sakit*** 'that's why he's sick'.

Discourse deictic uses of ***itu*** 'D.DIST': Marker of a result relation

(37) *pace de tida makang ... lapar,* ***itu*** ***yang*** *de sakit*
man 3SG NEG eat be.hungry D.DIST REL 3SG be.sick
'he (my) husband hadn't eaten … (he was) hungry, ***that's why*** he's sick'

2.5.4. Exophoric

The exophoric use refers to an entity in the speech situation surrounding the interlocutors, rather than in the discourse itself. Diessel (1999:94) distinguishes between two types of exophoric use: 'gestural', which refers to a concrete entity in the speech situation and is often accompanied by a pointing gesture, and 'symbolic', which may be more abstract or refer to an entity in the wider surrounding situation.

Examples of both types are given below. In (38) Speaker 1 shows his spear gun, a gestural exophoric use; while in (39), an exophoric demonstrative is used symbolically, referring to the town where the discourse was taking place.

(38) a. Speaker-1: *tangkap ka ato de mancing?*
Catch maybe or 3SG fish.with.rod
Speaker-1: '(does he) trap them, or fish with a rod?'

b. Speaker-2: *a pake* ***ini***
ah use D.PROX
Speaker-2: 'ah, use ***this***'

c. Speaker-1: *sumpit~sumpit*
RDP-spear
Speaker-1: ‘(he) spears (them)’

d. Speaker-2: *yo*
yes
Speaker-2: ‘yes’

(39) *baru* *de* *bilang,* *ini* *satu* *misionaris* *perna* *turung*
and.then 3SG say D.PROX one missionary once descend

di ***Sarmi*** ***ini***
at Sarmi D.PROX

‘then he said, ‘there was a missionary (who) once came to ***this Sarmi***’’

2.5.5. Recognitional

The recognitional use is unique in that the demonstratives do not have a referent in the preceding discourse or surrounding speech situation, but rather the speaker is referring to something or someone that is also familiar to the hearer, some specific shared knowledge (Diessel 1999 and Himmelmann 1996). This function is only attested for the adnominal demonstratives.

The example in (40) is taken from the beginning of a story about a younger relative’s death. The recognitional demonstrative is used because her identity is common knowledge between the speaker and hearers of the story. It is referring to shared knowledge rather than to a referent in any preceding discourse: ***ade yang … itu*** “that younger relative …”.

(40) *tapi* *kasiang* ***ade*** ***yang*** ***kemaring*** ***meninggal*** ***dengang*** ***baik*** ***itu***
but pity ySb REL yesterday die with be.good D.DIST
‘but pity on ***that younger relative who died so well recently***’

2.6. Placeholder uses

The demonstratives are also employed pronominally as “placeholders” in the context of “word-formulation trouble”: they function “as temporary substitutes for specific lexical items that have eluded the speaker” (Hayashi and Yoon 2006:499). Attested as placeholders, however, are only the long demonstrative forms.

As placeholders the demonstratives can substitute for any lexical item such as nouns as in (41), pronouns as in (42), or verbs as in (43). Further investigation is needed, however, to account for the alternation of ***ini*** ‘D.PROX’ and ***itu*** ‘D.DIST’ in this context. In most cases, as in (42) and (43), the demonstrative is set off from the subsequently produced target word by a comma intonation. Frequently, however, there is no audible pause between the placeholder and the target word as in (41).

Placeholder for a proper noun

(41) ... *saya* *ingat* ***ini*** ***Meri***
1SG Remember D.PROX Meri
‘(at that particular time) I remembered, ***what’s-her-name, Meri***’

Placeholder for a pronoun

(42) *wa* ***ini***, ***kitong*** *lari~lari* *kemaring* *sampe* ...
wow D.PROX 1PL RDP~run yesterday arrive
'wow, ***what's-their-name***, ***we*** drove yesterday all the way to …'

Placeholder for a verb

(43) *skarang* *sa* ***itu,*** ***simpang*** *sratus* *ribu*
now 1SG D.DIST store one.hundred thousand
'now I (already) ***what's-its-name***, ***set aside*** one hundred thousand (rupiah)'

While in (41) to (43), the demonstratives are used referentially to substitute for a lexical item, this does not seem to be the case in (44). In this example, ***itu*** 'D.DIST' appears to be used as a non-referential "interjective hesitator" that signals "the speaker's hesitation in utterance production" (Hayashi and Yoon 2006:512–513). This is evidenced by the fact that ***itu*** 'D.DIST' does not agree with adnominally used ***ini*** 'D.PROX' which modifies the head nominal ***pace*** 'man'. Further investigation is required, however, to explore whether and in which ways Papuan Malay makes a distinction between the placeholder and non-referential hesitator uses of its demonstratives and whether it may in fact be using this right-displacement as a deliberate construction for emphasis in some contexts (Dooley 2000:36).

Interjective hesitator

(44) *yo,* ***itu***, ***itu***, ***pace*** ***ini*** *de* *baru* *ambil* ...
oh! D.DIST D.DIST man D.PROX 3SG be.new get
'oh, ***umh***, ***umh***, ***this man***, he just took …'

3 Summary

Papuan Malay has a distance-oriented two-term demonstrative system. Syntactically the demonstratives have adnominal, pronominal, or adverbial uses; they can also be stacked. Their functions may be spatial, temporal, psychological, identificational, textual, or as a placeholder. In their psychological uses, the demonstratives indicate emotional involvement, vividness, or contrast. In their textual uses, the demonstratives are used anaphorically, as discourse deictics, and have exophoric and recognitional uses. There are three types of anaphora used specifically in the context of narratives: integration, contrastive prominence, and restaging.

References

Anderson, Stephen R. and Edward L. Keenan. 1985. Deixis. In Timothy Shopen (ed.) *Language typology and syntactic description. Volume 3*: *Grammatical categories and the lexicon*. Cambridge, New York: Cambridge University Press, 259–308.

Diessel, Holger. 1999. *Demonstratives: Form, function, and grammaticalization* (Typological Studies in Language 42). Amsterdam, Philadelphia: John Benjamins Publishing Company.

Diessel, Holger. 2006. Demonstratives. In Keith Brown (ed.) *Encyclopedia of language and linguistics*, 2nd edn.. Amsterdam: Elsevier, 430–435.

Donohue, Mark. to appear. Papuan Malay. In David Gil (ed.) *Studies in Indonesian and miscellaneous languages of Indonesia* (NUSA – Linguistic Studies of Indonesian and other Languages in Indonesia). Jakarta: Universitas Atma Jaya.

Dooley, Robert A. and Stephen H. Levinsohn. 2001. *Analyzing discourse. A manual of basic concepts*. Dallas: SIL International.

Givón, T. 1983. *Topic continuity in discourse: A quantitative cross-language study*. Amsterdam, Philadelphia: John Benjamins Publishing Company.

Hayashi, Makoto and Kyung-Eun Yoon. 2006. A cross-linguistic exploration of demonstratives in interaction: With particular reference to the context of word-formulation trouble. *Studies in Language* 30(3): 485–540.

Himmelmann, Nikolaus P. 1996. Demonstratives in narrative discourse: A taxonomy of universal uses. In Barbara A. Fox (ed.) *Studies in anaphora* (Typological Studies in Language 33). Amsterdam, Philadelphia: John Benjamins Publishing Company, 205–255.

Kim, Hyun, Christian G. Nussy, Benny E. W. Rumaropen, Eleonora L. Scott and Graham R. Scott. 2007. A survey of Papuan Malay: An interim report. Paper presented at the Eleventh International Symposium on Malay/Indonesian Linguistics – ISMIL 11, 6-8 August 2007 in Manokwari. Online URL: http://email.eva.mpg.de/~gil/ismil/11/abstracts/KimShonRumaropen.pdf.

Lakoff, Robin. 1974. Remarks on 'this' and 'that'. *Papers from the Regional Meetings, Chicago Linguistic Society* 10: 345–356.

Lichtenberk, Frantisek. 1996. Patterns of Anaphora in To'aba'ita Narrative Discourse. In In Barbara A. Fox (ed.) *Studies in anaphora* (Typological Studies in Language 33). Amsterdam, Philadelphia: John Benjamins Publishing Company, 379-411.

Longacre, Robert E. and Shin Ja J. Hwang. 2012. *Holistic discourse analysis*. Dallas: SIL International.

3 *Correlative clauses in Seediq*

NAOMI TSUKIDA

1 Introduction

A correlative clause is a kind of relative clause, which appears not adjacent to the head noun, but outside the matrix clause. It is also called an adjoined relative clause. Correlative clauses are reported in languages such as Hindi. An example of the Hindi correlative is given below.

(1) Hindi Correlative (Keenan 1985: 164)

Jis	*a:dmi*	*ka*	*kutta*	*bema:r*	*hai,*
COREL	man	GEN	dog	sick	is
us	*a:dmi*	*ko*	*mai*	*ne*	*dekha.*
that	man	DO	I	ERG	saw

'I saw the man whose dog is sick.'
(lit. 'Which man's dog is sick, that man I saw.')

Seediq has externally-headed relative clauses and internally-headed relative clauses, and in addition, has the third type of construction which can function as a relative clause. In this third type of construction, the clause which functions as a relative clause appears outside the matrix clause, and hence looks like a correlative. In this paper, I will show what this third type of construction is like, and try to characterize it.

2 Brief sketch of the Seediq language

First let me give an outline of the Seediq language.

Seediq is spoken by the Seediq tribe, which is one of Formosan indigenous peoples. They live in the northeastern part of Taiwan. The tribe has more than thirty thousand members, but not all of them can speak the language; young people and children, especially, cannot speak it. There are three dialects: *Teruku*, *Tegedaya* and *Te'uda*. This study is based on the *Teruku* dialect.

Let us look at the basic word order of Seediq. It is strictly predicate-initial, and subject-final.

(2)

Malu	*ka*	*hiyi=su?*
AV.good	NOM	body =2SG.GEN

'Is your body well? How are you?'

Asia-Pacific Linguistics, 2014.

Malu is the predicate. *Ka hiyi=su* is the subject. *Ka* marks an independent nominative, regardless of whether it is pronominal, a common noun, or a proper name.=*su* is a genitive enclitic, which is used when the possessor is expressed by a pronoun.

Modifiers follow the head noun, except for quantity expressions.

(3) *deha huliŋ kumu gaga*
two dog Kumu DIST
'Those two dogs of Kumu's'

The head noun is *huliN* 'dog', *Kumu* is a possessor, *deha* denotes a quantity, and *gaga* is a demonstrative ('that'). Demonstratives occupy the last position in an NP: they are simply juxtaposed, with no linker introducing them. 'Adjectives' are also juxtaposed, usually after the modified noun, again without any linkers. Seediq has prepositions, but no postpositions.

An NP may appear before the clause. In that case, one of the conjunctions *'u/ga/de'u/dega* appears after the NP, and a pause follows it.

NP *'u/ga/de'u/dega*, main clause.

The interpretation differs slightly according to which of *'u/ga/de'u/dega* appears. I limit the examples below to those with *'u* in between.

Such an NP can be regarded as a topic; the construction may be called a topic construction. The function of the topic construction is as follows:

A. Temporal or spatial framework (examples (4), (5))
B. Left-dislocated subject (example (6))
C. Left-dislocated non-subject Actor (example (7))
D. Framework for answers of alternative questions (example (8))

Examples are below.

(4) *Saman 'u, mawsa=nami Tehaypaq da.*
tomorrow PTC AV.FUT.go=1PLEX.NOM Taipei NS
'As for tomorrow, we will go to Taipei.'

(5) *Sapah=na 'u, hebaraw ka patas=na.*
house=3SG.GEN PRT AV.many NOM book=3SG.GEN
'As for his house, there are many books.'

(6) *Laqi gaga 'u, malu (ka hiya).*
child DIST PTC AV.good NOM 3SG
'As for that child, s/he is good.'

(7) *Kumu 'u, b-en-arig=na laqi=na ka patas niyi.*
Kumu PRT CV.PRF-buy=3SG.GEN child=3SG.GEN NOM book PRX
'As for Kumu, she bought this book for her child.'

(8) *Deha laqi niyi 'u, tege'ima ka sewayi=su?*
two child PRX PRT which NOM younger.sibling=2SG.GEN
'As for these two children, which is your younger sibling?'

In the case of left-dislocation of non-subject Actor, there appears an overt resumptive pronoun, =*na*, which is genitive, after the predicate of the main clause (example (7)).

3 Clause linking by *'u* and 'correlative'

In Seediq, there are several ways to connect clauses. One is to put one of the particles *'u*/*ga/de'u/dega* between the clauses, with an intonational break following the particle.

Clause-A *'u*/*ga/de'u/dega*, Clause-B.

The interpretation or function of the construction varies slightly according to which of *'u*/*ga/de'u/dega* appear. In this paper I limit the discussion to those cases involving *'u.*

This looks similar to the topic construction mentioned in the previous section, except that what comes before the particle is a clause. Several interpretations are possible for the topic construction, as we saw previously. Similarly, there are also several possible interpretations concerning the semantic relationship between Clause-A and Clause-B in the construction above. Here are possible semantic interpretations for clause-A and clause-B.

E. Clause-A expresses some kind of condition or framework for clause-B
   (E1) Reason (E2) Temporal
   (E3) Concession (E4) Hypothetical
F. Clause-A and clause-B express contrastive situations.
G. Clause-B is the explanation of the situation depicted by clause-A.
H. Clause-B expresses some kind of cognition and clause-A expresses its content.
I. Clause-A functions as a correlative clause for an NP in clause-B.

Formal and semantic similarities between conditional constructions and topic constructions have already been pointed out cross-linguistically (see Haiman (1978), among others).

A 'correlative' interpretation is one of the various semantic relationships that may hold between the two clauses. *The 'u* construction does not have an exclusively correlative interpretation, but it can be a functional equivalent of the correlative under certain conditions.

Sentence (9) is an example of a Seediq sentence with a 'correlative' interpretation.

(9)

Ga	*reŋag-an*	*yudaw*	*ga*	*ka*	*lupuŋ*	*'u*
DIST.PRG	talk-GV2	Yudaw	DIST	NOM	guest	PTC
pena'ah	*'amirika*	*(ka*	*lupuŋ*	*gaga).*		
from	America	NOM	guest	DIST		

'The guest to whom that Yudaw is talking over there is from America.'

In this example, Clause-A, *ga reŋag-an yudaw ga ka lupuŋ*, functions as correlative clause for the subject NP in clause-B, that is, *ka lupuŋ gaga*.

3.1 Comparison with correlatives of other languages

I will now compare Seediq sentences with correlative interpretations with those in other languages.

Hindi is said to have a correlative construction, as exemplified in (1), repeated as (10) below:

(10) Hindi Correlative (Keenan 1985: 164)

Jis	*a:dmi*	*ka*	*kutta*	*bema:r*	*hai,*
COREL	man	GEN	dog	sick	is

us	*a:dmi*	*ko*	*mai*	*ne*	*dekha.*
that	man	DO	I	RF	saw

'I saw the man whose dog is sick.'
(lit. 'Which man's dog is sick, that man I saw.')

Jis is a correlative marker. It marks the NP_{rel}, one of the coreferent NPs that appears in the relative clause. *Us*, a demonstrative, marks the NP_{mat}, which is the other of the two coreferent NPs appearing in the matrix clause.

Downing (1973) and Keenan (1985) call such a construction 'correlative', but recent studies, such as Andrews (2007), call this type an 'adjoined relative clause'. These terms are used with slightly different nuances.

The term adjoined relative clause was used by Hale (1976) in a description of some Australian languages. In some Australian languages, Diyari (Austin 1981) or Warlpiri, for example, one and the same form is used to express not only a correlative clause, but also a temporal, conditional, etc., clause. So the relevant construction is not exclusively for correlative clauses, but functions as correlative under certain conditions. The Seediq case is thus similar to the Australian cases, as the relevant construction has some usages other than correlative. I continue to use the term correlative to refer to the function or interpretation at hand. Compared to Australian languages, the relevant construction in Seediq has still wider usage.

Below are examples of the adjoined relative clause in Warlpiri.

(11) Warlpiri (Andrews 2007: 216, quoted from Hale 1976)

a.	*Ngatjulu-rlu*	*kapi-rna*	*wawirri*	*purra-mi*
	1-erg	FUT-1sg	kangaroo	cook-nonpast
	kutja-npa	*parntu-rnu*	*nyuntulu-rlu].*	
	rel-2sg	spear-past	you-erg	
b.	*[Nyuntulu-rlu*	*kutja-npa*	*wawirri*	*pantu-rnu]*
	you-erg	rel-2sg	kangaroo	spear-past
	ngatjulu-rlu	*kapi-rna*	*purra-mi.*	
	I-erg	FUT-1sg	cook-nonpast	

'I will cook the kangaroo that you speared.'

In this language, the adjoined relative clause may be on the right side of the main clause, as in sentence (11a), or on the left side of the main clause, as in sentence (11b). The meaning is the same in both cases. In the Warlpiri adjoined relative clause, there is no special marking for NP_{rel}, while *kutja* seems to mark the relative clause.

Below is an example of the adjoined relative in Diyari.

(12) Diyari (Austin 1981: 210)

ŋaṇi	*wiḷa-ni*	*yaṭa-ḷa*	*ŋana-yi/*	*yinda-ṇaṇi.*
1sg	woman-LOC	speak- FUT	AUX-PRES	cry-RELds

'I'll talk to the woman who is crying.'

Adjoined relative clauses in Diyari 'can have a second semantic function, namely, providing information about the temporal or logical conditions holding at the same time as or before the situation expressed in the main clause (Austin 1981; 212)'.

The characteristics of Seediq sentences with correlative interpretations compared to those of other languages are listed as below.

1. There is no marker specific to correlative. The several examples in Keenan (1985) do not have one either; e.g., Wappo, Medieval Russian, and Warlpiri.
2. Conjunctions that appear after the clause-A are limited to *'u*.
3. Correlative clauses appear only at the left side of the main clause (i.e., they are left-adjoined).
4. Both NP_{mat}, one of the coreferent NPs in the matrix clause (i.e., clause-B), and NP_{rel}, the other of the coreferent NPs in the relative clause (i.e., clause-A,) must be the subject in their respective clauses. When the NP_{rel} is not the subject of clause-A, NP_{mat} is not the subject of clause-B, or neither is the subject of their respective clause, a correlative interpretation is totally impossible, and the clause is interpreted as having another meaning, such as conditional or contrastive.
5. NP_{mat} is often realized as a zero anaphor in Seediq, but zero anaphora are observed in examples in Keenan(1985:166) also, so this is not unique to Seediq.
6. Correlatives are said to be limited to 'loose' verb-final languages (Keenan 1985: 164), but word order in Seediq is rather tight.
7. Correlatives are said to be limited to verb-final languages (Downing 1973, Keenan 1985: 164), but Seediq is verb-initial.

   I will explain the 4th point in the next subsection.

3.2 Conditions for 'correlative' interpretation

The 'correlative' interpretation is impossible when NP_{rel}, one of the coreferents in the relative clause, is not the subject of the relative clause (that is, clause-A).The subject of the relative clause (13), *yudaw*, is interpreted as being coreferent with the subject of matrix clause (that is, clause-B). Perhaps this American guest has a Seediq name.

(13)

Ga	*r-em-eŋaw*	*lupuŋ*	*ka*	*yudaw*	*ga*	*'u,*
DIST.PRG	AV-talk	guest	NOM	Yudaw	DIST	PTC
pena'ah	*'Amirika*					
from	America					

'Yudaw who is speaking to a/the guest is from America.'

* 'The guest to whom Yudaw is talking to, he is from America.'

When NP_{mat}, the NP that appears in the matrix clause (that is, clause-B), is not the subject of the matrix clause, the correlative interpretation is impossible. In (14), *patas* 'book' is not the subject in the matrix clause, so the correlative interpretation is not taken.

(14)

a.	**Sepeg-un*	*rubiq*	*ka*	*patas*	*'u,*
	read-GV1	Rubiq	NOM	book	PTC
	m-en-egay	*(patas)*	*ka*	*tama=na.*	
	AV-PRF-give	book	NOM	father=3SG.GEN	
b.	*Sepeg-un*	*rubiq*	*ka*	*patas*	*'u,*
	read-GV1	Rubiq	NOM	book	PTC
	b-en-egay	*tama=na.*			
	CV.PRF-give	father=3SG.GEN			

'Her father gave (her) this book which Rubiq will read.'

Another pair of examples is shown below. Sentence (15a) is not grammatical, because $NP_{mat}$is not the subject, but the non-subject Actor. Sentence (15b) is grammatical, because NP_{mat} is the subject in the matrix clause (though realized as a zero anaphor).

(15) a. **Ga reŋag-an yudaw ga ka lupuŋ 'u,*
DIST.PRG talk-GV2 Yudaw DIST NOM guest PTC
b-en-iq-an=na patas ka rubiq.
PRF-give-GV=3SG.GEN book NOM Rubiq
b. *Ga reŋag-an yudaw ga ka lupuŋ 'u,*
DIST.PRG talk-GV2 Yudaw DIST NOM guest PTC
m-en-egay patas rebiq-an.
AV-PRF-give book Rubiq-OBL
'The guest to whom Yudaw is talking to will give a/the book to Rubiq.'

Another pair of examples is given below. In sentence (16a), NP_{mat} is not the subject of the matrix clause, so a 'correlative' interpretation is impossible. In sentence (16b), on the other hand, NP_{mat} is the subject in the matrix clause, so a correlative interpretation is possible.

(16) a. **Saw 'adi mpeke-geruŋ ka peratu 'u,*
like NEG AV.FUT-break NOM plate PTC
m-barig=ku.
AV.FUT-buy=1SG.NOM
b. *Saw 'adi mpeke-geruŋ ka peratu 'u,*
like NEG AV.FUT-break NOM plate PTC
berig-un=mu.
buy-GV1=1SG.GEN
'I will buy such a plate as will not break.'

In some cases, even when both of the coreferent NPs are the subjects of their respective clauses, not a correlative interpretation, but another interpretation is taken, or another interpretation is taken in addition to a 'correlative' interpretation. For the (17a) sentence, two interpretations are possible: a temporal and a correlative one. For the (17b) sentence, one can only interpret the former clause to express a reason. When a clause beginning with *saw 'adi*is combined to another via *'u*, it may have the same function as the English *lest* clause.

(17) a. *Saw mpe-takur ka laqi 'u,*
like AV.FUT-stumble NOM child PTC
pelekuh-un=na da.
support-GV1=3SG.GEN NS
'When the child is likely to stumble, he supports the child.'
'He supported the child who was likely to stumble.'

b. *Saw 'adi mpeke-geruŋ ka peratu 'u,*
like NEG AV.FUT-break NOM plate PTC
seku-'un=mu kulu.
put-GV1=1SG.GEN box
'I put the plate in a box lest it should break.'

Note that clause-A is identical in (16b) and (17b). There are certain conditions necessary for the form to have a correlative interpretation, but they are not sufficient to produce such an interpretation.

The condition that the relevant NP must be the subject in clause-A is shared with externally-headed RCs and internally-headed RCs also. This will be dealt with in the next section.

3.3 Comparison with other types of RCs

There are other ways for a Seediq clause to modify a noun. Seediq has externally-headed RCs as well as internally-headed RCs. In externally-headed RCs, word order is that of a head noun followed by relative clause. No marker appears between relative clause and head noun. In externally-headed and internally-headed RCs also, the NP_{rel} must be the subject of the relative clause.

Here are examples of Seediq externally-headed and internally-headed RCs. These two sentences have the same meaning.

(18) a. External RC

Wada	*qeduriq*		
is.gone	AC.escape		
ka	*se'diq*	*[p-en-huqil*	*rebiq-an].*
NOM	person	AV.PRF-kill	Rubiq-OBL

b. Internal RC

Wada	*qeduriq*			
is.gone	AV.escape			
ka	*[p-en-huqil*	*rebiq-an*	*ka*	*se'diq]*
NOM	AV.PRF-kill	Rubiq-OBL	NOM	person

'The person who killed Rubiq escaped away.'

In a correlative clause, NP_{mat} also must be the subject of the matrix clause, as we saw in section 3.2, but in externally-headed and internally-headed RCs, it need not. In (19a), 'the man who gave Rubiq a book' is expressed by an externally-headed RC, *se'diq m-en-uway patas rebiq-an*, and it is a non-subject argument in the matrix clause. In (19b), 'the man who gave Rubiq a book' is expressed by an internally-headed RC, *m-en-uway patas rebiq-an ka se'diq*, and it is a non-subject argument in the matrix clause.

(19) a. *S-em-en-teruŋ=ku*
AV-PRF-meet=1SG.NOM

se'diq	*[m-en-uway*	*patas*	*rebiq-an].*
person	AV-PRF-give	book	Rubiq-OBL

'I met [the man who gave Rubiq a book].'(externally-headed RC)

b. *S-em-en-teruŋ=ku*
AV-PRF-meet=1SG.NOM

[m-en-uway	*patas*	*rebiq-an*	*ka*	*se'diq].*
AV-PRF-give	book	Rubiq-OBL	NOM	person

'I met [the man who gave Rubiq a book].'(internally-headed RC)

4 Characterization of Seediq 'correlative'

As an anonymous reviewer writes, translatability is not a good criterion for calling something a 'correlative'. English (20a) can be translated to (20b), which is a case of RC, but we do not regard (20a) as 'correlative'.

(20) a. The child came first in the examination; he is very talented.
b. The child who came first in the examination is very talented.

Are Seediq 'correlative' clauses similar to (20a), in that two clauses are just juxtaposed and interpreted as being related somehow? Is the 'correlative' interpretation semantically related to or derived from other interpretations of '*Clause-A 'u, Clause-B*' sentences? How about their syntax? I will try to characterize Seediq 'correlative' clauses by examining these characteristics in turn below.

4.1 Comparison with other cases of '*Clause-A 'u, Clause-B*' in terms of semantics

Sentences with the form '*Clause-A 'u, Clause-B*' are interpreted in several different ways, as I mentioned at the beginning of section 3. I list the possible semantic interpretations again.

E. Clause-A expresses some kind of condition or framework for clause-B.
(E1) Reason (Example (21)) (E2) Temporal(Example (22))
(E3) Concession(Example (23)) (E4) Hypothetical (Example (24))
F. Clause-A and clause-B express contrastive situations. (Example (25))
G. Clause-B is the explanation of the situation depicted by clause-A.(Example (26))
H. Clause-B expresses some kind of cognition and clause-A expresses its content.(Example (27))
I. Clause-A functions as a correlative clause for an NP in clause-B.(Example (9))

Examples are below.

(21) *Adi* *mpe-sepug* *ka* *laqi* *'u,*
NEG AV.FUT-read NOM child PTC
seku-i *haya.*
put.away-GV.NFIN BEN
'As the child will not read, put it away.' (E1 Reason)

(22) *M-uduh* *siyaŋ* *ka* *sehiga* *'u,*
AV-roast pork NOM yesterday PTC
'eme-'ima *ka* *m-eniq* *hiya?*
PL-who NOM AV-exist there
'When they roasted meat, who were there?'(E2 Temporal)

(23) *M-iyah* *q-em-ita* *tema-'an* *ka* *rubiq* *'u,*
AV-come AV-see father-OBL NOM Rubiq PTC
wada *ŋaŋut* *ka* *tama.*
is.gone outside NOM father
'Though rubiq came to see father, father went outside.'(E3 Concession)

(24) *M-iyah=su* *hini* *'u,* *me-qaras=ku.*
AV-come=2SG.NOM here PTC AV-be:glad=1SG.NOM

'If you come, I would be glad.'(E4 Hypothetical)

(25) *T-em-egesa* *'uyas* *kelemukan* *ka* *tiwaŋ* *'u,*
AV-teach song Taiwanese NOM Ciwang PTC
t-em-egesa *'uyas* *nihuŋ* *ka* *daway* *'uri.*
AV-teach song Japanese NOM Daway also
'Ciwang taught Taiwanese songs, and Daway taught Japanese songs.'(F Contrast)

(26) *Pe-kedus-un=ta=na* *ka* *'ita* *'u,*
CAUS-live-GV1=1PLIN.NOM=3SG.GEN NOM 1PLIN PTC
m-en-da *huway* *yisu* *kirisuto* *tehawlaŋ.*
AV-PRF-pass grace Jisus Christ lord
'That we are made to live by him, it is through the grace of Lord Jesus Christ.' (G Explanation)

(27) *M-usa* *sapah* *rubiq* *ka* *kumu* *'u,*
AV-go house Rubiq NOM Kumu PTC
me-kela *ka* *rubiq.*
AV-know NOM Rubiq
'Rubiq knows that Kumu went to Rubiq's house.' (H Content of cognition)

The semantic connection between the clauses is rather vague generally. E (Condition or Framework) and F (Contrast) may be grouped together, as clause-A describes something which occurs at the same time as or before the clause-B event[1], but it is so vague that one can include G (Explanation), H (Content of cognition), and I (Correlative) in the same group. G (Explanation), H (Content of cognition) and I (Correlative) may be subgrouped together, as clause-A functions as an argument of the predicate verb of clause-B.

It is possible to say that the Seediq 'correlative' interpretation is related to other interpretations, but in a very vague way.

4.2 Comparison with other cases of '*Clause-A 'u, Clause-B*' in terms of syntax

Comparing the '*Clause-A 'u, Clause-B*' sentences in terms of syntax, we can see another possibility.

Let us pay attention to the restriction that the NP_{rel} must be the subject of Clause-A and that NP_{mat} must be the subject of Clause-B. This restriction is not shared with '*Clause-A 'u, Clause-B*' sentences with interpretations other than 'correlative'. Sentences with interpretations other than 'correlative' do not have any restriction on what becomes the subject in Clause-A or Clause-B, as we can see from examples in 4.1. The restriction cannot be related to those '*Clause-A 'u, Clause-B*' sentences with interpretations other than 'correlative'.

The restriction is shared rather by other RCs, as we saw in 3.3.Internally-headed RC, in particular, has the same word order with *Clause-A* of correlative.

The Seediq 'correlative' may be a left-dislocated internally-headed RC. If we left-dislocate the internally-headed RC in (18b), the sentence would be as follows.

---

[1] It is just like Diyari relative clause. See Austin (1981: 209).

(27) *[P-en-huqil rebiq-an ka se'diq] 'u,*
AV.PRF-kill Rubiq-OBL NOM person PTC
wada qeduriq.
is.gone AV.escape
'The person who killed Rubiq escaped.'

One cannot distinguish such a sentence from sentences such as (9) or (14b).

Correlatives and dislocated internal RCs are hard to distinguish. In a discussion on Bambara, Keenan shows that the distinction between correlatives and (dislocated) internal RCs is not always easy to make (Keenan 1985: 165).According to Keenan, one can make a distinction between dislocated internal RCs and correlatives in Wappo because there are two properties that argue that something is a correlative. One of the two properties is that the main verb of S_{rel} is in its main clause form. The other is that S_{rel} is not treated as an NP, which means that it does not take any case marker.

In Seediq it is also hard to distinguish correlatives from dislocated internal RCs. Seediq correlatives have the first property identified by Keenan: the main verb of S_{rel} is in its main clause form. As for the second property, it is hard to judge whether it applies or not: As we saw earlier, either an NP or a clause can appear before *'u*.

There seems to be a difference in distribution, however. Non-subject Actor can be left-dislocated if it is a nominal, as we saw in section 2 (example (28a)=(7)). In the 'correlative', however, NP_{mat} must be the subject, and cannot be non-subject Actor of the matrix clause, as shown in (28b).

(28) a. *Kumu$_i$ 'u, b-en-arig=na$_i$ laqi=na ka patas niyi.*
Kumu PTC CV.PRF-buy=3SG.GEN child=3SG.GEN NOM book PRX
'As for Kumu, she bought this book for her child.' (=(7))

b. **Ga reŋag-an yudaw ga ka lupuŋ 'u,*
DIST.PRG talk-GV2 Yudaw DIST NOM guest PTC
b-en-arig=na laqi=na ka patas niyi.
CV.PRF-buy=3SG.GEN child=3SG.GEN NOM book PRX
I.m. 'As for the guest to whom Yudaw is talking to, s/he bought this book for her/his child.'

In fact, sentences with internally-headed RCs as non-subject Actor are not allowed from the beginning (example (29)).

(29) **B-en-arig [ga reŋag-an yudaw ga ka lupuŋ]*
CV.PRF-buy DIST.PRG talk-GV2 Yudaw DIST NOM guest
laqi=na ka patas niyi.
child=3SG.GEN NOM book PRX
I.m. 'The guest to whom Yudaw is talking tobought this for her/his child.'

Internally-headed RCs can certainly become a non-subject argument, as shown in (19b), but cannot be a non-subject Actor, for some unknown reason; perhaps it is too confusing. So it is natural that left-dislocated version of (29) (example (28b)) is not allowed.

It is possible to say that the Seediq 'correlative' is the result of left-dislocation of internally-headed RC. This happens to take the form '*clause-A 'u, clause-B*', which is used to express other semantic relations as well.

5 Summary

Seediq clause-linking, in the form *clause-A 'u, clause-B,* has several functions, and is not restricted to correlative clauses. It is similar to adjoined relative clauses in Australian languages, but has still wider usage. A necessary condition for the 'correlative' interpretation to hold is that the coreferent NPs must be the subject in both clause-A and clause-B. Even when the condition above is fulfilled, a 'correlative' interpretation may not obtain, and other interpretations may be possible.

Seediq has externally-headed RCs and internally-headed RCs are all subject to the condition that NP_{rel} must be the subject of the RC. Such a condition does not apply to *Clause-A 'u, clause-B* constructions with other interpretations.

I have attempted to characterize the Seediq 'correlative', comparing the 'correlative' with *Clause-A 'u, clause-B* sentences with other interpretations. Semantically, the 'correlative' may be related to other interpretations, but in a very vague way. From the syntactic point of view, it seems possible to regard the Seediq 'correlative' as the result of left-dislocation of internally-headed RCs.

References

Andrews, Avery D. 2007. Relative clauses. In Timothy Shopen (ed.) *Language typology and syntactic description, second edition, Vol. II Complex constructions.* Cambridge/New York: Cambridge University Press. 206-236.

Austin, Peter. 1981. *A grammar of Diyari, South Australia*. Cambridge/New York: Cambridge University Press.

Downing, Bruce. 1973. Correlative clauses in universal grammar. *Minnesota Working Papers in Linguistics and Philosophy of Language* 2. 1-17.

Haiman, John. 1978. Conditionals Are Topics. Language 54(3). 564-589

Hale, Kenneth L. 1976. The adjoined relative clause in Australia. In Robert M.W. Dixon (ed.). *Grammatical Categories in Australian Languages*. Canberra: Australian Institute of Aboriginal Studies, and New Jersey: Humanities Press. 78-105.

Keenan, Edward. 1985. Relative clauses. In Timothy Shopen (ed.) *Language typology and syntactic description, Vol. II Complex constructions*. Cambridge/New York: Cambridge University Press. 141-170.

Tsukida, Naomi. 2009. Sedekkugo (Taiwan) no Bunpou (Grammar of Seediq, Taiwan). Ph.D. dissertation submitted to the University of Tokyo. (In Japanese)

4 *Binding constraints on pronouns and reflexive pronouns in Javanese*

YANA QOMARIANA

1 Introduction

Javanese is an SVO language of the Austronesian family, used mainly on the island of Java in Indonesia. Java Island consists of three provinces: West Java, Central Java, and East Java. Javanese is spoken in Central Java and East Java,while people in West Java speak Sundanese. There are several different dialects of Javanese, for example, Pekalongan, Banyumas, Malang and Banyuwangi. Among these dialects, Central Javanese (especially the variety spoken in Yogyakarta and Surakarta) is considered to be the standard.

This paper is a preliminary study of pronouns and reflexive pronouns in East Javanese, especially in Malang. The data in this paper was taken from speakers of Malang dialect, and only deals with active sentences in *ngoko,* the low level of Javanese.[1]

Across languages, reflexives and pronouns both depend on an antecedent expression to determine their reference; they are therefore both anaphors. Anaphors are subject to restrictions on their possible antecedents: a given antecedent may, or may not, bind its anaphor (i.e., determine its reference) within a given binding domain, such as the clause.

This behavior is captured by binding conditions, such as the requirement, within traditional binding theory, that a reflexive pronoun be bound within its clause. If these binding conditions hold for all anaphors cross-linguistically, then we would expect that the distribution of anaphors is similar across languages. However, there are languages whose anaphors exhibit behavior different from that predicted by traditional binding theory. This suggests the necessity of having more binding constraints, to capture the variation in the distribution of anaphors across languages. Such constraints are accommodated by Lexical Functional Grammar (henceforth LFG). An overview of binding constraints within the LFG framework is presented in the second part of this paper.

Pronouns and their distribution are described in detail in section three. Pronouns in Javanese are inflected for person and number. The categorization of pronouns here is based on Cysouw's (2003) theory. There are three singular pronouns and two plural pronouns. Among all pronouns, it is only the third person singular pronoun *deke* that can takean antecedent.

[1] Javanese has three levels: *ngoko*, the lowest level;*madyo*, the middle level; and *kromo*, the highest level. The choice of level in a conversation depends on the status of interlocutors in a community. For example, the high level is used by a younger person when speaking to an older person: the middle level is used between close friends. This study is only concerned with the lowest level of Javanese spoken in Malang.

Asia-Pacific Linguistics, 2014.

There are six reflexive pronouns in Javanese. Distribution of the reflexives is described in section four. Reflexive pronouns behave similarly to pronouns. Although the issue of logophoric reflexives is mentioned, I will not discuss them specifically. Binding conditions concerning the distribution of pronouns and reflexive pronouns are summarized at the end of sections three and four, respectively.

2 Reflexives and the LFG Framework

While traditional binding theory assumes that anaphoric expressions share the same characteristics across languages worldwide, a more detailed study of anaphors reveals that more parameters are necessary to capture cross-linguistic variety in anaphors' behavior. LFG provides such parameters, allowing for more categories of anaphors to be distinguished.

Following Falk (2001), who works within an LFG framework, I assume that reflexives, reciprocals and pronouns are all anaphors. Further, Kroeger (2004) mentions that anaphors are expressions whose semantic interpretation is based on the interpretation of other expressions; the other expression is called the antecedent. The crucial features for the binding of pronouns and reflexive pronouns are agreement in terms of person, number and gender with their antecedents.

A reflexive expression describes a situation in which 'a participant acts on himself or herself, rather than on any other' (Lichtenberk, 1994). Lichtenberk mentions three strategies to mark reflexivity: (1) Nominal reflexives, where the marker possesses characteristics of nouns or pronouns in a language, such as English reflexives; (2) Verbal reflexives, where the marker is part of the morphology associated with verbs, such as an affix, a clitic or a particle; (3) Possessive reflexives, which use a possessive form, such as possessive adjectives.

Faltz (1985) distinguishes two classes of reflexives: Noun Phrase (NP) reflexives, which are derived from nouns or possessive pronouns, and Verbal reflexives, which are attached to verbs and show coreferentiality with the subject.

Bresnan's (2001) Thematic Hierarchy Condition on Anaphors:
A binds B if A and B are coindexed and A is higher than B on the thematic hierarchy.

The thematic hierarchy reflects the prominence of different semantic roles. The agent is said to be the most prominent role; other roles, such as patient, or locative, are less prominent. The thematic hierarchy is as follows, with the leftmost role being the most prominent, and the rightmost role the least prominent:

Agent > beneficiary > experience/goal > instrument > patient/theme > locative

Another constraint is related to the grammatical function of the anaphors. This constraint is called the Relational Hierarchy, defined below:

A binds B if A and B are coindexed and A is higher than B on the relational hierarchy.

The Relational Hierarchy also follows a left-to-right order of prominence as in following:

SUBJ > OBJ > OBJ2 > OBL…

([Bresnan, 1982] in Arka, 2003)

The binding constraint on anaphors assumes that the antecedents that bind the anaphors are more prominent on the Thematic and Relational Hierarchies.

Manning ([1994, 1996]) in Arka, 2003) suggests that anaphoric binding is also constrained by syntactic prominence with respect to termhood (terms vs non-terms).

Below is the comparison of the syntactic prominences:

Grammatical Function (GF)-prominence : SUBJ > OBJ > OBJ2 > OBL …
Termhood prominence : TERMS NON-TERMS

Therefore the binding constraint for anaphoric expressions can be characterized in three independent parameters below (Falk, 2001):

1. c-structure : (f)-precedence
2. f-structure : relational hierarchy
3. a-structure : thematic hierarchy

Antecedents must outrank their anaphors in terms of c-structural, f-structural and a-structural prominence.

3 Pronouns in Javanese

Pronouns across languages are mostly inflected for person, number and gender (Cysouw, 2003, Kroeger, 2004). A person feature may take three values: speaker or first person, addressee or second person, and others which refer to third person. In Javanese, pronouns and reflexive pronouns are marked for person and number, but not for gender.

Sudaryanto (1991) found five categories of pronouns in Javanese spoken in Semarang[2], as summarized in Table 3.1 below.

Table 3.1: Pronouns in Semarang

Person / Group Category	Pronouns
1	*Aku, awakku, kene*
2	*Kowe, awakmu, kono, sliramu, slirane*
3	*Dewee, dheknene*
1+2	*Kito*
Other plural	Ngoko Sg Pr + *kabeh*, e.g. *aku kabeh*, *kowe kabeh*

However, Javanese spoken in Malang has different pronouns. There are first, second and third person singular pronouns, and also two pronouns that express plurality[3].

[2] Semarang is a city in Central Java.

[3] The term group for plural feature is based on group method of categorization proposed by Cysouw (2003:72). There are seven groups altogether: (1) 1+1, we, mass speaking; (2) 1+2, we, including addressee, excluding other; (3) 1+3, we, including other, excluding addressee; (4) 2+2, you-all, only present audience; (5) 2+3, you-all, addressee(s) and others; (6) 3+3, they; (7) 1+2+3, we, complete. Among seven groups, the first and fourth group are considered as 'feasible categories but they are not linguistically salient'.

Table 3.2: Singular Pronouns in Malang

Person	Singular
1	*Aku*
2	*Koen, awakmu*
3	*Deke*

Tabel 3.3: Plural Pronouns in Malang

Group Category	Plural
2+3	*Koen kabeh*
1+2, 1+3, 1+ 2+3	*Awake dewe*

Javanese does not have any pronouns that express the third person plural, such as English 'they'. The reference to the third person plural is done by repeating the people who are included in the group, as illustrated below:

(1) *Wong-wong* *Amerika* *bangga banget* *Karo* *Patung* *Liberty*
Person-person Amerika proud very of Statue Liberty
"The Americans/they are very proud of the Statue of Liberty".

(2) *Koen kabeh* *bangga banget* *Karo* *Patung* *Liberty*
You all proud very of statue Liberty
"You all are very proud of the Statue of Liberty".

The first sentence above shows the repetition of *wong* 'people' to denote third person plural reference, which in English is represented by 'they'.

Among pronouns in Malang, it is only the third person singular pronoun *deke* that may take an antecedent. As sentences (3-5) make clear, in sentences with transitive verbs*, deke* can occur in subject position, or in object position:

(3) *Budi*$_i$ *ng-gepuk* *deke*$_{j/k}$.
Budi Act-hit 3Sg
"Budi hit him/her".

(4) *Deke*$_i$ *ng-gepuk* *Budi*$_j$
3Sg Act-hit Budi
"He/she hit Budi"

Sentence (3) above shows *deke* in object position; it cannot take the subject *Budi* as its antecedent. Similarly, in sentence (4), *deke* is in subject position, and cannot take *Budi* as its antecedent. In both cases, the pronoun refers to a participant who is not mentioned in the sentence, but who is a discourse referent. Therefore, *deke* is not bound within its domain.

Sentences below illustrate the pronoun *deke* serving as an argument of ditransitive verbs. In sentence (5) *deke* is a complement to a noun, while in sentence (6) it is an oblique.

These examples shows that where *deke* occurs in simple sentences with ditransitive verbs, it can refer to the subject, object, or a discourse referent.

(5) *Budi*$_i$ *ng-(k)anda-ni* *Ali*$_j$ *berita* *soal* *deke*$_{i/j/k}$
Budi Act-told-appl Ali news about 3Sg
"Budi told Ali news about him"

(6) *Budi*$_i$ *ng-(k)anda-ni* *Ali*$_j$ *soal* *deke*$_{i/j/k}$
Budi Act-told-Appl Ali about 3Sg
"Budi told Ali about him

In a situation where both subject and object are possible antecedents, such as above, the subject is the first preference. The discourse referent is in the second, while object is the least likely to be the antecedent.

In sentences (7-9*),deke* is an embedded subject, and cannot take a local antecedent (i.e.,an antecedent within the same clause):

(7) *Budi*$_i$ *ng-omong* *lek* *deke*$_i$ *loro.*
Budi Act-say that 3Sg Sick
"Budi said that he was sick".

(8) *Budi*$_i$ *cerito* *lek* *deke*$_i$ *ketemu* *ambek* *Ali*$_{k.}$
Budi Act-told That 3Sg met With Ali.
"Budi told that he met Ali".

(9) *Budi*$_i$ *ng-andhani* *Siti*$_j$ *lek* *deke*$_{*i/j}$ *kudu* *tuku* *mobil anyar.*
Budi Act-told Siti that she must Buy car new
"Budi told Siti that she must buy a new car"

By now we can see that *deke* can function as either a short-ora long-distance anaphor. In LFG terms, *deke* might be characterised as [+ Nucleus],as it can take a local antecedent, and [- Nucleus], because it may also take a long-distance antecedent. Arka (2003:223) shows that Balinese has a similar type of anaphor, and suggests this type is not subject to a nucleus constraint. Following Arka, I consider this property, the lack of a nucleus constraint, to be the general (syntactic) binding constraint on the pronoun *deke.*

The distribution of the pronoun *deke* is summarized in Table 3.4. *Deke* occurring in the subject position of simple sentences cannot be co-referent with other arguments in the sentence; yet, as an embedded subject, it can corefer to the subject or object of the matrix clause, and is disjoint from other arguments that share the same predicate.

In a complement clause, the antecedent of *deke* may be either the subject or the object. In a situation where both subject and object are available to be antecedent, the subject is preferred over the object. This leads to the conclusion that the pronoun *deke* follows the relational hierarchy. This also coheres with the thematic hierarchy, because *deke* prefers to be bound by the agent, rather than the beneficiary, as illustrated in (9). *Deke* also respects the thematic hierarchy, as it always bound by the antecedent which is semantically most prominent.

Table 3.4

Function	Antecedent	Type of Verb
OBJ	▪ Disjoint to subject ▪ Discourse referent	Simple sentences with primary transitive verbs
OBL	▪ Bound to subject ▪ Bound to object ▪ Discourse referent	Simple sentences with ditransitive verbs
SUBJ	▪ Disjoint to object ▪ Discourse referent	Simple sentences with an verb
Embedded Subject	▪ Bound to subject and object of the matrix sentence ▪ Discourse referent ▪ Disjoint to the GF of the same predicate	Compound sentences with intransitive and transitive matrix verbs

The binding constraint on the pronoun *deke* is summarized below:

Table 3.5

Antecedent	Domain	Relational Hierarchy	Thematic Hierarchy
Gf-command	No nucleus constraint	Yes	Yes

4 Reflexive Pronouns in Malang

Following Sudaryanto (1991), reflexive pronouns in Malang are nominal reflexives. The reflexives derive from the noun *awak* 'body'. Sudaryanto illustrates this in the sentence below:

(10) *Aku* *arep* *ng-asoake* *awak* *sedelo*
I will act-rest body a while
"I will rest my body for a while"

There are several forms in Malang that express reflexivity. The first form consists of the noun *awak* + possessive marker + *dewe*. The possessive marker of the first person singular is *–ku*, second person singular *is –mu*, while third person singular is *–e*. The possessive pronoun cliticizes to the noun *awak* and shows agreement with the person category. The form is used for the first person singular reflexive *awakku dewe*, second person singular *awakmu dewe*, third person singular *awake dewe* and the unified 'we' *awake dewe*. Recall that this form is also used as the third person singular pronoun.

The second person plural reflexive uses the second person singular reflexive form plus *dewe* which yields *awakmu dewe dewe;* meaning literally 'each of yourselves'. The same strategy is used to form the reflexive for the third person plural *awake dewe dewe*.

The reflexive forms are summarized in tables 4.1 and 4.2 below:

Tabel 4.1: Reflexives of Singular Category in Malang

Person Category	Reflexive Pronoun
1	*awakku dewe*
2	*awakmu dewe*
3	*awake dewe*

Tabel 4.2: Reflexives of Plural Category in Malang

Person Category	Reflexive Pronoun
1+2+3	*awake dewe*
2	*awakmu dewe dewe*
3	*awake dewe dewe*

The above forms are reflexives because the only way to create the reflexive meaning is by using those reflexive forms.

(11) *Aku* *n-(t)embak* *awakku dewe.*
I Act-shot 1Sg Reflex
"I shot myself"

(12) **Aku*$_i$ *n-(t)embak* *aku*$_i$
1Sg Act-shot 1Sg
"I shot me"

(13) **Aku*$_i$ *n-(t)embak* *akudewe*
1Sg Act-shot 1Sg alone
"I shot me"

(14) *Koen* *n-delok* *awakmu dewe* *ndek koco*
You Saw 2Sg Reflex on mirror
You saw yourself on the mirror.

(15) *Deke* *n-delok* *awake dewe* *ndek koco*
3Sg Saw 3Sg Reflex on mirror
"She/he saw her/himself on the mirror"

(16) *Koen*$_i$ *n-(t)embak* *koen*$_{*i/j}$
2Sg Act-shot 2SG
"You shot you (other person in the audience)"

(17) *Deke*$_i$ *n-(t)embak* *deke*$_{*i/j}$
3Sg Act-shot 3SG
"He/She shot he/she"

In sentences (11), (14) and (15), the reflexive forms are used in the object position, which produces the reflexive meaning, whereas while the use of pronouns will either yield ungrammatical sentences, as in (12), or sentences without reflexive interpretation, as in

(16) and (17). Sentence (13) shows that the use of pronoun + *dewe* does not give a reflexive interpretation. Plural reflexives show a similar distribution to singular reflexives.

We have seen the occurrence of reflexives as the Object of transitive verbs. Javanese also allows reflexives to occur in the oblique position. The use of pronouns in this similar position does not yield a reflexive interpretation, as illustrated below:

(16) *Aku*$_i$ *ng-gawe* *ngombe* *gawe* *awakku dewe*$_i$
1Sg make a drink For 1Sg Reflex
"I made a drink for myself"

(17) **Aku* *ng-gawe* *ngombe* *gawe* *aku*
I Act-make a drink For Me
"I made a drink for me"

(18) *Koen* *ng-gawe* *ngombe* *gawe* *koen dewe*
2Sg Act-make a drink For yourself
"You made a drink for yourself"

However, the occurrence of reflexives as OBJ2 and OBJ in sentences with ditransitive verbs is not allowed:

(19) **Aku* *ng-gawe* *ngombe* *awakku dewe*
1Sg act-make a drink myself
"I made a drink for myself"

(20) **Aku* *ng-gawe* *awakku dewe* *ngombe*
1Sg act-make a drink Myself
"I made a drink for myself"

The use of reflexives above is similar to that in sentences with applicative process (using suffix *–no*), except the no-OBJ2 condition does not hold. The use of reflexives in OBJ2 position is allowed after the verb has undergone applicativization, as illustrated in the following:

(21) *Aku* *ng-gawek-no* *Ngombe* *gawe* *awakku dewe*
1Sg act-make-Appl a drink for 1Sg Reflex
"I made drink for myself"

(22) **Aku* *ng-gawek-no* *awakku dewe* *ngombe*
1Sg act-make-Appl 1Sg Reflex a drink
"I made drink for myself"

(23) *Aku* *ng-gawek-no* *ngombe* *awakku dewe*
1Sg act-make-Appl a drink 1Sg Reflex
"I made a drink for myself"

The sentences above also illustrate binding in applicative constructions. The reflexive pronouns occur as obliques and they take the agent as their antecedent, which occurs in subject position. The reflexive always takes a more prominent antecedent.

In the causative construction, reflexives also take more prominent antecedents. The following sentences illustrate binding of reflexive pronouns in causative constructions:

(24) *Adik*	*n-(t)ibak-no*	*awake dewe*	*tekok*	*bandulan*
Little brother	act-fall-appl	3Sg Reflex	off	swings

"Little brother caused himself to fall off the swings"

In (24) above, is the subject causer in the causative construction usually has the agent role. The causer in the sentence above is the subject *Adik* and the causee is the reflexive pronoun *awake dewe*, which occurs in an object position. *Awake dewe* takes the agent *Adik* as its antecedent. *Adik* is an agent and is more prominent than *awake dewe*, which is a patient.

The causer in the causative construction usually has the agent role. The causer in the example above is the subject *adik* and the causee is the reflexive pronoun *awake dewe*, which occurs in object position *Awake dewe* takes the agent *adik* as its antecedent. *Adik* is an agent and it is more prominent that the *awake dewe* which is a patient.

So far we have seen reflexives function as short-distance anaphors. In fact, reflexives can functionas long-distance anaphors. In compound sentences, a reflexive can take the subject of the matrix clause as its antecedent.

(25) *Ali*$_i$	*ng-(k)ongkon*	*deke*$_j$	*n-(t)inju*	*awake dewe*$_{i/j}$
Ali	act-told	3Sg	Act-punch	3Sg Reflex

"Ali told him/her to punch himself/herself"

In the above sentence, *awake dewe* can take either *Ali* or *deke* as its antecedent. The fact that *awake dewe* can refer to Ali,the subject of the matrix clause,shows that it has the characteristics of a long distance anaphor.

Awake dewe can also occur in a construction with two embedded subjects.

(26) *Ali*$_i$	*ng(k)ongkon*	*deke*$_j$	*n-(c)yrita-ni*	*Nina*$_k$	*berita*	*Soal*	*awake dewe*$_{i/j/k}$
Ali	Act-told	3Sg	Act-told-Appl	Nina	news	about	3Sg Reflex

"Ali told him/her to tell Nina a story about him/himself"

Sentence (26) shows *awake deke* in the object complement position. It can take *Ali,deke,* and also *Nina* as its antecedents.

The followingsentences illustrate that the embedded subject in a higher clause has a different person feature to the subject in the highest and lowest clause.Observe the following:

(27) *Ali*$_i$	*ng(k)ongkon*	*koen*$_j$	*ng-(k)andani*	*Nina*$_k$	*soal*	*awake dewe*$_{i/j/k}$
Ali	Act-told	3Sg	Act-told-Appl	Nina	about	3Sg Reflex

"Ali told you to tell Nin about him/heself"

Sentence (27) shows that the embedded subject *koen* does not affect the corefence of *awake dewe* to *Ali*, the latter being the subject of the higher clause. In terms of semantic roles, the theme, *awake dewe*, is bound by a patient, *Nina*, or agent *Ali*.

In compound sentences, as below, reflexives takes objects as their antecedents.

(28) *Ali*$_i$ *Kondo Nang deke*$_j$ *lek deke*$_j$ *kudu n-(t)embak awake dewe*$_j$
Ali Told To 3Sg that 3Sg must act-shoot 3Sg reflex
"Ali told him that she/he must shoot her/himself"

Sentence (28) shows that *awake dewe* is bound by the oblique benefactive; while in sentence (29), the reflexive is bound by the object benefactive.

The distribution of reflexive pronouns in Malang is summarised below:

Table 4.3: Distribution of Reflexive Pronouns

Function	Antecedent	Domain
Object	Subject	Simple sentence with transitive verbs
Oblique	Subject	Simple sentences with ditransitive verbs
Object 2	Subject	Simple sentences with ditransitive causative/applicative verbs
Object of a lower clause	▪ Subject of matrix clause ▪ Object of matrix clause ▪ Subject of lower clause	Compound sentences
Complement	▪ Subject of matrix clause ▪ Object of matrix clause ▪ Subject of lower clause	Compound sentences

Based on their distribution, the function of the reflexives as short- and long-distance anaphors is captured by their behavior in compound sentences. Similar to pronoun *deke*, the reflexives do not have a nucleus constraint.

All the reflexive pronouns are allowed to take not only a subject antecedent, but also an object antecedent. However, the reflexive pronouns always obey the GF command condition (Arka, 2003), as the grammatical function of the antecedent is higher than the grammatical function of reflexives according to the relational hierarchy.

In terms of thematic hierarchy, the occurrence of reflexive pronouns follows the general rule that the antecedent outranks the reflexive. All antecedents are more prominent than reflexive pronouns in thematic hierarchy, i.e. agent antecedent > patient reflexive pronoun. In the case of reflexive pronouns, when they occur in the object position of the lower clauses they take matrix object antecedents; the object/patient of the matrix sentence is more prominent than the object/patient in the lower clauses.

Binding constraints on reflexive pronouns in Malang is summarized below:

Table 4.4: Binding Constraint of Reflexive Pronouns

Antecedent	Domain	Relational hierarchy	Thematic Hierarchy
GF-commanded antecedent	No nucleus constraint	Yes	Yes

5 Conclusion

In this paper I have discussed the distribution of the third person singular pronoun *deke* and reflexive pronouns in Javanese, and also their binding constraints, using the LFG approach.

The third person singular pronoun *deke* can take either a local or non-local antecedent. This means that *deke* can function as either a long-or a short-distance anaphor. In LFG terms, *deke* appears to have the f-structural features [+ minimal complete nucleus] and [- minimal complete nucleus], which leads to the conclusion that it is not subject to the nucleus constraint. In terms of the relational and thematic hierarchy, *deke* follows the general assumption that the antecedent is more prominent than the anaphor. The ability to take a more prominent antecedent in terms of the relational hierarchy assigns the feature [+GF-command] to *deke.*

Reflexive pronouns in Javanese can occur in subject, object and complement position. Their appearance in OBJ and OBJ2 positions in ditransitive constructions is not allowed. Reflexive pronouns can only appear in OBJ and OBJ2 positions when the verb has undergone applicativisation. All reflexive forms have similar distribution to pronoun *deke,* which shows that they can be either short- or long-distance anaphors; therefore, they too have no nucleus constraint.

References

Arka, I Wayan. 2003. *Balinese morphosyntax: a lexical-functional approach.* Canberra, Pacific Linguistics.

Bresnan, Joan. 2001. *Lexical Functional Syntax*. Massachusetts, Oxford, Blackwell Publisher.

Cole, Peter; Hermon, Gabriella: Kim, Chongyuck; Sim, Chang-yong; and Tsai, Yaping. 2003. *Anaphoric Expression in Javanese*. In Rakowski, Andrea, and Richard, Norvin (Ed). Proceedings of AFLA VIII: The Eight Meeting of The Austronesian Formal Linguistics Association, MIT 44, Working Papers in Linguistics 44, Spring 2003.

Cysouw, Michael. 2003. *The Paradigmatic Structure of Person Marking.* Oxford, OUP.

Dalrymple, Mary. 1993. *The Syntax of Anaphoric Binding*. Stanford, CSLI

Falk, Yehuda. 2001. *Lexical-Functional Grammar*. Stanford, CSLI

Kroeger, Paul R. 2004. *Analying Syntax – A lexical-functional approach*. Cambridge, Cambridge UP

Lichtenberk. Frantisek. 1994. *Reflexives and Reciprocals.* In R.E. Asher and J.M.Y. Simpson (eds). The Encyclopedia o Language and Linguistics. Vol 7. Oxford, Pergamon Press, 3504-3509

Soedaryanto et al. 1991. *Tata Bahasa Baku Bahasa Jawa*. Yogyakarta, Duta Wacana UP

Suhandono. 1994. *Grammatical Relation in Javanese*: A Short Description. Master Thesis. The Australian National University, Canberra

5 *The distribution of unmarked cases in Samoan*

JAMES N. COLLINS

1 Introduction[1]

Legate (2008) proposes a class of ergative case marking systems, in which sole arguments of intransitives receive nominative case, transitive patients receive accusative case, and transitive agents receive ergative case. The system does not exhibit absolutive as a distinct case. The appearance of a single morphological case on both intransitive sole arguments and transitive patients is due to a morphological syncretism between the nominative and accusative morphological cases.

I argue that the Polynesian language Samoan falls into this category of ergative languages, arguing that both nominative and accusative case in Samoan are realised by the absence of any phonologically overt case marking on full DPs, but that nominative and accusative pronouns are morphologically distinguished. The evidence for this comes from nominalised clauses, where the distribution of the morphologically null case (ordinarily taken to be absolutive case) follows a nominative-accusative like distribution: P may retain the morphologically null case under nominalisation, but S and A may not. I argue this pattern falls out from positing two distinct morphologically null abstract Cases, nominative and accusative, which are licensed by distinct mechanisms; only accusative case is licensed under nominalisation. I also demonstrate previous analyses of Polynesian ergativity within the generative tradition (e.g., Bittner and Hale 1996a; Massam 2001) predict the wrong results for these data.[1]

2 An introduction to Samoan morphosyntax

Samoan is spoken in Samoa and American Samoa and by significant immigrant populations in New Zealand, Australia, the USA and elsewhere. It is an Austronesian

[1] With sincerest thanks to Vince Schwenke Enoka, Reuben Mauga, Rev. Tauoa Mauga Head, Siatua Vailele, Iakopo Leleimalefaga and LJ Siu for their time and generosity as consultants. Thanks also to Vera Gribanova, Beth Levin, Paul Kiparsky, Maria Polinsky and Ivan Sag, the audiences at Austronesian Formal Linguistics Association 19, the 12th International Conference on Austronesian Linguistics, the Stanford University Syntax-Morphology Circle and the UC Berkeley Syntax-Semantics Circle for helpful and insightful comments.

Asia-Pacific Linguistics, 2014..

language of the Polynesian sub-branch. It is head-initial, with predicate-initial word ordering.

The basic paradigm for case marking finite Samoan clauses follows in (1). The preposition *e* only occurs on the more agentive argument of a transitive predicate. Throughout I use the abbreviation A for the more agentive argument of a transitive predicate, P for the less agentive argument and S for the sole argument of an intransitive predicate, a convention borrowed from typological literature on ergativity (Comrie 1978). In (1a), the preposition *e* occurs on the more agentive argument of a transitive. Both the patient of the transitive verb in (1a), and the sole arguments of the intransitive verbs in (1b) and (1c) are not marked with any preposition. This pattern of behaviour has led the descriptive literature on Samoan (including Mosel and Hovdhaugen 1992, Milner 1976) to classify the absence of case marking on a DP as 'absolutive' case. Where the sole argument of an intransitive predicate is agentive, it does not take ergative marking, (1c).[2]

(1) a. *'Olo'o* *fafao* *e* *le* *tama* *le* *pusafa'i*
PROG pack ERG the boy the banana-case
'The boy is packing the banana-case.'(Milner 1976; 59)

b. *'Ua* *to'a* *le* *vai*
PERF settle the water
'The water settled down.'(Milner 1976; 269)

c. *'Ua* *'ata* *le* *tama*
PERF laugh the boy
'The boy laughed.'(Milner 1976; 26)

Tense, mood and viewpoint aspect in Samoan are marked on an auxiliary-like element which precedes the predicate. The predicate may fall into a large range of semantic categories. For example, it could be an event-denoting predicate, as in (2a), a stative predicate (2b), or a numeral (2c).

(2) a. *'Ua* *mele* *le* *vai* *e* *A'opo*
PERF throw.away the water ERG A'opo
'A'opo threw away the water.' (Milner 1976: 144)

b. *Sā* *fiafia* *tele* *le* *ali'i*
PAST happy very the chief
'The chief was very happy.' (Mosel and Hovdhaugen 1992: 421)

c. *E* *lua* *ta'avale* *a* *Feleti*
PRES two car.PL GEN Feleti
'Feleti has two cars.' (Mosel and Hovdhaugen 1992: 422)

---

2 Abbreviations: ACC accusative, ALIEN alienable, ANTICAUS anticausative, CAUS causative, DAT dative, DIR directional particle, ERG ergative, GEN genitive, HON honorific, INAL inalienable, LOC locative, NOM nominative, NMZR nominaliser, PERF perfective, PL plural, PRES present, PROG progressive, REDUP reduplication, SG singular, SUBJNCT subjunctive, TOP topic marker

As is clear from the above examples, the basic clause structure in Samoan is predicate-initial. Roughly speaking, the ordering of constituents within a clause headed by a transitive verb is VAP, though the relative ordering of the post-verbal arguments is largely variable, with the rightmost element receiving a focused interpretation. In (3), variances between VAP and VPA ordering have an effect of focussing A and P respectively.

(3) a. *Sā* *'ai* *e* *le* *tamaloa* *le* *i'a*
PAST eat ERG the man the fish
'It was the man that was eating the fish.'

b. *Sā* *'ai* *le* *i'a* *e* *le* *tamaloa*
PAST eat the fish ERG the man
'It was the fish that the man was eating.'

The structure of the nominal complex is similar to the clausal structure exemplified above. I assume the nominal complex is a DP, headed by a determiner located at the left periphery of the constituent. The determiner precedes the nominal head, which is followed by any arguments or adjuncts. As in non-nominal clauses, the ordering of the arguments within a DP is variable.

(4) *le* *ulua'i* *fa'apāina* *e* *tagata* *o* *le* *pulu* *atomika*
the first CAUS.explode ERG person GEN the bomb atomic
'The first exploding of the atomic bomb by people.' (Mosel and Hovdhaugen 1992: 545)

Arguments in Samoan are often dropped if their referent may be inferred from context. This is most common in conjoined clauses where arguments overt in one clause may be dropped from any following clauses. In (5), there is a sequence of conjoined clauses. After the first occurrence of the nominals which refer to Sina and the pigeon, any further references to these individuals are implicit.

(5) *Tū* *atu* *loa* *lea* *'o* *Sina,* *tag* *'i* *le* *lupe,*
o
stand up then that TOP Sina hol DAT the pigeon
d
titina, *togi* *'i* *fafo*
strangle Throw DAT outside
'Then Sina stood up, [she] took hold of the pigeon, [she] strangled [it], [she] threw [it] outside.' (Mosel and Hovdhaugen 1992: 706)

Turning now to clauses with nominal predicates, a DP predicate precedes its subject. The predicational DP is prefixed with the particle *'o*. In the following clause, the predicational DP *se fale Sāmoa*, 'a Samoan house', is marked with *'o*, while the referential DP *lo'u aiga fou*, 'my new house', is left unmarked.

(6) *'O* *se* *fale* *Sāmoa* *lo'u* *aiga* *fou*
TOP a house Samoa my house new
'My new house is a Samoan house.' (Mosel and Hovdhaugen 1992: 502)

The same *'o* particle is also seen in the following structure, where a core argument may be fronted to a position to the left of the tense marking auxiliary. In the example below, the argument *le mea* precedes the tense marker and is marked with *'o*. The post-verbal position where *le mea* would appear in a canonical transitive structure is left empty.

(7)	*'O*	*le*	*mea*	*na*	*fai*	*e*	*le*	*tama*
	TOP	the	thing	PAST	do	ERG	the	boy

'It is the thing that the boy did.' (Cook 1991: 85)

Having introduced a set of data relating to Samoan morphosyntax, I will now lay out the data of central import to this paper, relating to the distribution of morphological case on core arguments.

3 A tripartite case system in Samoan

This section lays out the evidence for a case assignment mechanism which assigns two distinct cases to S and P (nominative and accusative respectively). The two cases are morphologically syncretic. This gives rise to the appearance of a single morphological case on both S and P, often termed absolutive. The evidence for two distinct but homophonous morphological cases comes from two distinct patterns of the distribution of morphological case on S and P. In explaining the data, I will discuss how an apparent ergative-absolutive pattern arises in the morphological case system.

The layout of the argument within this section proceeds as follows — first, I demonstrate that S and P have distinct patterns with respect to their morphological case. The data comes from nominalised clauses and pronominal forms. Based on these data, I argue that a system of morphological case assignment must not treat S and P in Samoan as a unified category, contra any analysis which posits an absolutive abstract Case. I also argue that the data presented is consistent with the system of morphological case assignment proposed by Legate (2008) for a subclass of ergative languages, which demonstrate three distinct abstract Cases according to her analysis: nominative, accusative and ergative. The approach stands in opposition to analyses which analyse ergative systems as assigning two cases to core arguments: the marked ergative and the unmarked absolutive. Under Legate's system, there are two unmarked cases, nominative and accusative, as well as the marked ergative case. The availability of accusative case allows us to account for the complicated range of data presented in this section, which demonstrates that each of S, A and P takes a unique set of morphological realisations.

The primary focus of the entire paper is the class of verbs in Samoan which I label as transitive. It is these verbs, and these verbs only, which can co-occur with ergative case marked arguments. These verbs fall into two broad categories each of which display some degree of semantic homogeneity. The first category includes verbs which exhibit high semantic transitivity in the sense of Hopper and Thompson (1980). They involve two or more participants, denote dynamic (non-stative) events, entail that one participant is the causer or initiator of the event. The other participant is in some way affected by the event. The class includes verbs whose arguments undergo some kind of change of state as a result of the event denoted by the verb. It also includes arguments which denote recipients of some force, including the point of contact within events of surface contact (e.g., *hit, sweep*), and the participant upon which force is exerted in events of force exertion (e.g., *pull, push, yank*). Some examples of verbs in this large category which co-occur with ergative and (what I will argue to be) accusative case are listed below

(8)	*tipi*	'cut'	*muāvae*	'kick'
	poi	'sever'	*moto*	'punch'
	vae	'divide'	*pani*	'strike'
	sala	'slice'	*tu'i*	'strike, blow'
	tutū	'cut down'	*toso*	'pull'
	soni	'cut up'	*fālō*	'pull down'
	pena	'butcher'	*fai*	'build'
	polo	'cut into pieces'	*vali*	'paint'
	tafa	'lance (a boil)'	*lalaga*	'weave'
	fa'i	'break, snap'	*tusi*	'write'
	gau	'break'	*me'i*	'draw'
	motu	'break (s.t. brittle)'	*'ai*	'eat'
	talepe	'break (s.t. long)'	*inu*	'drink'
	sasa	'beat'		

Verbs in the second category are necessarily morphologically complex. They are derived from verbs which in isolation fall under a wide set of categories, including intransitive verbs, dative case assigning verbs and transitive verbs from the category exemplified in (8). With the addition of the suffix *–Cia*[3], the pattern of morphological case on the verb's arguments must be ergative-accusative. The semantics of the suffix -*Cia* is complex, and for simplicity is not discussed to any great extent in this paper. This set of -*Cia* suffixed verbs demonstrate a lower semantic transitivity in the sense of Hopper and Thompson than the class exemplified in (8). These verbs do not always denote dynamic events (e.g, *fa'alagaia* denotes the state of A desiring P), but generally entail the event was caused or initiated by one participant, which is mapped to the A argument. The P argument is often but not always affected by the event.

(9)	*iloa*	'find out, know'	*talosia*	'pray for'
	va'aia	'look at'	*la'asia*	'step over'
	tagofia	'intentionally touch'	*fulisia*	'turn over'
	fa'agalaia	'desire'	*inumia*	'drink'
	fofōia	'massage'	*mana'omia*	'need'
	lafoia	'throw, cast away'	*masalomia*	'suspect'
	sōloia	'move s.t. forward'	*si'omia*	'surround'
	usuia	'woo'	*fa'alanumia*	'wash off'
	tāgisia	'cry over'		

The primary focus of this paper is the verbs in the category exemplified in (8). The remainder of this section is devoted to presenting the data which I argue are evidence for treating S and P distinctly for the purposes of the assignment of morphological case.

3.1 The syncretism of S and P

I now present the key data demonstrating that S and P pattern differently in terms of the range of morphological cases available to them. Having demonstrated this, I argue that

[3] The suffix is termed -*Cia*, where *C* is a lexically specified consonant. The initial consonant of the suffix, referred to as the thematic consonant in Mosel and Hovdhaugen (1992), varies lexically.

any analysis of Samoan must not lead to generalisations which treat S and P as a natural class for the purposes of morphological case assignment. The data contradict any analysis which posits that S and P take a single morphological case (absolutive) whose assignment is mediated by a single mechanism.

Legate proposes a typological split in ergative languages: there are absolutive as nominative (ABS=NOM) languages and absolutive as a morphological default (ABS=DEF) languages. ABS=NOM languages are analysed as having a single abstract Case assigned to both intransitive subjects and transitive objects by a unified mechanism. This abstract Case on S and P is labelled nominative. ABS=DEF languages assign different abstract Cases to both S and P. In ABS=DEF languages, S and P receive distinct abstract Cases: S receives nominative and P receives accusative. The appearance of absolutive case is a product of both nominative and accusative being mapped to a default morphological form. I argue that Samoan falls into the category of ABS=DEF languages, where both nominative and accusative are realised by the lack of a case marking preposition.

In both ABS=NOM and ABS=DEF languages, ergative case is an inherent case licensed in the specifier position projected by a functional head determining the transitivity of the clause, termed *v* (or little *v*).

Legate's system crucially differs from any system where absolutive case is assigned by a unified mechanism. For example, Bittner and Hale (1996a) argue that absolutive in Samoan is assigned to both S and P via government by C in finite clauses. Massam (2001) argues that absolutive case in Niuean is licensed on S and P by a dedicated functional head Abs, embedded below *v*. I will demonstrate that any such analysis which asserts that a single morphological case is assigned to both S and P has difficulties accounting for Samoan data presented in this section.

3.2 Case in nominalised clauses

This section outlines the syntax of nominalised clauses in Samoan. In particular, I show that P, but not S, is able to retain its phonologically null case marking (ordinarily labelled absolutive) under nominalisation of its selecting predicate. If the unmarked case on S in Samoan is nominative, we predict it to be absent in finite clauses. If the unmarked case on P is accusative, it should have no interaction with the finiteness of the clause.

Under Legate's system, supported with data from Warlpiri, Enga, Niuean and Hindi, nominative and accusative case are licensed by distinct functional heads. Nominative is assigned to subjects by T (the functional head controlling tense), and accusative is assigned to transitive objects by *v* (the functional head controlling transitivity). Her key piece of evidence for this proposal is the distinct case distributions in non-finite contexts. The languages investigated do not allow S to take the morphological case ordinarily labelled absolutive in a non-finite clause, while P is free to take this morphological case. According to Legate, analyses where S and P receive case via the same mechanism are unable to extend to these data.

Samoan frequently employs nominalisations. Most often exclamative clauses and clauses which set the background state of affairs in narratives are nominalised (Mosel 1991b). Samoan nominalised clauses are similar to verbal clauses in terms of constituent order. Nominalised clauses never contain tense marking morphemes. Rather, the clause is preceded by any of the articles found in ordinary DPs. In (10a), the clause is preceded by the determiner *le*. Tense marking is categorically excluded from appearing within the clause. Compare (10a) with its finite counterpart (10b), where tense marking is available.

(10) a. *le* *(*sā)* *kī-ina* *o* *le* *leitio*
the PAST turn.on-INA GEN the radio
'The turning on of the radio.' (Mosel and Hovdhaugen 1992: 533)

b. *Sā* *kī-ina* *le* *leitio*
PAST turn.on-INA the radio
'The radio was turned on.'

Although nominalised clauses are never tensed, they may include various other functional material. For example, nominalised clauses may contain negation (11a). Negation (realised by the particle *lē*) occurs in the same position as in finite clause counterparts (11b).

(11) a. *lona* *lē* *fia* *'ai*
her NEG want eat
'Her not wanting to eat.'

b. *Sā* *lē* *fia* *'ai*
PAST NEG want eat
'She didn't want to eat.' (Mosel and Hovdhaugen 1992: 558)

The nominalised verb retains the aspectual morphology typical of verbs in finite clauses, such as reduplication for a frequentative interpretation.

(12) *le* *oteote* *o* *tinā* *o* *tamaiti*
the scold.REDUP GEN mother.PL GEN child.PL
'The scolding by the children's mothers.' (Mosel and Hovdhaugen 1992: 533)

The distribution of morphological case on arguments within a nominalisation differs from the distribution of morphological case in a finite clause. In order to make the comparison, I partially repeat the paradigm from (1) below. The case pattern in finite clauses clearly fits the definition of an ergative alignment: A receives a marked case, while P and S are both unmarked.

(14) a. 'Olo'o fafao e le tama le pusafa'i
PROG pack ERG the boy the banana-case
'The boy is packing the banana-case.'(Milner 1976; 59)

b. 'Ua to'a le vai
PERF settle the water
'The water settled down.' (Milner 1976; 269)

Where S may take the morphologically null case in a finite clause (as in 14b), it may not under nominalisation. Under nominalisation, S must take one of the two genitive case markers, the inalienable genitive marker *o* or alienable *a*. Examples (15) and (16) show two nominalised intransitive predicates which co-occur with an obligatorily genitive marked S argument. The S argument in (15) takes *a*, while the S argument in (16) takes *o*.

(15) *'ua* *i'i* *vale* *[le* *fetagisi* **(a)* *namu]*
PERF squeak stupid the cry.PL GEN.ALIEN mosquito.PL
'The cry of the mosquitoes was a stupid squeak.' (Mosel and Hovdhaugen 1992: 542)

(16) *le* *taunu'u* **(o)* *le* *ulua'i* *misionare* *'o* *Ioane*
the arrive GEN the first missionary TOP John
'The arrival of the first missionary, John.' (P. L. Tauiliili, 2009, *Anoafale o le Gagana ma le Aganu'u*, p. 19)

The choice between *o* or *a* appears to depend roughly on the agentivity of the argument. The prediction is that the agentive sole participants of unergative predicates appear with the alienable genitive marker *a*, while patientive sole participants of unaccusative and stative predicates would appear with the inalienable genitive marker *o*. The alienable genitive marker *a* is associated with arguments denoting participants which can exert control or agency over the event, initiate or cause the event, are animate and/or are propelled by self-directed motion. Roughly speaking, the choice between the *a* and *o* genitive could be considered a first pass diagnostic for determining whether an intransitive verb is unergative or unaccusative.

For a more thorough understanding of the distribution of *o* and *a* on S amongst nominalised intransitive predicates, consider the following lists of verbs. The verbs which appear with an *o*-marked S under nominalisation often denote change of state events, or stative properties of S. The verbs which appear with an *a*-marked S include active processes enacted by the S participant, including but not limited to manners of motion, bodily functions, and sound and light emission. A more thorough discussion of this distinction is a topic for future research.

(17) **Intransitive Verbs with *o*-Marked S in nominalisations**

a'alogo	'hear'	*leai*	'lack, not exist'
pala	'rot'	*taunu'u*	'arrive'
mūmū	'burn'	*maualuga*	'be high'
gagau	'break'	*alofa*	'be kind, feel love'
lelei	'be good'	*mago*	'be dry'

(18) **Intransitive Verbs with *a*-Marked S in nominalisations**

pese	'sing'	*sau*	'come'
fa'alogo mai	'listen to'	*tagi*	'cry'
fetagisi	'cry, whine'	*tusi*	'write'
u'u	'make hollow sound'	*susu*	'suck'
alu	'go'	*galue*	'work'

Within nominalised clauses, P may also take genitive case. Earlier, I stated that P is causally affected or a force recipient. It therefore does not fit the prototype for an argument which takes the alienable genitive marker *a*. In fact, if P is marked with genitive, it is the *o* genitive. In (19), the P arguments are marked with the *o* genitive.

(19) a. *le* *fafaga* ***o*** ***le*** ***pepe*** *i* *le* *fagu* *susu*
the feed GEN the baby LOC the bottle milk
'The feeding of the baby with the milk bottle.' (Mosel and Hovdhaugen 1992: 546)

b. *'O le ala lena 'o [le fau* ***o ni potu]***
TOP the reason that TOP the build GEN some room
'It is for that reason some rooms are built.' (Mosel and Hovdhaugen 1992: 546)

Unlike the sole argument of an intransitive, P may take the morphologically null case in a nominalisation if it does not take genitive. In (20), the P arguments of the nominalised transitive predicates appear without any morphological case marking.

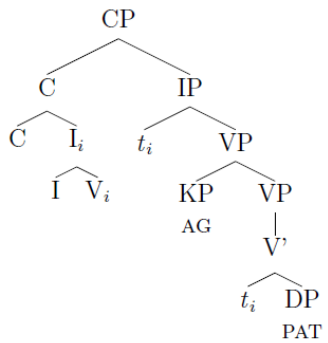
(20) a. *'Ole'ā lē fa'atauina le masini e [la'u 'ai* ***le vaumago]***
FUT NEG sell.INA the machine ERG my grip the straw
'The machine will not get sold by my gripping the straw.' (https://www.lds.org/general-conference/2010/10/the-divine-gift-of-gratitude?lang=zho&clang=smo)

b. *E matamata le tamaitiiti i [le siiane e lona tama* ***le matatao]***
PRES watch the child DAT the lift up ERG his father the spear
'The child watches his father lifting up the spear.' (Mosel and Hovdhaugen 1992: 546) lit. 'The child watches the lifting the spear by his father.'

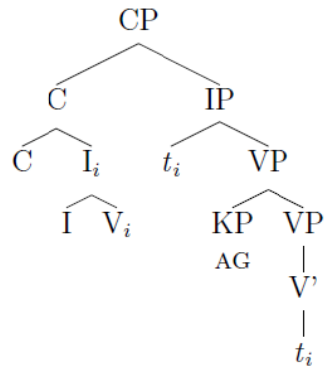
Here is a clear instance where the grammatical properties of S and P diverge. Where intransitive sole arguments must take genitive case under nominalisation, transitive patients may optionally take genitive case or the morphologically null case. This pattern is not predicted by previous accounts such as the model for Samoan proposed by Bittner and Hale (1996a) or the Massam (2001) model for Niuean.

Bittner and Hale (1996a) posit the following structures for Samoan transitive clauses (21a) and intransitive clauses (21b).

(21) a. Transitive



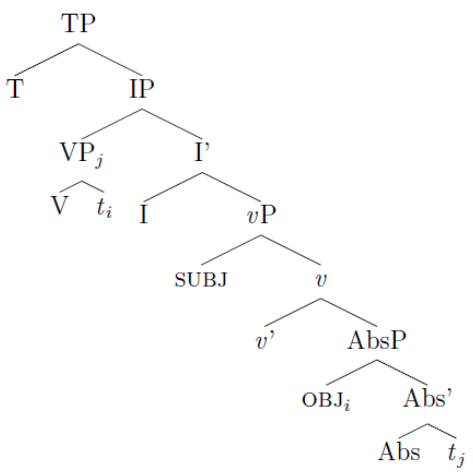
b. Intransitive



Bittner and Hale have the morphologically null case on S and P (which they identify as nominative) assigned by C. For them (contra Massam (2001)), verb-initial word order is derived via head movement of V to I to C. Head movement of the verb all the way to C renders the whole clause transparent to government by C. As nominative is licensed by C, nominative case is licensed on both S and P. This predicts that the distribution of the morphologically null case on S and P should pattern together in nominalised and finite clauses alike. In finite clauses, S and P are both governed C and both receive nominative case. In nominalised clauses, S and P are both governed by a functional head K, which Bittner and Hale propose is also a nominative case licensor. Therefore, they do not predict any split in the distribution of the morphologically null case on S and P.

Under Massam's (2001) system, absolutive case is assigned by a dedicated functional head, Abs, within the *v*P. The structure in (22) represents Massam's proposal for clause structure and case assignment. For Massam, verb-initial word order is derived via fronting the VP constituent to a specifier position above the subject. Note that in her system, the object also moves to a position outside of the VP.

(22)



Massam proposes that the morphologically null case in Niuean (which she labels absolutive) is licensed in one of two ways. Firstly, the internal argument in a transitive clause or an unaccusative clause raises to the specifier position of AbsP, in which it receives absolutive case via specifier-head agreement with the functional head Abs. The external argument of an unergative intransitive also receives absolutive case however, as she assumes that unergative subjects are underlyingly positioned in the specifier of *v*P. Massam maintains that this position too assigns absolutive case. By these assumptions, absolutive case must be available within any vP. Massam's clause structure requires the *v*P constituent to appear in both nominalised and finite clauses, therefore absolutive should be freely available in both nominalised and non-nominalised clauses, contrary to what we see.

However, the observed pattern of data presented in this section is entirely consistent with Legate's proposal. For Legate, the case ordinarily termed absolutive is actually accusative on P and nominative on S. Nominative case is a structural case assigned by T, the functional category determining the tense of the clause. If the morphologically null case on S is nominative, it is therefore expected that S may not take the morphologically null case in any non-finite syntactic environment, such as a nominalisation.

In order to import Legate's analysis to Samoan, I take T to be instantiated by the class of clause initial tense marking free morphemes. The obligatory absence of these particles in a nominalised clause correlates with the obligatory absence of the morphologically null

case on the sole arguments of intransitives in examples like (15) and (16). If the morphologically null case on S is defined as nominative case assigned by T, this correlation is predicted.

Furthermore, if the morphologically null case on P is not dependent on T, it is expected that the absence of T in a nominalised clause will have no effect on the patient's case. Legate's system has accusative case assigned by transitive *v*, a functional category which assigns an agentive thematic role to the external argument. The prediction is that internal arguments will only receive accusative case if the predicate is able to take an external argument, thereby capturing Burzio's Generalisation. This prediction is supported by the data from nominalised clauses. Wherever the morphologically null case appears on P under nominalisation, the nominalised verb denotes a transitive event with an external argument which denotes an agentive participant (as in 20).

Samoan demonstrates several morphological alternations where verbs gain or lose the ability to co-occur with an external argument. Under nominalisation, these alternations demonstrate that only in instances where the verb co-occurs with an external argument may the internal argument take the morphologically null case. The first alternation is the prefixation of monomorphemic transitives with *ma-* to form an anticausative. When prefixed with *ma-* the predicate denotes an event which is not necessarily initiated by any identifiable participant. Where a transitive predicate is nominalised, the P argument may take the morphologically null case. When the same predicate is anticausativised with *ma-*, the patientive sole argument of the now intransitive predicate must take genitive case. The morphological case available to the bolded arguments in (23) differs based on the transitivity of the predicate, despite the fact that their thematic roles are equivalent.

(23a) demonstrates a nominalisation of a transitive verb, while (23b) is a nominalisation of its anticausativised variant.

(23) a. *Sā matamata le teine 'i [la'u goto* ***le va'a]***
PAST write the girl DAT my sink the boat
'The girl watched me sink the boat.'

b. *E gata ai [le magoto *(o)* ***se va'a]***
PRES complete LOC the ANTICAUS.sink PAST a boat
'A boat has sunk there.' lit. 'the sinking of a boat has completed there.' (http://208.109.238.104/viewstory.php?storyid=30342)

The second alternation is causativisation of an intransitive with the prefix *fa'a* to form a transitive. With this prefix, the verb denotes an event necessarily caused by a participant, denoted by the A argument. Where the non-prefixed form is nominalised, the internal argument must take genitive case. Where the causativised form is nominalised, the internal argument may take the null case. Again, the morphological case available to the bolded arguments in (24) differs based on whether the predicate is causativised. (24a) demonstrates a nominalisation of an intransitive verb, while (24b) is a nominalisation of its causativised variant.

(24) a. *Fa'afetai 'i [le tupu *(o)* ***le atunu'u Sāmoa]***
thanks DAT the grow GEN the country Samoa
'Thanks for the growth of the country of Samoa.' (www.youtube.com/all comments?v=1NMOt vY-dQ)

b. *le* *fa'atupu* ***le*** ***taofiofi*** ***o*** ***tagata***
the CAUS.grow the restraint GEN person.PL
'The raising the restraint of the people.'
(http://42976.activeboard.com/t47860333/)

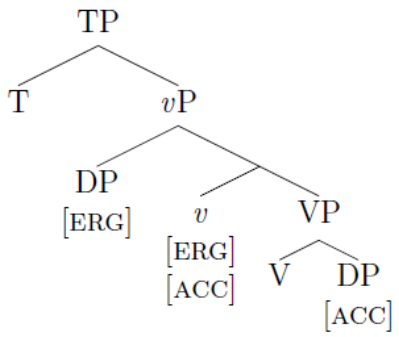
Under Legate's system, the null case on internal arguments is accusative case assigned by *v*. The data in (20)-(24) support this hypothesis. Only in instances where the verb-form co-occurs with an A argument is accusative case available. However, although accusative case is available for the P argument, it is not obligatory. The examples in (19) demonstrate that P is able to take genitive case in a nominalised clause. We can therefore conclude that the assignment of accusative case in a nominalised clause is governed by an optional rule.

The next major element of Legate's analysis is that ergative case is an inherent case licensed in the specifier position projected by *v* in a transitive clause, following Woolford (1997). As with accusative case, this hypothesis predicts that ergative case is insensitive to the presence of T within a clause. It is therefore predicted that ergative case is licensed within nominalised clauses. It is simple to find data confirming this prediction. In (25), the A argument (*e tagata*) is able to take ergative case within a nominalisation.

(25) *le* *uluai* *fa'apāina* *o* *le* *pulu* *atomika* *e* *tagata*
the first CAUS.explode.INA GEN the bomb atomic ERG person.PL
'The first expoding of the atomic bomb by people.' (Mosel and Hovdhaugen 1992: 544)

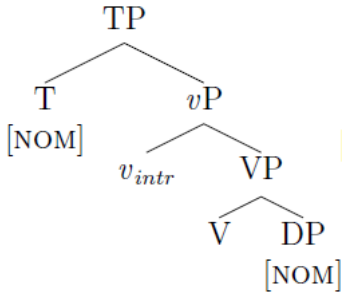
Schematising Legate's (2008) proposal structurally, in a transitive clause, ergative and accusative are licensed by the same source, the functional head which determines the transitivity of the clause, *v*. As the two arguments in a transitive clause take accusative and ergative case, the nominative case is left unassigned by T.

(26)

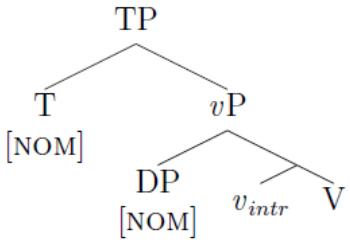


Legate's model relies on multiple variants of the functional category *v*. She posits a transitive variant v_{tr}, which licenses ergative and accusative, and an intransitive variant, v_{intr}, which does not license any case. The trees in (27) are Legate's proposed structures for intransitive predicates (unaccusative in (27a) and unergative in (27b)). In both structures, the S argument does not receive case from v, and therefore must receive nominative case from T.

(27) a. Unaccusative



b. Unergative



If it is assumed that finiteness is determined by functional heads higher than *v*P, and that ergative and accusative are licensed *v*P internally, it may be concluded that the licensing of both ergative and accusative is insensitive to whether the clause is finite or nominalised. As nominative is licensed by T, it is licensed only in finite clauses (where T is present), and not licensed in nominalised clauses (which lack T).

3.3 Case and pronouns

The hypothesis that nominative and accusative are two distinct cases in Samoan with two distinct assignment mechanisms neatly accounts for a large range of data relating to the morphological forms of Samoan pronouns. In this section I briefly summarise that data set and show that instances where the pronoun cliticises to the tense marking morpheme coincide with the predicted distribution of nominative case.

Pronouns may be realised in the same positions as full DPs, that is, following the verb. In this position, a pronoun is a morphologically free-standing unit, able to be marked with a case-marking preposition. Pronouns may also be realised cliticised to the tense marking morpheme. Pronouns take distinct forms depending on whether they are morphologically free-standing or cliticised to the tense marker. The following table lists the forms for each person and number combination.

Table 3.1

	Morphological Realisation of Personal Pronouns	
	Clitic	Free-standing
1SG	*ʻou*	*aʻu*
2SG	*ʻe*	*ʻoe*
3SG	*ia* or *na*	*ia*
1DU-INC	*tā*	*tāʻua*
1DU-EXC	*mā*	*māʻua*
2DU	*lua*	*ʻoulua* or *lua*

3DU	*lā*	*lā ʻua*
1PL-INC	*tātou*	*tātou*
1PL-EXC	*mātou*	*mātou*
2PL	*ʻoutou*	*ʻoutou* or *tou*
3PL	*lātou*	*lātou*

(data from Mosel and Hovdhaugen 1992)

The distribution of morphological forms of Samoan pronouns appears to be tripartite, with unique instantiations for each of S, P and A.

The S argument of both unaccusatives and unergatives may surface as the pronominal clitic form, forming a prosodic unit with the tense marking morpheme.

(28) a. *ʻOlo ʻo=ʻou* *siva*
PROG=1SG dance
'I am dancing.'

b. *ʻOlo ʻo=ʻou* *pa ʻū*
PROG=1SG fall
'I am falling.'

A pronominal S argument may also appear post-verbally as a morphologically free-standing unit. It may or may not be marked with the particle *ʻo* — identical to the topic/focus marking element identified in section 2.1 (cf. Mosel 1991a: 187).

(29) a. *ʻOlo ʻo* *siva* *(ʻo)* *a ʻu*
PROG dance TOP/FOC 1SG
'I am dancing.'

b. *ʻOlo ʻo* *pa ʻū* *(ʻo)* *a ʻu*
PROG fall TOP/FOC 1SG
'I am falling.'

The P argument of a transitive may also appear post-verbally. The P argument may also optionally appear with the *ʻo* marker. (30a) is a causativised version of (29a). However, a pronominal P argument is never able to surface cliticised to the tense marker (30b).

(30) a. *ʻOlo ʻo* *fa ʻasiva* *e* *le* *ta ʻita ʻi* *(ʻo)* *a ʻu*
PROG CAUS.dance ERG the leader TOP 1SG
'The chief is making me dance.'

b. **ʻOlo ʻo=ʻou* *fa ʻasiva* *e* *le* *ta ʻita ʻi*
PROG=1SG CAUS.dance ERG the leader
'The chief is making me dance.'

A pronominal A argument may occur post-verbally. It may be either marked with *e* or left bare. Native speaker consultants dispreferred sentences where post-verbal A was marked with *ʻo*.

(31) *'Olo'o* *fa'asiva* *(e)* *a'u* *le* *teine*
PROG CAUS.dance ERG 1SG the girl
'I am making the girl dance.'

A pronominal A argument may also appear cliticised to the tense marker (32). In this instance, the transitive verb generally takes the suffix *-ina*, though this is not a categorical rule. Speakers differ on the level of grammaticality of sentences with clitics where the *-ina* suffix is absent, ranging from vaguely ungrammatical to dispreferred but acceptable.

(32) *'Olo'o='ou* *fa'asivaina* *le* *teine*
PROG=1SG CAUS.dance.INA the girl
'I am making the girl dance.'

The following table summarises the possible realisations of S, A and P where they are expressed as pronouns. Only the A argument is ever marked with the preposition *e*, and only S and P are marked with *'o*. To paraphrase, the appearance of *e* and *'o* marking on an argument relies on the conjunction of the transitivity of the argument's predicate and whether the argument is or is not an external argument.

Case Possibilities for Pronouns

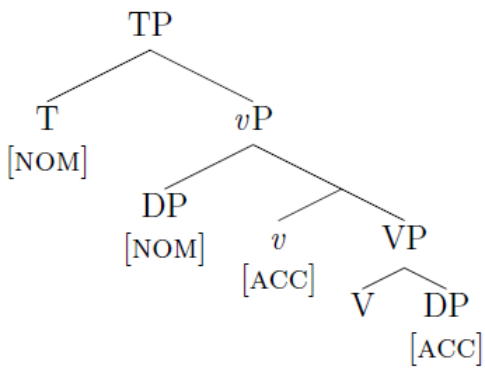
Table 3.2

	S	P	A
Bare + Post-Verbal	Yes	Yes	Yes
Clitic	Yes	No	Yes (with *-ina*)
e DP	No	No	Yes
'o DP	Yes	Yes	No

The data summarised in this table is consistent with a hypothesis where the clitic form is the pronominal realisation of nominative morphological case, while the post-verbal bare form is the pronominal realisation of accusative case. Only S and A are able to appear in the clitic form, while only P is able to appear in the post-verbal bare form. Under the assumption that both S and P take a single, absolutive case, the divergence in the morphological forms of S and P is entirely unexpected.

As with the nominalised clause data, the pronominal data is consistent with Legate's proposal. Recall that Legate has nominative case assigned by T and accusative by *v* in her system. Given that these two assignments of cases are determined by independent functional categories, there is nothing ruling out their co-occurrence in a configuration like in (33) below.

(33)



Clauses such as (32), with A realised as a subject clitic, have no instantiation of the ergative morphological case. Both A and P appear in unmarked forms. Pronominal A is realised in the same morphological form as pronominal S. This pattern is predicted if (32) instantiates the structure in (33), where A takes nominative case and P takes accusative.

The data in the table 3.2, where S, A and P have unique morphological instantiations, is consistent with a hypothesis that all three of nominative, accusative and ergative are available to be licensed. The clitic form of the pronoun is licensed precisely where nominative case is predicted to be licensed, on S or A in a finite clause. The morphological paradigm in which free-standing differ from clitic pronouns is understood as the differing morphological realisation of accusative and nominative case.

A surprising fact emerging from the data in this section is that A, when expressed as a pronoun, is often realised as a nominative pronoun, rather than with ergative case, for example the clause in (32). This is a productive alternation in which A is variously realised with nominative or ergative case. This alternation is explored in the next section.

3.4 Optional Ergative Marking

In the structure (33), nominative and accusative may both be assigned within the same clause to A and P respectively, as with a typical nominative-accusative language. Clauses where A and P are expressed as what I analyse as nominative and accusative pronouns respectively are common in Samoan.

(34) *Sā=‘outou* *sasaina* *a‘u*
PAST=2PL hit.INA 1SG
‘You (pl.) hit me.’

Given the assumption that nominative and accusative are not morphologically distinguished on non-pronominal DPs, there is now a principled explanation for the apparent optionality of ergative case in finite clauses. In a transitive clause, the A argument may appear with or without the ergative preposition *e*.

(35) *Sā* *fa‘amoe* *(e)* *le* *tinā* *lana* *pepe*
PAST CAUS.sleep ERG the mother her baby
‘The mother put her baby to sleep.’

This alternation is especially common in more casual registers of the language. The dropping of ergative often coincides with other features of casual Samoan speech, most prominently the replacement of all instances of alveolar stops with velar stops (commonly represented in writing). Example (36) shows both the dropping of ergative and the replacement of alveolar stops with velar stops.

(36) *E* *kau* *fa‘asiva* *lo‘u* *keige* *lea* *sole*
PRES try CAUS.dance my girl that lad
‘My girlfriend tries to make that lad dance.’ (www.veengle.com/s/Keige‘s/3.html)

I propose that examples like (36) realise A with nominative case and P with accusative case. As nominative is realised without any overt preposition, it gives the appearance that the ergative preposition marking A in the basic clause type has been dropped. Under this view, these types of clauses have a clear analogy with clauses such as (34), where nominative and accusative pronouns are morphologically distinguished.

This data is problematic for a system where absolutive case is assigned via a unified mechanism to S and P. These analyses must account for why their absolutive case is assigned twice in these clauses. One possible explanation is that the ergative preposition is elided by some process of phonological reduction. This hypothesis loses the theoretically appealing connection between such clauses and clauses with nominative-accusative pronouns like (34). Furthermore, the hypothesis is contradicted by data from nominalised clauses. Recall that nominalised clauses may contain ergative marked A arguments. (37) repeats (20b), while (38) is a finite paraphrase of the nominalised clause.

(37) *E matamata le tamaitiiti i le si'i ane e lona tama le matatao*
PRES watch the child DAT the lift up ERG his father the spear
'The child watches his father lifting up the spear.' (Mosel and Hovdhaugen 1992: 546) lit. 'The child watches the lifting the spear by his father.'

(38) *Sā ane (e) lona tama le matatao*
PAST up ERG his father the spear
'His father lifted up the spear'

In nominalised clauses, the ergative preposition may not be 'dropped'. If the dropping of ergative were a phonological phenomenon, this restriction is unexpected as the phonological environment of the ergative preposition would be identical in (37) and (38). However, if the paradigm is instead characterised as a productive alternation between nominative and ergative case, this pattern is expected; nominative case is unavailable in a nominalised clause and therefore the A argument is forced to take an overt morphological case marker.

4 Summary

Data from the morphological forms of pronouns, the distribution of case marking prepositions in nominalised clauses and the ability of transitive agents to 'drop' their ergative preposition in finite tensed clauses give evidence in favour of the hypothesis that the grammar or Samoan does not make reference to any coherent notion of absolutive case. Rather, what is generally taken to be absolutive case is in fact the manifestation of two distinct cases, nominative and accusative, which are morphologically syncretic when marking full DPs, but distinguished when marking pronouns. This analysis demonstrates a three-way case marking system, allowing for ergative, nominative and accusative. As full DPs do not distinguish nominative and accusative, but do distinguish ergative, the appearance of an ergative-absolutive pattern in the morphological case system is derived.

The study suggests that the underlying, abstract system of Case in Samoan is in fact a tripartite system. This suggestion is perhaps surprising considering the relative scarcity of tripartite-aligned systems manifested overtly in morphological case cross-linguistically. The diagnostics in this paper for determining the presence of such an underlying system may lead us to re-evaluate other languages which demonstrate a regular ergative-absolutive pattern in their morphological case system. If similar systems present themselves cross-linguistically, the scarcity of tripartite systems may need to be questioned.

References

Bittner, M. and K. Hale. 1996a. The structural determination of Case and agreement. *Linguistic Inquiry* 27. 1-68.

Bittner, M. and K. Hale. 1996b. Ergativity: Towards a theory of a heterogeneous class. *Linguistic Inquiry* 27. 531-604.

Comrie, B. 1978. Ergativity. In W.P. Lehmann ed., *Syntactic Typology: Studies in the Phenomenology of Language*, 329-394. Austin: University of Texas Press.

Cook K.W. 1991. The search for subject in Samoan. In R. Blust ed., *Currents in Pacific Linguistics: Papers on Austronesian Languages and Ethnolinguistics in Honour of George W. Grace*, 77-98. Pacific Linguistics.

Hopper, P.J. and S.A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56. 251-299.

Legate, J.A. 2008. Morphological and abstract Case. *Linguistic Inquiry* 39. 55-101.

Massam, D. 2001. Pseudo noun incorporation in Niuean. *Natural Language and Linguistic Theory* 19. 153-197.

Milner, G.M. 1976. *Samoan Dictionary*. London: Oxford University Press.

Mosel, U. 1991a. Transitivity and reflexivity in Samoan. *Australian Journal of Linguistics* 11. 175-194.

Mosel, U. 1991b. The Samoan construction of reality. In R. Blust ed., *Currents in Pacific Linguistics: Papers on Austronesian Languages and Ethnolinguistics in Honour of George W. Grace*, 293-303. Pacific Linguistics.

Mosel, U. and E. Hovdhaugen. 1992. *Samoan Reference Grammar*. Oslo: Scandinavian.

Woolford, E. 1997. Four-way case systems: Ergative, nominative, objective and accusative. *Natural Language and Linguistic Theory* 15. 181-227.

6 *Clitic pronouns in Seediq*

ARTHUR HOLMER AND LOREN BILLINGS

1 Background

Modern Seediq, we argue, orders its pronominal clitics relative to each other based on both morphological case (a notion that can also be restated in terms of syntactic relations) and grammatical person. A clitic cluster must begin with a pronoun encoding a speech-act participant (SAP), but if both pronouns encode SAPs, then grammatical person is *not* the relevant property that determines their relative order.[1] Rather, other factors emerge: definitely morphological case (or syntactic relations), possibly semantic roles, and even phonological weight. This collaborative study builds on proposals about Seediq by the first author and places this language within a typological framework developed by his collaborator. We conclude the paper with speculations about how this ordering in Seediq developed: from Proto Austronesian, through Proto Atayalic, into the current situation.

This section situates Seediq geographically and historically. It also presents the constituent order (if there is no more than one clitic pronoun) and pronominal inventories.

Seediq is spoken in north-central Taiwan, in the counties of Nantou and Hualian, from the area near Puli in an arc past Wushe across the Central Mountain Range. The majority of speakers are today found on the Pacific side of the mountains. Seediq is spoken in three dialects: Truku (Hualian, as well as Jingguan in Nantou), Toda (Chunyang and Pingjing, northeast of Wushe, in Nantou; Cholin, in Hualian), and Tgdaya (southwest of Wushe, as well as in Qingliu and Zhongyuan in the Guoxing Valley, all in Nantou). Most of our data come from Tgdaya, but we understand the facts to be representative of the other dialects.[2]

---

[1] We use these additional abbreviations: aug (a binary feature) augmented; AV Actor voice; CNG connegative (discussed in Holmer 2004:124, 126, 2006:86–87, 117 n. 8, 2010:166, 2012:905–907, 912); EXCL exclusive; (f)n. (foot)note; GEN genitive/ergative; INCL inclusive; LV Location voice; NEG negation; NOM nominative/absolutive; pl (a binary feature), PL (a gloss) plural; PRF perfect; PST past; PV Patient voice; Q question; and SG singular. In addition, angle brackets delineate infixes, as in (1c), (2c), (5a), (6b), (9a), and (10a); hyphens indicate other affixal boundaries, including between infixes, as in (5a); and equals signs precede morphosyntactic (though not necessarily phonological) clitics. Our Seediq data are presented in the standard orthography (where <y, ng, g, c> correspond in the International Phonetic Alphabet to /j, ŋ, ɡ, ʦ/, respectively, and italicized <*a*> is identical to the language's only low vowel /a/). We also gratefully acknowledge the assistance of our colleagues Malcolm Ross, Naomi Tsukida, Edith Aldridge, Ross Errington, Douglas M. Fraiser, Haowen Jiang, and Celeste Chia-Yen Lee. Naturally, our greatest debt is to our language consultants Dakis Pawan, Temi Nawi, Pawan Nawi, Habo Mangis, and Temi Puhuk (Seediq), as well as Kibin Samad and Pandikar Padi (Iranun, for the data in table 4 below), without whom this work would have been impossible. Regarding any errors, standard disclaimers apply.

[2] The only dialectal difference we have detected is the acceptability, only in Tgdaya, of =*mian* in apparently free variation with =*nami* '=EXCL1PL'. (See table 1 below.) Without listing examples here, we have confirmed that in this dialect =*mian* can be used wherever =*nami* is used. (For instance, only in Tgdaya can =*mian* precede either of the 3.GEN pronouns in cells G and H of table 2 further below.) Inter-dialectal differences exist, but not ones that hamper the overall applicability of the following exposition.

As for Seediq's prehistory, the Atayalic subgroup of Austronesian "is regarded as self-evident," Blust (1999:46) writes, "and has been adequately demonstrated in[,] e.g., Li (1981)." Ross has since argued for a Nuclear Austronesian primary subgroup of Austronesian, where Atayalic is a primary branch of Nuclear Austronesian (2009:316), noting (2009:315 fn. 31) that the Atayalic member languages "are so similar that their relationship is obvious by inspection." Even more recently, Ross has assessed the evidence for an Atayalic subgroup, concluding that there is "respectable if weak evidence" in its favor (2012:1282). Seediq is then a primary branch of Atayalic, the only other branch being Atayal proper, itself divided into Squliq and Ts'uli' (Li 1981, 1985a, 1985b – each republished in Li 2004). With regard to the issues in this paper, we know of data from Squliq (which is, to our knowledge, internally homogeneous with regard to pronominal ordering) and two quite distinct Ts'uli' subdialects: Plngawan and Mayrinax. (The comparative evidence surrounding Atayalic is further considered in section 3 below.)

Seediq has unmarked subject-final word order for nonpronominal arguments, as in (1a). However, this word-order pattern is clouded by arguments being more frequently realized as clitic pronouns, as (1b–c) show.[3] There are two paradigms, NOM and GEN, of clitics.[4]

(1) a. *M-n-ekan* *bunga* *ka* *Pawan.*
AV-PRF-eat sweet.potato NOM Pawan
'Pawan ate sweet potatoes.' (Holmer 2005:177; cf. also Starosta 1974:336)

b. *M-n-ekan=**ku*** *bunga.*
AV-PRF-eat=1SG.NOM sweet.potato
'I ate sweet potatoes.' (Holmer 2012:906)

c. *P<n>uq-an=**mu*** *ka* *bunga.*
<PRF>eat-LV=1SG.GEN NOM sweet.potato
'I ate the sweet potatoes.'

A NOM clitic pronoun encodes the Actor in the so-called Actor voice and the Undergoer in any other voice (i.e., the argument mirrored in the voice affix of the verb). By contrast, a clausal GEN clitic pronoun always encodes the Actor referent (only in a non-Actor voice). Seediq clitic pronouns *co-occur* only in the non-Actor voices. As such, in all pronominal-clitic clusters the GEN and NOM cases encode the Actor and Undergoer, respectively. In this respect, the clitics behave as in most other languages of the Formosan/Philippine type.

In (1b–c), the clitic pronouns attach to the verb. More generally, however, they attach to the first element in the clause, provided that it is a syntactic head. Namely, they attach obligatorily to whichever free-form head is linearly first among negation, tense/aspect markers, subordinators, and interrogative markers, exemplified in (2a–d), respectively.

---

[3] The verb root for 'eat' is suppletive (Holmer 1996b:66 including fn. 2). Suffixation requires /(p)uq-/, shown in (1c), (2b), (3a), and (6b) below (where initial /p/ is obligatory if the PRF infix is also present). Otherwise, /ekan/ is used; see (1a-b), (2d), (4a-b), (5b), and (6a). Unprefixed /ekan/ is the AV.CNG form (not exemplified in this study but see Holmer 1995:76, 2005:192, 194–195, 2006:102, 109, 111).

[4] We follow traditional practice among Austronesianists in referring to these clitic pronouns as NOM and GEN, in part because one of the uses of the latter is as a Possessor (within a nominal expression). Referring to a clitic Possessor as *ergative* is likely to be more confusing than referring to an Actor clitic as *genitive*. However, nothing in our analysis hinges on this choice. In practice (at the clausal level), these NOM and GEN clitic-pronominal paradigms reflect absolutive and ergative arguments, respectively.

(2) a. *Ini=**ku*** *imah* *sino.*
NEG=1SG.NOM drink[AV.CNG] wine
'I don't drink wine.'

b. *Wada=**mu*** *puq-un* *ka* *damac.*
PST=1SG.GEN eat-PV NOM food
'I ate the food.'

c. […] *ado=**ku*** *m-beyax* *t<m>alang* *yaku.*
because=1SG.NOM AV-strong <AV>run 1SG
'[…] because I am good at running.'

d. *Ye=**ku*** *ini* *huwa* *m-ekan* *tmaku* *hini?*
Q=1SG.NOM NEG how.CNG AV-eat tobacco here
'Is it okay if I smoke here?' (Holmer 2005:190)

However, initial elements that are not positioned as heads do not attract clitics, as in (3a-b).

(3) a. *Paye* *nii* *puq-un=**daha*** *klaali,* […]
rice this eat-PV=3PL.GEN always
'They always eat this rice, […]' (Henningsson and Holmer 2008:38)

b. *Kiya* *bkey-an=**daha*** *knraga.*
thus tie-LV=3PL.GEN purlin
'That's how they tie the roof purlins.'

To sum up, the bound pronouns in Seediq attach to the verb, in (1b–c) and (3a–b); NEG, in (2a); tense/aspect markers, in (2b); subordinators, in (2c); and interrogative markers, in (2d). More dramatically, bound GEN pronouns also attach to nouns, to encode the Possessor (e.g., *tama=mu* father=1SG.GEN 'my father'). This promiscuity of attachment strongly supports the idea that these bound pronouns are (morphosyntactic) clitics, contrary to arguments in Chang (1997:97–107, 1999b:356–363) that these forms are affixes.[5]

Table 1 shows the repertoire of pronominal paradigms discussed in the remainder of this paper. The formal features in tables 1, 2, and 4 (as well as the tableaux in §3) follow the spirit of McKaughan (1959). If both of the arguments in a clause are represented by pronominal person/number features, the situation is more complex and we delay discussion of it until section 2 below. There are two important points to note about this table. First, as in many other Austronesian languages of the area, in the third person, only the GEN clitic pronouns are overt. Usually third-person NOM pronominal arguments are unexpressed (as, e.g., sentences in Lee 2011:53–54 show). If for whatever reason an overt 3.NOM form is required, the case-neutral free form is used instead, as in (4a–b). The free form is located in the clause in the same position as a nonpronominal subject, in clause-final position.

---

[5] See Yen (2012:53, citing Li 2006:34 fn. 8) and Zwicky and Pullum (1983:503–504) for discussion of promiscuity as a test of clitichood in other languages. Holmer (2002:343) also points out that unlike affixes, bound pronouns in Seediq are not part of the same stress domain as the preceding word, also supporting their analysis as clitics rather than affixes (Zwicky and Pullum 1983:504). Chang's other arguments in support of an affixal analysis have to do with arbitrary gaps in pronominal co-occurrences and the morphological idiosyncrasy of portmanteaux (1997:104–105, 1999b:360, both citing Zwicky and Pullum 1983). Though we consider arbitrary gaps to be a robust test in the world's languages, we argue below in section 2 that the gaps in table 2 are not arbitrary as such. As for the idiosyncrasy criterion, we consider it to be not as probative as promiscuity (to which we know of no exceptions in the literature).

Table 1: Partial pronominal inventory in Seediq

PERSON/NUMBER		**CASE**		
Traditional labels	**Formal features**	**NOM *clitic***	**GEN *clitic***	**NEUTRAL *free***
1SG	[+me, –you, –pl]	=ku	=mu	yaku
EXCL1PL	[+me, –you, +pl]	=nami		yami
INCL1PL	[+me, +you, +pl]	=ta		ita
2SG	[–me, +you, –pl]	=su		isu
2PL	[–me, +you, +pl]	=namu		yamu
3SG	[–me, –you, –pl]	Ø	=na	heya
3PL	[–me, –you, +pl]	Ø	=daha	dheya

(Holmer 1996a:32, 2002:343; rows arranged as in Asai 1953:41–42 and Chen 1992:34–36)

(4) a. *M-n-ekan* *bunga* ***heya***.
AV-PRF-eat sweet.potato 3SG
'She/He ate sweet potatoes.'

b. *M-n-ekan* *bunga* ***dheya***.
AV-PRF-eat sweet.potato 3PL
'They ate sweet potatoes.'

Such free-form pronouns are not case-marked and can equally well be used with an object relation, as in (5a). Free forms for first- and second-person pronouns also exist, as in (5b), and when they are used, they co-occur with the clitic pronoun of the same person/number.

(5) a. *Q<m-n>ita=**ku*** ***heya***.
<AV-PRF>-see=1SG.NOM 3SG
'I saw her/him.'

b. *M-n-ekan=**ku*** *bunga* ***yaku***.
AV-PRF-eat=1SG.NOM sweet.potato 1SG
'Me, I ate sweet potatoes.'

Next, all clitic forms in the first and second persons except 1SG syncretize for NOM and GEN, as indicated in table 1 by the yellow highlighting (also used in tables 2 and 4 below). Thus, without the context established primarily by the voice of the verb, a form like *=nami* '=EXCL1PL' is ambiguous as to whether it encodes NOM, as in (6a), or GEN, in (6b).

(6) a. *M-n-ekan=**nami*** *bunga.*
AV-PRF-eat=EXCL1PL sweet.potato
'We ate sweet potatoes.'

b. *P<n>uq-an=**nami*** *ka* *bunga* *kiya.*
<PRF>eat-LV=EXCL1PL NOM sweet.potato that
'We ate that sweet potato.'

This concludes our preliminaries. The remainder of this paper is devoted to the relative order – within the clitic cluster – of two clause-mate pronouns, as well as the complicating factor posed by portmanteau pronouns. We present the ordering facts in section 2 then explain them in section 3 from typological, theoretical, and historical perspectives.

2 Pronominal-clitic combinations

This section starts by introducing the idea of how a pronominal-combination table is useful as a method of sorting out how bound pronominal forms are ordered relative to each other. Next, such a table is presented, illustrating the facts in Seediq. Various details are then exemplified: twelve combinations with overt co-occurrences of two clitic pronouns, four of person/number features where no clitic cluster or portmanteau is possible, and three where pronominal portmanteaux are used. (One combination allows two of these types.)

Studies growing out of the tagmemic framework (chronologically: Pike and Erickson 1964 and Erickson 1965), known as Matrix Theory, have argued for principled decisions regarding the order in which the rows and columns of two-dimensional tabulations (or matrices) of features or forms are arranged to describe a language. Pittman (1965) also uses such tabulations without going into the benefits of such decisions (or even citing the previous studies). Using mostly data from Algonquian languages, Pike and Erickson (1964:203, 208–209), Erickson (1965:227), and Pittman (1965:36–37) show arrangements of pronominal forms in which the rows and columns encoding [+me] are consecutive and the rows and columns encoding [+you] are likewise contiguous. Pittman's table uses the order as in the rows of our pronoun inventory above (in table 1), with the [+me] rows/columns preceding the [+you] ones (where the INCL1PL rows/columns are the last of [+me] and the first of [+you]) and the [–me, –you] rows/columns appear last. (We also use this arrangement below in tables 2 and 4.) The other two studies actually show the [+you] rows/columns *preceding* the [+me] ones: arguably the preferable order for Algonquian, where [+me] morpheme is overt only if there is no [+you] feature present. As it were, the [+you] morpheme displaces the [+me] form. The principle shared by these studies and ours is that the rows and columns are ordered primarily according to person rather than number.

The same matrix-theoretic ordering principles are evident in our table 2. For instance, if both arguments in the clause must be represented by person/number features (rather than, say, by nouns), then there are restrictions on the possible combinations of clitic pronouns actually realized. Consider first the seventeen empty cells with diagonal lines in table 2. The block of nine cells in the upper, left-hand corner are the intersection of [+me] (i.e., both above and to the left). Similarly, the nine mid-table cells are the intersection of [+you]. (In one cell both intersections are found, hence diagonal lines in both directions. As such, there are seventeen rather than eighteen such empty cells because these two blocks overlap by one cell.) If either (i) both entities are [+me] or (ii) both are [+you], then no cluster of clitic pronouns is possible. This is true in numerous Austronesian languages and perhaps universal (as in the Algonquian pronominal matrices alluded to in the preceding paragraph). In table 2 any GEN pronouns are underlined; any NOM pronouns, in bold type. In the subsequent discussion, we refer to the cells of this table using capital letters.[6] Only in cell K is there variation, discussed separately below in this section.[7]

---

[6] Keeping the third-person rows together in our table 1 also follows matrix-theory principles since the lack of overt NOM forms (or of clitic combinations in tables 2 and 4 below) also results in a contiguous block.

[7] Another potential source of variation is the cluster of =*ku* '=1SG.NOM' preceding =*su* '=2SG', reported as acceptable only in Chang (1997:74–75, 1999a:623–624). For instance, neither Chang (1999b), a study devoted in large part to pronominal ordering, nor Chang (2000), a reference grammar, mentions this,

Table 2: Combinations of clitic pronouns in Seediq

NOM \ GEN		+me –you –pl	+me –you +pl	+me +you +pl	–me +you –pl	–me +you +pl	–me –you –pl	–me –you +pl
		=mu	=nami	=ta	=su	=namu	=na	=daha
+me –you –pl	**=ku**				A =saku *=**ku**=su *=su=**ku**	B =**ku**=namu *=namu=**ku**	C =**ku**=na *=na=**ku**	D =**ku**=daha *=daha=**ku**
+me –you +pl	**=nami**				E *=**nami**=su *=su=**nami**	F *=**nami**=namu *=namu=**nami**	G =**nami**=na *=na=**nami**	H =**nami**=daha *=daha=**nami**
+me +you +pl	**=ta**						I =**ta**=na *=na=**ta**	J =**ta**=daha *=daha=**ta**
–me +you –pl	**=su**	K 1 =**su**=mu 2 =misu *=mu=**su**	L *=**su**=nami *=nami=**su**				M =**su**=na *=na=**su**	N =**su**=daha *=daha=**su**
–me +you +pl	**=namu**	O =maku *=**namu**=mu *=mu=**namu**	P *=**namu**=nami *=nami=**namu**				Q =**namu**=na *=na=**namu**	R =**namu**=daha *=daha=**namu**
–me –you ±pl	**Ø**	N/A	N/A	N/A	N/A	N/A	N/A	N/A

[Based on another such table in Li (2010:28). Sources: [A] Chang (1997:104, 1999b:360), Holmer (1996a:70, 89), Tseng and Sung (2010:187), and Tsukida (2005:303); cf. Holmer (2002:343). [B] Chang (2000:81 fn. 6), Tseng (2009:128, forthcoming), (10a) below, and (iv) in fn. 11. [C] Aldridge (2004:76, 83), Chang (1997:99 [cited in Holmer 1999:10], 109 [cited in Chang 2010:203], 190 [cited in Holmer 1999:10], 1999b:356), Holmer (1995:85, 1996a:63, 69, 87, 111, 114, 147, 2002:344, 2012:906, 909), Tseng (2009:36, 154, 2010:97, 103, forthcoming), and Tseng and Sung (2010:187). [D] Aldridge (2004:41, 43, 214), Chang (1997:107, 1999b: 365), Tseng (2009:37, 2010:42, 203), and (9a) below. [E] See (7a–b) below. [F] See (8a–b) below. [G] Tseng (forthcoming). [H] *Truku Bible* (Acts 16:37), via Naomi Tsukida (personal communication, 1 October 2006). [I] See (9b) below. [J] Tseng (2009:37). [K1] Asai (1953:46), Chang (1997:20, 83, 100, 102, 108, 110, 1999a:630, 1999b:357–359, 366–367, 2000:114, 135 [cited in Holmer 2010:178]), Chen (1992:39), Holmer (1995:85, 1996a:42, 69, 110, 2002:336), Lin (2005:125), Tseng (2009:26, 31, 2010:92), and (10b) below; cf. Holmer (1996a:70). [K2] Chang (1997:109, 1999b:366) and Holmer (1996a:71, 2004:126, 2006:90 [cited in Chang 2010:198]), Tseng and Sung (2010:187), and Tsukida (2005:309); cf. Holmer (1996a:70, 2002:343–344) and Tsukida (2005:303). [L] See (7a–b) below. [M] Chang (2000:81), Tseng (2009:145, 174, forthcoming), and Tseng and Sung (2010:187). [N] Chang (1997:107, 1999b:365) and Tseng (forthcoming). [O] Chang (1997:104, 108, 1999b:360, 366), Tseng (2010:93), and Tseng and Sung (2010:187); cf. Holmer (1996a:70, 89, 2002: 343) and Tsukida (2005:303). [P] See (8a–b) below. [Q] Tseng (2009:175). [R] Holmer (1996a:70, 2002:344, 351).]

---

though the opposite order, *=*su*=*ku* '=2SG=1SG.NOM', is listed as unacceptable (1999b:360) – along with *=*namu*=*mu* '=2PL=1SG.GEN', *=*mu*=*su* '=1SG.GEN=2SG', *=*na*=*ku* '=3SG.GEN=1SG.NOM', and *=*na*=*su* '=3SG.GEN=2SG' (1999b:360, 366, 2000:81). The speakers whom we consulted have consistently rejected *=*ku*=*su*. In tables 2 and 3 we list this order as unacceptable, reflecting their judgments. See also Ochiai (2009:56). However, if some Seediq speakers accept it, there is no deleterious effect on our proposals in section 3 below. (Both it and =*su*=*mu* '=2SG=1SG.GEN' behave alike in the relevant clustering respects.) Ochiai (2009:57, 62, 65) also reports a 2SG>EXCL1PL portmanteau; the speakers we consulted rejected it.

The second co-occurrence restriction is unique (as far as we know) to Seediq. No combination of two syncretic pronouns is grammatical, as indicated by the yellow shading in table 2. Now, most yellow-shaded cells are empty there because of the aforementioned restriction represented by the diagonal lines. Only four cells (E, F, L, and P: i.e., those with both orders unacceptable, shown in red type, where *=nami* '=EXCL1PL' might combine with either *=su* '=2SG' or *=namu* '=2PL') still require explanation. Namely, all four orders in (7a–b) and (8a–b) are unacceptable under either mapping of roles to cases, represented by the unacceptability of *both* free translations for each. (That the yellow-shaded cells are a contiguous block in table 2 is also consistent with the principles of Matrix Theory.)

(7) a. **Qta-un***=nami=su**.*
see-PV=EXCL1PL=2SG
('We see you.'/'You see us.')

b. **Qta-un***=su=nami**.*
see-PV=2SG=EXCL1PL
('You see us.'/'We see you.')

(8) a. **Qta-un***=nami=namu**.*
see-PV=EXCL1PL=2PL
('We see you.'/'You see us.')

b. **Qta-un***=namu=nami**.*
see-PV=2PL=EXCL1PL
('You see us.'/'We see you.')

The lack of overt clitic clusters in the yellow-shaded cells of table 2 is reported in Holmer (1996a:70), where it is argued that at least one of the two pronouns must distinguish between its NOM and GEN forms. We merely restate these facts here.[8] These combinations constitute a natural class of gaps. They are gaps in the matrix but clearly not arbitrary ones.

Only the fourteen cells of table 2 that contain acceptable combinations (namely: A–D, G–K, M–O, and Q–R) remain to be discussed. First, as the rightmost two columns of table 2 show, in any cluster that involves a third-person clitic pronoun, that pronoun must be final, in (9a–b). Recall that overt third-person clitic pronouns are always GEN, so this order is actually also invariably NOM before GEN. What exactly the underlying generalization is, be it based on the person features (i.e., SAP preceding third person) or morphological case (i.e., NOM preceding GEN), is an open question that we address in the following section.

---

[8] There is just one correction to the record here. Holmer (1996a:70) writes that instead of structures like (7a–b) or (8a–b) "we are forced to use" an AV verb, as in (i). We have since determined, inspired by the Plngawan data in Chang (2012), that a way exists to express this idea without changing the voice, in (ii).

(i) *Q<**m**-n>ita**=nami*** ***yamu***.
<AV-PRF>see=EXCL1PL 2PL
'We saw you.' (Holmer 1996a:70)

(ii) […] *skxl-**un=nami*** *balay* ***ka*** ***yamu*** […]
like-PV=EXCL1PL really NOM 2PL
'[…] we really like you […]' (*Truku Bible* [I Thessalonians 2:8], via Naomi Tsukida, personal communication, 17 April 2013; our free translation)

In both (i) and (ii) the only clitic is the EXCL1PL Actor (whereas *yamu* '2PL', encoding the Undergoer, is from the case-neutral free paradigm). Only in (ii), *ka* (used to mark NOM case on free forms) precedes *yamu*. We have verified that structures like (ii), relevant to cell P of table 2, are also possible relevant to cells E, F, and L (respectively: …*=su* … *ka yami* … '…=2SG … NOM EXCL1PL …', …*=namu* … *ka yami* … '…=2PL … NOM EXCL1PL …', and …*=nami* … *ka isu* … '…=EXCL1PL … NOM 2SG …'). In this connection, as of our 12-ICAL talk, we were unaware of the Seediq data reported in Ochiai (2009:39, 57, 61–67) analogous to (ii). Ochiai has labeled this the *single-clitic pattern*. Inasmuch as the current study is about *combinations* of clitic pronouns (including portmanteaux), and Ochiai's data differ in several key respects from our findings, predicting even unacceptability in structures like (ii), with either a 2SG or 2PL Undergoer after *ka* (2009:39), we defer until a later study the exact properties of the single-clitic pattern.

(9) a. *B<n>be-an=**ku=daha**.*
<PRF>hit-LV=1SG.NOM=3PL.GEN
'They hit me.' (cf. Chang 1997:107, 1999b:365)

b. *Biq-un=**ta=na*** *clokah.*
give-PV=INCL1PL=3SG.GEN ease
'He will give us relief.' (unpublished church hymn)

The only remaining combinations in table 2 not yet discussed are the four cells where [+me, –you, –pl] intersects with [–me, +you, ±pl]. Of these four, only two cells (namely, B and K) are realized as overt, segmentable clusters and both with NOM-initial internal order:

(10) a. *Q<n>ta-an=**ku=namu**.*
<PRF>see-LV=1SG.NOM=2PL
'You saw me.'
(Holmer 1996a:70 [cited in Chang 1997:109, 1999b:366], 2002:344, 351)

b. *Qta-un=**su=mu***
see-PV=2SG=1SG.GEN
'I will see you.' (Holmer 2002:344; cf. *Qta-un=**misu***, in Holmer 1996a:71)

Parallel to this, there are three opaque portmanteau forms (shown with green highlighting in table 2 above) that combine the two sets of person/number features into a single unsegmentable form. (Note the glossing convention here 'X>Y', read as *X acting upon Y*.)

(11) a. *=saku* '=2SG>1SG' b. *=misu* '=1SG>2SG' c. *=maku* '=1SG>2PL'

These three forms cannot be divided synchronically into separate clitics, contrary to suggestions about (11a) in Chang (2000:81) and about (11b–c) in Chang (1999b:366). Our rationale is twofold. In terms of morphophonology, there is no rule that can convert an overt co-occurrence of two pronouns into these forms. That is, the initial syllables [sa], [mi], and [ma] in (11a–c) only occur in these combinations, not independently as clitics. Moreover, the second syllable [ku] of *=maku*, if it is to be seen as reflecting 1SG, would unexpectedly be NOM rather than GEN. The second reason for these forms' inseparability is syntagmatic, where only the forms in (11a–b) – if somehow analyzed as deriving from /=su/ '=2SG' plus /=ku/ '=1SG.NOM' as well as /=mu/ '=1SG.GEN' plus /=su/ '=2SG', respectively – constitute an exception to the NOM-before-GEN relative ordering of two clitic pronouns. An analysis that considers the forms in (11) as unsegmentable, and therefore stored separately as single morphological forms, thus allows for a single, simple generalization. For these reasons, it is clear that there is a fundamental difference in kind between clitic clusters and portmanteaux. Table 3 summarizes the various realizations.

Table 3: Combinations involving [+me, –you, –pl] and [–me, +you, ±pl]

Participant person/number and roles	Portmanteaux	Overt Clusters
1SG Undergoer + 2SG Actor (cf. cell A)	*=saku*	not attested
1SG Undergoer + 2PL Actor (cf. cell B)	not attested	*=ku=namu* '=1SG.NOM=2PL'
2SG Undergoer + 1SG Actor (cf. cell K)	*=misu*	*=su=mu* '=2SG=1SG.GEN'
2PL Undergoer + 1SG Actor (cf. cell O)	*=maku*	not attested

Nearly every row of table 3 allows either only a portmanteau or only a cluster. Only one row shows variation: either *=misu* '=1SG>2SG' or *=su=mu* '=2SG=1SG.GEN'. Of the 70-odd languages of the area studied in this regard, Seediq is quite unusual in having an optional portmanteau. In most languages, if a portmanteau is attested for a particular combination of person/number features, then the overt cluster is impossible. The only other optional-portmanteau language that we know of is Mantauran Rukai (Yen and Billings 2011:172).

As Chang (1997:107, 1999b:366) observes (in our view, especially if the portmanteaux are disregarded), the two overt clusters in table 3 prohibit a formal analysis of cluster-internal ordering based only on grammatical person: *=ku=namu* '=1SG.NOM=2PL' shows first person preceding second, whereas *=su=mu* '=2SG=1SG.GEN' exhibits the opposite order of persons.[9] Nor is grammatical number of any use at least in the *=su=mu* cluster (cell K) because both pronouns have are marked with the same (i.e., SG) number.[10] The simplest (and thus most economical) characterization of cluster-internal ordering is that the NOM form invariably precedes the GEN one. Because portmanteaux and overt clitic clusters occur only in the voices where the Undergoer encodes the subject, this characterization can also be restated in terms of syntactic relations, with the subject appearing first in the clitic cluster, or in terms of semantic roles, where the Undergoer precedes the Actor.[11]

---

[9] In a different study, Chang argues that second-person clitic pronouns precede first-person ones (and either of these precede third-person ones) but concedes that *=ku=namu* '=1SG.NOM=2PL' is an unexplained exception (2000:81 including fn. 6). Elsewhere, Chang (1997:107 fn. 53, citing unpublished work by Shang-yi Tsai) criticizes such an approach for this very reason. Liao (2004:285–296, 2005:51–53) makes much the same argument about Squliq Atayal. As we discuss below in section 3, the cross-linguistic evidence also does not support such a second-before-first characterization. The only other Atayalic dialect where grammatical person does decide the order of two SAP clitic pronouns, Plngawan Ts'uli' Atayal, requires the first-person clitic pronoun to precede the second-person one (Chang 2012:178–179).

[10] We mention (in fn. 25, §3.3, below) that Huang's analysis of Mayrinax Ts'uli' Atayal (1995a, 1995b) utilizes grammatical number (in addition to person and semantic roles) in ordering pronominal clusters.

[11] Note, however, that the same principles also seem to hold for equational clauses, where the semantic-role labels of *Undergoer* and *Actor* are not relevant. As an answer to (i), either (ii) or (iii) can be used.

(i) *Ima=**saku*** *yaku?*
who=2SG>1SG 1SG
'What relation am I to you?' (literally: 'Your who am I?')

(ii) *Tama=**su=mu*** *isu.*
father=2SG=1SG.GEN 2SG
'You are my father.'

(iii) *Dangi=**misu*** *isu.*
friend=1SG>2SG 2SG
'You are my friend.'

In (i) and (iii) we use the '=X>Y' glossing convention from the rest of this paper even though the meaning is not X *acting upon* Y as such. Rather, X is the noun's Possessor and Y is the Theme/subject of the clause. We have confirmed that right after the initial noun any of the attested combinations in table 2 above are allowed. Inasmuch as the NOM-case pronoun in equational clauses still bears the syntactic relation of subject, within overt clusters like (ii) our generalization that the subject is initial still holds. Nonetheless, newly uncovered data like (iv) cloud the definitions of *subject* and *Undergoer* somewhat:

(iv) *Gaga* *inu* *ka* *patis* [*wada=**ku=namu*** *biq-un* ___]*?*
be.located where NOM book PST=1SG.NOM=2PL give-PV
'Where is the book that you gave to me?'

In relativization (which in Seediq, as in many other Austronesian languages, is subject-oriented) the NOM-case argument (if there were no relativization), upon extraction, necessarily forms a gap. In (iv) *ka patis* 'NOM book' has been extracted. Nevertheless, with ditransitive verbs like 'give' pronominal-clitic clusters can still occur, and the clitic pronoun that is NOM-marked encodes the Recipient. If, as we argue,

One final point about the portmanteaux in Seediq deserves mention here. The existence of any of these portmanteaux in the synchronic lexicon is arbitrary. (For historical reasons perhaps, various frequently used clusters have fused into portmanteaux. However, the modern lexicon contains such forms for no principled reason. They are as arbitrary as any other lexical item, in the Saussurean sense.) As we discuss in the following section, the existence of a portmanteau can entail the absence of an overt cluster. Still, the gap in the co-occurrence matrix is not arbitrary as such. Along with the four gaps in cells E, F, L, and P of table 2 (discussed above in this section), none of the gaps in overt clusters is arbitrary.

To summarize section 2, we have presented the rather complicated facts in Seediq regarding cluster-internal clitic ordering and portmanteaux. One co-occurrence restriction, apparently unique to Seediq, prohibits a clitic cluster if neither clitic pronoun exhibits a distinction between NOM and GEN. In certain other combinations portmanteau pronouns are found. The one characterization that captures all overt co-occurrences of clitics in Seediq is that the NOM/subject/Undergoer pronoun is invariably initial in the clitic cluster.

3 Typological, theoretical, and historical considerations

Whereas the preceding section describes the facts in the modern language, here we explore Seediq from three additional perspectives. First, we show that this language belongs to a relatively rare type in the region, along with only three other known languages, where the clitic cluster begins with the subject pronoun. Next, we show that because there are few overt co-occurrences of clitic pronouns in Seediq, it is impossible to formalize the language's grammar sufficiently by looking only at the modern language. As an indirect means of investigation, in the final subsection we compare Seediq to three of its closest relatives in order to see at least how this language came to be the way it is. We show that there is evidence for lexical replacements but not grammatical changes as such.

3.1 Typology

Quite a few languages of Taiwan, the Philippines, and nearby parts of Malaysia and Indonesia – an area henceforth dubbed *northern Austronesian* (following Donohue 2007) – have been examined with regard to cluster-internal ordering of clitic pronouns.[12] These languages share the property of attesting two paradigms of clitic pronouns corresponding to the Actor and Undergoer. In all except for Rukai (southern Taiwan) and Bunun (central and southern Taiwan), GEN and NOM encode these two semantic macroroles, respectively. These languages order two pronouns within the clitic cluster based on semantic roles, syntactic relations (or morphological case), grammatical person, or even prosodic weight.

We begin by sketching the languages that might use semantic roles, syntactic relations, or morphological cases for such ordering. There are two, essentially opposite, such types.

The clearest case of the first type of ordering is Mantauran Rukai (Yen and Billings 2011), where it is the grammatical subject that is invariably cluster-initial. We know that it

---

the property relevant to ordering two clitics within the cluster is subjecthood, then in (iv) the NOM-marked Recipient must be the subject (at some level of representation). We leave it for future research exactly how the theoretical details are to be worked out. Incidentally, Holmer (1996a:89) implies that such an extra NOM argument is possible only if both sets of person/number features form a portmanteau. As (iv) shows, in a relative clause where an Undergoer has been extracted an overt clitic cluster is also possible.

[12] This subsection draws heavily from Hung and Billings (2009/1010:11–13, including footnotes 9–14) and updates the more recent publications that have appeared since that work. Several unpublished talks cited there are not mentioned here, where only published (or at least currently downloadable) works are listed.

is the syntactic relation, subjecthood, that is operational in this language (rather than either semantic roles or morphological cases) for two reasons. First, Mantauran has pronominal clitic clusters in both active and passive clauses. Regardless of the clause's voice, it is the subject pronoun, not necessarily the Actor, that is cluster-initial. Thus, we know that semantic roles are not relevant to ordering in that language. Next, Mantauran utilizes two distinct morphological cases to encode the subject: NOM and GEN. (The latter is used for Possessors within nominal expressions and for the subject of certain kinds of clauses, such as nominalizations, including subordinate clauses, and polar structures, including yes/no-interrogative and negated clauses.) However, regardless of its morphological case, the subject pronoun is invariably cluster-initial. For reasons of economy, Yen and Billings conclude that the pronoun's subjecthood (a syntactic property), and not its morphological case, is what determines the cluster-internal ordering in Mantauran (2011:169). Two other languages that order their pronouns with the NOM-case subject in cluster-initial position are the two Kalamianic languages of the western Philippines (Quakenbush and Ruch 2008). However, in those languages (as in Seediq) it is impossible to eliminate roles and cases as the ordering factor, because if there is a pronominal-clitic cluster, then the subject always encodes the Undergoer role and invariably bears NOM case.[13] The Kalamianic microgroup (i.e., Agutaynen and Kalamian Tagbanwa), Mantauran Rukai, and Seediq are the only known languages in which the subject pronoun is invariably cluster-initial. We know of no language where crucially either the pronoun with the semantic macrorole of Undergoer or the one bearing NOM case must be cluster-initial. Based just on Mantauran, then, in the framework used here the constraint dominant in these languages is dubbed SUBJECT-1ST.

In several other northern-Austronesian languages the cluster-internal ordering is the exact opposite: with the GEN Actor preceding the NOM Undergoer. The languages of this kind that we know of are Kavalan, spoken on Taiwan's Pacific coast (Yen 2012; Yen and Billings, forthcoming); Mamanwa and Tausug, argued to be a subgroup of Central Philippine (Lee and Billings 2008:194–197); and the Central Luzon microgroup (Billings and Kaufman 2004:16–17). Because in these languages the semantic roles of Actor and Undergoer always map respectively to the morphological cases GEN and NOM, it is impossible to tease the roles and cases apart.[14] Still, based on the notion that the relevant constraint is the one that picks the most prominent member of the given morphosemantic

---

[13] As Quakenbush and Ruch point out using Agutaynen data (2008:222), though the NOM/subject/Undergoer pronoun is invariably a morphosyntactic clitic (e.g., appearing between clause-initial NEG and the verb where, as in Seediq, NEG is a free form that precedes the verb), there is *optional* positioning of the Actor pronoun. For example, if the clause is negated, then the Actor pronoun can appear either (i) immediately after the clitic-NOM pronoun and before the verb or (ii) right after the verb. (The latter is far more likely to occur, Quakenbush and Ruch also point out.) Of relevance to the current study is only option (i), with both pronouns preceding the verb. (The Actor pronoun also varies in its morphological form: clitic GEN if it does not follow the NOM/subject/Undergoer pronoun within the clitic cluster and free OBL elsewhere. The latter is an instance of disformation, a notion discussed separately in connection with table 4 below.)

[14] Another language attests similar ordering: Cotabato Manobo (spoken in the southern Philippines) requires both pronouns to appear immediately after the verb (regardless of whether the verb is clause-initial). The GEN Actor pronoun is invariably right after the verb (e.g., even if negated). Similarly to Agutaynen (mentioned in the preceding fn.), in Cotabato Manobo the *Undergoer* pronoun varies in its form, though without changing its case: clitic NOM if it does not follow an Actor/GEN pronoun in the clitic cluster and free NOM if it does. That is, unlike in Iranun (discussed below surrounding table 4) and Agutaynen, where there is at least the option of two clitic pronouns being adjacent, in Cotabato Manobo the NOM Undergoer is never in a clitic cluster. Thus, though ACTOR-1ST is true of this language, the Actor never co-occurs with the other pronoun in its clitic form. Crucial negated data appear in Kerr (1988:11, 35, 69–70), with additional discussion and data from Errington (1979:139, 142, 145–147, personal communication), Douglas M. Fraiser (personal communication), Johnston (1975:41, 46), and Kerr (1965:22–23, 26, 30, 32, 1988:12, 16, 18–20, 23–24, 26, 32–33, 35–37, 46, 50, 56–58, 66–67, 78, 85, 98, 103, 121).

category, where the Actor is the highest on the thematic hierarchy (for example, as argued, in Grimshaw 1990:7–30), this ordering type utilizes the constraint dubbed ACTOR-1ST.

The same reasoning extends to the subject being the most prominent syntactic relation (in describing Mantauran Rukai, the Kalamianic languages, and Seediq). See, for instance, the notion in Role and Reference Grammar of "privileged syntactic argument" (Van Valin 2005:94 including fn. 5 citing earlier work in the same framework). Hence, if semantic roles are the crucial factor, then it is the Actor pronoun that must be cluster-initial, and if syntactic relations, then it is the subject. Similarly, as the remainder of this subsection will also show, if grammatical person is the relevant factor, then constraints require SAP clitic pronouns to precede third-person ones (and, in some languages, even first- to precede second-person ones), based on the notion – perhaps most famously in Silverstein's Agency Hierarchy (1976) – that certain grammatical persons are more prominent than others.

Before proceeding, it is worth differentiating the two constraints introduced so far from those presented below. Because in a verbal clause there is exactly one Actor and at most one subject, if a language uses roles or relations to order its clitic pronouns relative to each other within the cluster, then such ordering is categorical. Namely, if either SUBJECT-1ST or ACTOR-1ST is the dominant constraint in a language, then *every* overt co-occurrence of clitic pronouns shows the same categorical order. In the types yet to be discussed (based on grammatical person and prosodic weight, where the two pronouns might, e.g., both encode third person or both be monosyllabic), there is not only a dominant constraint but also one or more other constraints that emerge as tie-breakers if in a given cluster the dominant constraint fails to rule out all the unattested orders, thus not selecting any optimum.

In Squliq Atayal and Mayrinax Ts'uli' Atayal a third-person clitic pronoun must be cluster-final. On this issue there is consensus (Huang 1989:123–124, 1993:18–19, 1995a:127, 1995b:32, 34–35; Huang et al. 1999:187; Liao 2004:287, 2005:54–55; Rau 1992:146–147). However, if *neither* clitic pronoun encodes third person (in other words, if both are SAPs), then some other factor decides the ordering. (We follow Liao, arguing against the previous studies' claims, who proposes that prosodic weight is the tie-breaking factor.) As in Seediq, the picture in these two Atayal dialects is clouded considerably by the absence of overt clitic 3.NOM pronouns. As such, there can be no cluster of two third-person clitic pronouns. It is therefore impossible to conclude that grammatical person is the relevant factor. There is, however, one other language where this complication is absent: Iranun (a Danao language spoken in Sabah, Malaysia). The heretofore-unpublished data in table 4 were collected in the field (from two speakers, hailing from Rampayan and Kaguraan, both in Kota Belud, in early 2007 by the current study's second author). As in table 2 on Seediq above, GEN/Actor clitic pronouns are underlined, NOM/Undergoer clitic pronouns appear in bold type, and any yellow shading indicates (potential) syncretism of certain forms: most of the [+you] pronouns. (Unlike in Seediq, there are no potential co-occurrences among such case-syncretic forms because both pronouns in such combinations would encode [+you]. In some of these combinations both pronouns even encode [+me] as well. That is, all yellow-shaded cells also show at least one set of diagonal lines.) In addition, Iranun attests a minimal-augmented, formally [±aug], system of grammatical number, as opposed to the more common [±pl] system found in Seediq.[15] Yet another difference in Iranun is so-called *disformation* (a term originally coined in Peng and Billings 2008:184). In Iranun, if the Undergoer pronoun is not cluster-initial, then the corresponding form from the free-NOM paradigm, shown in italics in table 4, is used. These details aside, table 4 shows clearly that grammatical person is relevant only if there is one

[15] Namely, in Iranun /=ta/ is used to denote only the speaker and one addressee, whereas /=tanu/ refers to the speaker and (i) more than one addressee, (ii) one addressee plus one or more individuals not present, or both. See Corbett (2000:166–169) for discussion. This is why table 4 has eight sets of rows/columns.

SAP pronoun and one third-person pronoun. As the orange-shaded cells in table 4 show, if only one of the clitic pronouns encodes third person, then that pronoun follows the SAP pronoun.[16] As the blue-shaded cells in the same table show, if *both* clitic pronouns in the cluster are SAPs, then the ACTOR-1ST constraint (already defined above in this subsection) emerges to decide the ordering. Similarly, as the three green-shaded cells show, if *neither* pronoun is an SAP, then ACTOR-1ST also emerges. In Iranun [–me, –you, +aug] clitic pronouns do not co-occur, hence the empty cell in the bottom-right corner of this table.[17]

Table 4: Combinations of clitic pronouns in Iranun of Sabah

NOM \ GEN		+me –you –aug =aken	+me –you +aug =ami	+me +you –aug =ta	+me +you +aug =tanu	–me +you –aug =(ŋ)ka	–me +you +aug =yu	–me –you –aug =yan	–me –you +aug =iran
+me –you –aug	**=aku** ~ *saken*					=(ŋ)ka =*saken*	=yu =*saken*	**=aku** =yan	**=aku** =iran
+me –you +aug	**=kami** ~ *sekami*					=(ŋ)ka =*sekami*	=yu =*sekami*	**=kami** =yan	**=kami** =iran
+me +you –aug	**=ta**							**=ta** =yan	**=ta** =iran
+me +you +aug	**=tanu**							**=tanu** =yan	**=tanu** =iran
–me +you –aug	**=ka** ~ *seka*	=aken =*seka*	=ami =*seka*					**=ka** =yan	**=ka** =iran
–me +you +aug	**=kanu** ~ *sekanu*	=aken =*sekanu*	=ami =*sekanu*					**=kanu** =yan	**=kanu** =iran
–me –you –aug	*sekayan*	=aken =*sekayan*	=ami =*sekayan*	=ta =*sekayan*	=tanu =*sekayan*	=(ŋ)ka =*sekayan*	=yu =*sekayan*	=yan =*sekayan*	=iran =*sekayan*
–me –you +aug	*sekiran*	=aken =*sekiran*	=ami =*sekiran*	=ta =*sekiran*	=tanu =*sekiran*	=(ŋ)ka =*sekiran*	=yu =*sekiran*	=yan =*sekiran*	

Iranun is thus characterized by a grammar in which LOCAL-1ST is paramount, dominating the ACTOR-1ST constraint (where *local*, occasionally used in the literature, is synonymous with *SAP*). Because Iranun attests overt third-person clitic pronouns in *both* morphological cases, it is possible to see LOCAL-1ST most clearly. We are aware of dominant (i.e., exceptionless) LOCAL-1ST only in Squliq Atayal, Mayrinax Ts'uli' Atayal, and Iranun.

---

16 The disformed free-NOM pronouns in table 4 exhibit *optional* morphosyntactic positioning (similarly to the situation in Agutaynen, discussed in fn. 13 above). For example, if the clause is negated (where, as in Seediq, NEG is a free form that precedes the verb), then the GEN clitic pronoun (regardless of its position within the cluster) or NOM clitic form (in cluster-initial position) in such clusters must both follow NEG and precede the verb. However, the disformed, free-NOM pronoun (shown in italics in the eight blue-shaded cells and both of the last two rows of table 4) can appear either (i) immediately after the clitic-GEN pronoun, before the verb, or (ii) right after the verb. Of relevance to the current study is only option (i), with both pronouns preceding the verb and thus both morphosyntactically positioned as clitics.

17 We adopt McKaughan's transcription, where <e, y, r> are a central mid/high vowel, a palatal glide, and an alveolar flap (respectively); as with Seediq, italicized <*a*> in Iranun is the language's only low vowel /a/ (1999:82 n. 6). The 2SG.GEN clitic pronoun in Iranun begins with [ŋ] only after a vowel (McKaughan 1999:55, 82 n. 11). In addition, *=ku* is acceptable in each environment in place of *=aken* (even if the [+me, –you, –aug] column in table 4 shows only *=aken*). McKaughan does not list the *=ku* variant and reports different forms: *=yu* ~ *=iu* '=2PL.GEN', *=(y)ian* '=3SG.GEN', *sekitanu* 'INCL1PL.NOM' (1999:55).

There is a considerably larger number of languages that distinguish between first- and second-person pronouns within a clitic cluster. Those that do also distinguish between SAP and third-person pronouns. The languages of this type that we know about are Plngawan Ts'uli' Atayal (Chang 2012); Maranao, a Danao language of the southern Philippines (Kaufman 2010); and all known Manobo languages of the southern Philippines (Brichoux and Brichoux 1977; DuBois 1976; Hung and Billings 2009/2010; Liao 2004:459–468; Peng and Billings 2008; Weaver and Weaver 1964) – except, that is, for Cotabato Manobo, already discussed (in fn. 14). Within the pronominal-clitic cluster in each of these languages (i) any first-person form must precede any second-person form and, as in the languages discussed in the preceding paragraph, (ii) any SAP clitic pronoun precedes any third-person form. In these languages, ME-1ST is the main constraint; it dominates YOU-1ST. Chang (2012) utilizes these two constraints in her analysis of Plngawan. (In all such languages except Plngawan there are overt clusters involving third-person pronouns encoding both morphological cases, allowing us to see that grammatical person – rather than semantic roles, syntactic relations, or morphological cases – is the crucial factor.) Note that some of the languages of this type are very closely related to those mentioned in the preceding paragraph, where LOCAL-1ST is the primary constraint: Plngawan and Mayrinax (Ts'uli' Atayal subdialects) as well as Maranao and Iranun (Danao languages). Li (2010:36–38) demonstrates that LOCAL-1ST is a combination (specifically, a conjunctive local tie, as defined by Müller 1999:6–7/2003:307–310) of ME-1ST and YOU-1ST.

The final type of cluster-internal ordering uses phonology rather than morphosemantic properties. In many of the Central Philippine languages the main requirement is that a monosyllabic clitic pronoun be cluster-initial (Billings 2005; Bloomfield 1917:143, 181 [299, 337]; Kaufmann 1916:37; Lee 2008; Lee and Billings 2008; McFarland 2001; Schachter 1973:217). The relevant constraint is LIGHT-1ST, which is also emergent (under LOCAL-1ST) in Squliq Atayal (Liao 2004:285–296, 2005) and Mayrinax Ts'uli' Atayal. In the theoretical framework adopted here (e.g., Hung and Billings 2009/2010 and Kaufman 2010), *light* vs. *heavy* is shorthand for 'monosyllabic' as opposed to 'prosodically larger' (whereas *short* vs. *long* is a morphologically arbitrary, paradigm-assigning designation).

This concludes the discussion of the typology of cluster-internal ordering of pronouns in northern Austronesian (mainly the languages of Taiwan and the Philippines). In Seediq neither SUBJECT-1ST nor LOCAL-1ST is violated. However, a grammar using the latter constraint cannot generate all the data. Only if SUBJECT-1ST is used are all overt co-occurrences of pronominal clitics adequately predicted. We come back to this issue below (in §3.3), where we also discuss the three aforementioned dialects of Atayal. Before that, we present a formal analysis using the optimality-theoretic constraints discussed so far.

3.2 Theoretical analysis

This subsection uses the principles of Optimality Theory (originally from Prince and Smolensky 1993/2004) to define the formal grammar of cluster-internal ordering in Seediq. To simplify the exposition considerably, we show ME-1ST and YOU-1ST conjoined into the single LOCAL-1ST constraint. Because extremely few overt co-occurrences of clitics are found (especially due to the empty cells along the bottom row of table 2 above), there are also relatively few constraint rankings that can be posited about this language.

As we mention above in section 2, the variation between *=misu* '=1SG>2SG' and *=su=mu* '=2SG=1SG.GEN' (in cell K of table 2) is cross-linguistically quite rare. Though the use of *=misu* is "considered archaic" (Holmer 1996a:70), either of these combinations is used, even by the same speaker. We therefore have a hint that the language is changing, with the overt cluster gradually taking over from the portmanteau as the way to express the combination of a 1SG Actor and a 2SG Undergoer. To account for this variation, we assume

that there are two co-existing lexicons with most speakers today: one *with* the portmanteau =*misu* '=1SG>2SG' and the other *without* it. (Both lexicons currently include the other two portmanteaux, =*saku* '=2SG>1SG' and =*maku* '=1SG>2PL', which are both obligatory.) In these two co-existing lexicons, the optimum realizations are =*misu* '=1SG>2SG' and =*su*=*mu* '=2SG=1SG.GEN', respectively. Positing co-existing lexicons (or grammars) is one way to account for synchronic optionality. (Below in this subsection we demonstrate that there is still just one grammar.) In most of this subsection we assume one lexicon (without =*misu*) then briefly account for the other lexicon (in which this portmanteau is attested).

Using just the overt co-occurrences in table 2 above (in cells B–D, G–J, K1, M–N, and Q–R) for the time being, the only conclusive ranking of the constraints in the preceding subsection that can be determined is that SUBJECT-1ST dominates ACTOR-1ST. Comparing attested =*su*=*mu* '=2SG=1SG.GEN' with the opposite, unacceptable cluster-internal order demonstrates this ranking.[18] Both orders satisfy LOCAL-1ST and LIGHT-1ST equally, since each begins with an SAP pronoun that is monosyllabic (respectively). Attested =*su*=*mu* satisfies SUBJECT-1ST but violates ACTOR-1ST, whereas *=*mu*=*su* violates SUBJECT-1ST but satisfies ACTOR-1ST. This demonstrates that SUBJECT-1ST dominates ACTOR-1ST in Seediq:

Tableau 1: Seediq (lexicon that excludes =*misu*) [See cell K in table 2 and (10b) above.]

☞=*su*=*mu* '=2SG=1SG.GEN'	*			*
Input: Actor [+me, –you, –pl]; Undergoer [–me, +you, –pl].	LOCAL-1ST	LIGHT-1ST	SUBJECT-1ST	ACTOR-1ST
*=*mu*=*su* '=1SG.GEN=2SG'	*		* W	L

For more on the notations used in our tableaux, see Li (2010:32–33) or Yen (2012:98–99).

Only one other, far less revealing ranking statement can be determined using just the overt clusters. Two co-occurrences demonstrate this: =*nami*=*na* '=EXCL1PL=3SG.GEN' and =*namu*=*na* '=2PL=3SG.GEN' (in cells G and Q of table 2, respectively). A tableau for =*nami*=*na* (not shown here) would show exactly the same arrangements of asterisks, *W*s, and *L*s as tableau 2 does. Each of these attested orders fares *better* than its unattested opposite order with regard to both LOCAL-1ST and SUBJECT-1ST but fares *worse* with regard to both LIGHT-1ST and ACTOR-1ST. Since we know from tableau 1 that SUBJECT-1ST dominates ACTOR-1ST, the only new ranking statement (i.e., in addition to the one already demonstrated in the preceding tableau) that these two clusters show is that at least one of LOCAL-1ST and SUBJECT-1ST dominates LIGHT-1ST. Tableau 2 demonstrates this.

Tableau 2: Seediq (both lexicons) [See cell Q in table 2 above.]

☞=*namu*=*na* '=2PL=3SG.GEN'	*	*		*
Input: Actor [–me, –you, –pl]; Undergoer [–me, +you, +pl].	LOCAL-1ST	LIGHT-1ST	SUBJECT-1ST	ACTOR-1ST
*=*na*=*namu* '=3SG.GEN=2PL'	** W	L	* W	L

Next, we consider the combinations involving portmanteaux. A disyllabic portmanteau, being a single form, clearly violates LIGHT-1ST. If we assume (along the lines of Chang 2012:184–185), that a portmanteau vacuously satisfies any of LOCAL-1ST, SUBJECT-1ST, and ACTOR-1ST, then it is possible to improve just a little on the ranking from tableau 2.

[18] We adopt the definition of LOCAL-1ST from Li (2010:38): "A single violation is incurred if the cluster-initial pronoun is either [+me, –you] or [–me, +you]. Two violations are incurred if the cluster-initial pronoun is [–me, –you]. No violation is incurred if the cluster-initial pronoun is [+me, +you]."

Comparing =*saku* '=2SG>1SG' (found in both co-existing lexicons) to *=*ku*=*su* '=1SG.NOM=2SG' (i.e., one of the unacceptable overt clusters), both combinations satisfy SUBJECT-1ST. However, whereas the attested portmanteau fares *worse* than *=*ku*=*su* with regard to LIGHT-1ST, =*saku* fares *better* with regard to both LOCAL-1ST and ACTOR-1ST. This demonstrates that at least one of LOCAL-1ST and ACTOR-1ST dominates LIGHT-1ST. (Because we already know that SUBJECT-1ST dominates ACTOR-1ST, the preceding sentence tells us slightly more than the ranking from the preceding paragraph does. Namely, we now know that LIGHT-1ST is dominated by at least one of LOCAL-1ST and not only SUBJECT-1ST but also ACTOR-1ST: a constraint clearly dominated by SUBJECT-1ST.)

Tableau 3: Seediq (both lexicons) [See cell A in table 2 and (11a) above.]

☞=*saku* '=2SG>1SG'		*		
Input: Actor [–me, +you, –pl]; Undergoer [+me, –you, –pl].	LOCAL-1ST	LIGHT-1ST	SUBJECT-1ST	ACTOR-1ST
*=*ku*=*su* '=1SG.NOM=2SG'	* W	L		* W
*=*su*=*ku* '=2SG=1SG.NOM'	* W	L	* W	

Comparing =*saku* '=2SG>1SG' to the other unacceptable order, *=*su*=*ku* '=2SG =1SG.NOM', shown on the bottom row of tableau 3, corroborates the ranking demonstrated in tableau 2.

So far in this subsection we have discussed the rankings for the lexicon of Seediq that excludes the portmanteau =*misu* '=1SG>2SG'. Here we briefly assess the other co-existing lexicon that includes this portmanteau. As it so happens, the comparison of =*misu* '=1SG>2SG' (the assumed optimum in this lexicon) with =*su*=*mu* '=2SG=1SG.GEN' (one of the assumed sub-optima) results in exactly the same rankings as the comparison of =*saku* with *=*ku*=*su* in the preceding lexicon, in tableau 3: at least one of LOCAL-1ST and ACTOR-1ST dominates LIGHT-1ST. Thus, though the lexicons differ, the grammars do not.

Tableau 4: Seediq (lexicon that includes =*misu*) [See cell K in table 2 and (11b) above.]

☞=*misu* '=1SG>2SG'		*		
Input: Actor [+me, –you, –pl]; Undergoer [–me, +you, –pl].	LOCAL-1ST	LIGHT-1ST	SUBJECT-1ST	ACTOR-1ST
=*su*=*mu* '=2SG=1SG.GEN'	* W	L		* W
*=*mu*=*su* '=1SG.GEN=2SG'	* W	L	* W	

The last row of tableau 4, comparing =*misu* with *=*mu*=*su*, again reveals nothing new. (In fact, the entirety of tableaux 3 and 4 reveal identical ranking information to each other.)

Moving to the third and final portmanteau, though the last rows of tableaux 3 through 5 show identical rankings, the penultimate row of tableau 5 (comparing =*maku* '=1SG>2SG' with *=*namu*=*mu* '=2PL=1SG.GEN') is so far distinct. This is because =*namu* '=2PL' (unlike either =*ku* '=1SG.NOM' or =*su* '=2SG' in tableaux 3 and 4, respectively) is disyllabic.

Tableau 5: Seediq (both lexicons) [See cell O in table 2 and (11c) above.]

☞=*maku* '=1SG>2PL'		*		
Input: Actor [+me, –you, –pl]; Undergoer [–me, +you, +pl].	LOCAL-1ST	LIGHT-1ST	SUBJECT-1ST	ACTOR-1ST
*=*namu*=*mu* '=2PL=1SG.GEN'	* W	*		* W
*=*mu*=*namu* '=1SG.GEN=2PL'	* W	L	* W	

The disyllabicity of =*namu* '=2PL' entails that both the optimum =*maku* and unattested cluster *=*namu*=*mu* violate LIGHT-1ST. A row with only *W*s and no *L*s (standing for 'optimum wins' and 'optimum loses', respectively) shows no ranking information at all. The optimum is said to *harmonically bound* the other form (meaning that under any ranking of the various constraints the optimum wins). Tableau 5 thus tells us nothing new.

To summarize briefly, very few rankings can be gleaned from the modern language:

(12) a. SUBJECT-1ST dominates ACTOR-1ST.

b. At least one of LOCAL-1ST and ACTOR-1ST dominates LIGHT-1ST.

The next subsection refines these ranking statements using data from related languages.

3.3 Prehistory

Fortunately for the current purposes, relatively complete pronoun-combination data are available for three closely related clitic-ordering systems: in the dialects of Atayal proper.[19] This subsection reconstructs the lexicon and grammar of Proto Atayalic by means of the Comparative Method, and even Pre-Proto Atayalic using Internal Reconstruction, in order to add insight to the somewhat sketchy modern constraint hierarchy in (12a–b) above.[20]

To begin, SUBJECT-1ST dominating ACTOR-1ST (from tableau 1) is not reconstructable for that earlier stage of prehistory because all three Atayal dialects require /=misuʔ/ '=1SG>2SG' (rather than an overt clitic cluster) for this pronominal combination. However, tableau 4's rankings above can be extended to Proto Atayalic. First, *=misuʔ '=1SG>2SG' can safely be reconstructed for Proto Atayalic; cf. cell K of table 2, (11b), and tableau 4 above. All four clitic-ordering systems attest this form (differing trivially only in the final glottal stop in only the three Atayal dialects). In addition, corresponding to =*su* '=2SG' and =*mu* '=1SG.GEN' in Seediq, all three modern Atayal dialects also attest /=suʔ/ '=2SG' and /=mu/ '=1SG.GEN'. (Mayrinax has variants /=siʔ/ and /=miʔ/, respectively; Squliq has a variant /=makuʔ/ '=1SG.GEN', discussed further below. These variants are innovations in only one dialect each and thus are irrelevant.) Tableau 4 is therefore modified as in the following tableau. Comparing the portmanteau optimum *=misuʔ '=1SG>2SG' to the sub-optimal cluster *=suʔ*=mu '=2SG=1SG.GEN' shows the same ranking as in (12b) above.

Tableau 6: Proto Atayalic [Compare with Seediq tableau 4 above.]

☞*=misuʔ '=1SG>2SG'		*		
Input: Actor [+me, –you, –pl]; Undergoer [–me, +you, –pl].	LOCAL-1ST	LIGHT-1ST	SUBJECT-1ST	ACTOR-1ST
=suʔ=mu '=2SG=1SG.GEN'	* W	L		* W
=mu=suʔ '=1SG.GEN=2SG'	* W	L	* W	

[19] Pronoun-combination tables for Squliq, Plngawan Ts'uli', and Mayrinax Ts'uli' come from, respectively, Li (2010:37), Chang (2012:189), and Lee and Billings (2012). Chang's table is complete, but Li's lacks Squliq data corresponding to cells E, F, L, O, and P in our table 2 and Lee and Billings are missing Mayrinax data corresponding to cells D, H–J, M and Q of table 2. Any reconstructions are based on cells where there are data from all four dialects: corresponding to cells A–C, G, K, N, and R. For example, tableaux 2 and 5 above, based on cells Q and O of table 2, respectively, cannot be used for Proto Atayalic. Furthermore, every *form* in our tableaux below is reconstructed, as the discussion shows.

[20] All modern data below from the dialects of Atayal proper are standardized to the International Phonetic Alphabet. We also use the same transcription (preceded by asterisks) for any prehistoric reconstructions.

The bottom row of tableau 6 shows the same ranking as the bottom rows in tableaux 3 through 5. Thus, these two ranking statements can be reconstructed for Proto Atayalic so far: at least one of LOCAL-1ST and ACTOR-1ST dominates LIGHT-1ST, identical to (12b) for Seediq, and at least one of LOCAL-1ST and SUBJECT-1ST dominates LIGHT-1ST. Yet another ranking statement can be made for Proto Atayalic, using an overt clitic cluster this time.[21]

Because in all four of the known modern clitic-ordering systems in Atayalic an SAP clitic pronoun must precede a 3.GEN pronoun in the cluster, we can reconstruct Proto Atayalic as having utilized LOCAL-1ST as a superordinate constraint. Perhaps the most probative combination corresponds to cell R of table 2 above, where Seediq attests *=namu* '=2PL' followed by *=daha* '=3PL.GEN'. For this combination, Squliq, Plngawan, and Mayrinax attest /=simu=nhaʔ/ '=2PL.NOM=3PL.GEN', /=mamu=nahaʔ/ '=2PL=3PL.GEN' (where /=mamu/ is syncretic with NOM), and /=ʦimu=nhaʔ/ '=2PL.NOM=3PL.GEN', respectively.[22] As these combinations show, like Seediq *=daha* '=2PL', the 2PL(.NOM) form in most of the three Atayal dialects is disyllabic (with the final two sounds as [mu]). We can therefore reconstruct a disyllabic form for Proto Atayalic, as tableau 7 demonstrates.[23]

---

[21] The portmanteaux discussed in this subsection are not found outside of Atayalic. However, certain pronouns in Proto Austronesian reconstructed in Ross (2006:532–534) resemble some of these three portmanteaux outside of Atayalic: (i) *(=)m-iSu, the 2SG member of his GEN2 set, (ii) =[S]aku, one of the 1SG forms in his NOM2 set, and (iii) *(=)m-aku, the 1SG form in his GEN2 set. In (i) and (ii) *S is a coronal obstruent of inexact quality (with almost exclusively fricative reflexes, and which Ross 2012:1275 reconstructs as post-alveolar), reflected as /s/ throughout Atayalic. Incidentally, Ross (2006) does not mention portmanteau pronouns at all. Of these, (i) and (ii) appear to be chance resemblances to the three portmanteaux in Atayalic. To begin, Ross lists reflexes of (i) only in Thao, Siraya, K[anak]anavu, Proto Rukai[c], and Proto Amis (2006:534). Next, though =[S]aku in (ii) resembles "*=[ca]ku" (which Ross posits for Proto Atayalic, where *c is a coronal affricate) '=1SG.NOM', Ross (personal communication, 19 May 2013) has clarified that the *c in his reconstruction of Proto Atayalic is not intended to reflect the *S in his reconstruction of Proto Austronesian and now agrees, based on reflexes in Takibakha Bunun (that were not available to him for Ross 2006), that the *S in his reconstruction of Proto Austronesian in (ii) should now be changed to *s (with possible concomitant rearrangement of his NOM1 and NOM2 paradigms). In all Atayalic languages *s (like *S) is reflected as /s/. Only in (iii) does Ross list reflexes outside Atayalic that might be cognate to *=maku* '=1SG>2PL' in Seediq. For (iii) Ross lists *=m-aku (Proto Atayalic), *=mau* (Siraya), *=maku* (K[anak]anavu), and *m-aku (Proto Amis). In none of those non-Atayalic (proto)languages does the form encode a portmanteau function (2006:534). As for his "*=m-aku" in Proto Atayalic, the modern reflexes are /=mu/ in all of Truku (Seediq), Wulai (Squliq), Plngawan, and Mayrinax; only Wulai (Squliq) is listed also with "*=[ma]kuʔ*" (2006:549). Below we offer an alternative analysis of the 1SG.GEN forms (i.e., of /=makuʔ/ as a variant of /=mu/) in Squliq Atayal.

[22] The Squliq datum is from Rau (1992:147, cited in Li 2010:37); the Plngawan, from Chang (2012:179); and the Mayrinax, from Huang (1995a:126, 1995b:31, both cited by Lee and Billings 2012).

[23] Two complications cloud the assessment of the LIGHT-1ST constraint in tableau 7 somewhat. To begin, though the 2PL.NOM form might be reconstructed as *=mu, a single syllable, in Proto Austronesian (as, e.g., in Ross 2006:532), in Proto Atayalic we reconstruct a disyllabic from: *=iamu '=2PL.NOM', also segmentally identical in form to the case-neutral 2PL free form. Compare these with "*cimu" and "*yamu" both glossed as 'you (pl.)', with neither specifying its case or whether it is a free or bound form, in Li (1981:297/2004:690) as well as "(*=cimu)" '=2PL.NOM' and "(*i-amu)" case-neutral free '2PL' in Ross (2006:533); cf. slightly divergent reconstructions in his appendix B.3 (2006:549), where *c is a coronal affricate in both authors' transcription, and Ross places parentheses around reconstructions to show uncertainty due to lack of supporting data (2006:524). The optimum order of clitics in tableau 7 therefore violates LIGHT-1ST. Next, we reconstruct *=nahaʔ '=3PL.GEN' for Proto Atayalic. Noncrucially, we reconstruct initial *n (rather than *d or *l) and final *ʔ here, neither of which affects the number of syllables. Of more consequence, we reconstruct a disyllabic form. Ross also reconstructs only disyllabic forms here: "(*=laha)" and "*(=na-haʔ)" (2006:537, 549). However, even if this protoform were monosyllabic, it would not affect our ranking statement in (13c). In tableau 7, under *LIGHT-1ST*, there would be an *L* but no asterisk, further entailing the additional ranking statement in (13c) below, arrived at using independent evidence: the last row of tableau 6 above. The upshot is that we can reconstruct the following ranking statements for Proto Atayalic: (13a-b) from tableau 6 alone and (13c) from tableau 7.

Tableau 7: Proto Atayalic [Corresponds to cell R in table 2 above.]

☞*=iamu*=nahaʔ '=2PL=3PL.GEN'	*	*		*
Input: Actor [–me, –you, +pl]; Undergoer [–me, +you, +pl].	LOCAL-1ST	LIGHT-1ST	SUBJECT-1ST	ACTOR-1ST
=nahaʔ=iamu '=3PL.GEN=2PL'	** W	*	* W	L

Tableau 7 tells us that at least one of LOCAL-1ST and SUBJECT-1ST dominates ACTOR-1ST. (Recall that we do not reconstruct SUBJECT-1ST » ACTOR-1ST in Proto Atayalic as such.)

The following are thus our full list of reconstructed rankings for Proto Atayalic itself.

(13) a. At least one of LOCAL-1ST and ACTOR-1ST dominates LIGHT-1ST [= (12b)].

b. At least one of LOCAL-1ST and SUBJECT-1ST dominates LIGHT-1ST.

c. At least one of LOCAL-1ST and SUBJECT-1ST dominates ACTOR-1ST.

Incidentally, (13a–c) are consistent with our modern rankings for Seediq, in (12a–b) above. As it were, there is no evidence of the *grammar* (i.e., the constraint hierarchy) changing. Instead, only the *forms* (and their combinations) have changed, allowing us to see perhaps the same grammar from different perspectives. The remainder of this subsection is devoted to reconstructing the portmanteaux even further back in time, to Pre-Proto Austronesian.

We begin with *=misu* '=1SG>2SG' in Seediq and its cognate: /=misuʔ/ in all three Atayal dialects. Egerod analyzes /=misuʔ/ in Squliq Atayal as "combining mu and isuʔ" (1980:1.377/1999:160). We extend Egerod's idea, reconstructing the cluster *=mu*=isuʔ '=1SG.GEN=2SG(.NOM)'. Corresponding to *=mu* '=1SG.GEN' and *isu* '2SG' in Seediq are /=mu/ '=1SG.GEN' in all four dialects of Atayal proper and /isuʔ/ '2SG' in both Squliq and Plngawan but /ʔisuʔ/ '2SG' in Mayrinax.[24] As has already been mentioned, Squliq and Mayrinax also attest the variants /=makuʔ/ and /=miʔ/ '=1SG.GEN', respectively. Each of these is an innovation in only one dialect and thus irrelevant to any comparative reconstruction. Because the portmanteau *=misuʔ '=1SG>2SG' itself can be reconstructed to Proto Atayalic, the fusion of *=mu '=1SG.GEN' plus *=isuʔ '=2SG(.NOM)' into *=misuʔ must have occurred earlier, in Pre-Proto Atayalic. (Here we've initially done comparative reconstruction of Proto Atayalic followed by internal reconstruction of the fusion in even more distant prehistory.) In Proto Austronesian the GEN pronouns were monosyllabic but the NOM forms were mostly disyllabic (Dahl 1973, cited in part by Blust 1977:1–2, 10). We further assume that in Proto Austronesian the NOM pronouns were free but the GEN forms were both phonologically enclitic and morphosyntactically positioned as clitics (e.g., before the lexical verb in the presence of NEG). Later in Pre-Proto Atayalic, we propose, the free-form NOM pronouns could also be positioned as clitics but were not phonologically deficient as such. Therefore, the reconstructed NOM pronouns discussed in the remainder of this subsection are morphosyntactic clitics (but not phonologically enclitic as such).

We reconstruct *=mu*=isuʔ '=1SG.GEN=2SG(.NOM)' as optimal in Pre-Proto Atayalic and compare it to the opposite (and sub-optimal) order in tableau 8, on the next page. Note that (i) both orders of clitics violate LOCAL-1ST equally, with one asterisk each; (ii) only the optimum order *=mu*=isuʔ violates SUBJECT-1ST; and (iii) only the sub-optimal order *=isuʔ*=mu violates the remaining two constraints (i.e., LIGHT-1ST and ACTOR-1ST). Thus, in Pre-Proto Atayalic at least one of LIGHT-1ST and ACTOR-1ST dominates SUBJECT-1ST:

[24] However, cf. Li (1981:250, 262, 269, 297/2004:639–640, 651–652, 658, 690), who lists consensus *ʔisuʔ* in all of Squliq, four Ts'uli' dialects (including Mayrinax and Plngawan), and four Seediq locations.

Tableau 8: Pre-Proto Atayalic [Compare with tableaux 4 and 6 above.]

☞*=mu*=isuʔ '=1SG.GEN=2SG(.NOM)'	*		*	
Input: Actor [+me, –you, –pl]; Undergoer [–me, +you, –pl].	LOCAL-1ST	LIGHT-1ST	SUBJECT-1ST	ACTOR-1ST
=isuʔ=mu '=2SG(.NOM)=1SG.GEN'	*	* W	L	* W

Incidentally, this lone ranking statement from tableau 8 remains consistent with both (12a–b), for Seediq, and (13a–c), for Proto Atayalic, above. (For reasons discussed below, tableau 8 is our only attempt at a ranking in Pre-Proto Atayalic.) Namely, the hierarchy LOCAL-1ST » LIGHT-1ST » SUBJECT-1ST » ACTOR-1ST is consistent with all three stages.[25] The constraint columns in all our tableaux are, for convenience, arranged in this order. Once again, what's going on might well be changes only in the forms (and their combinations) without any changes in the grammars (i.e., the hierarchy of constraints).

Similarly, we reconstruct *=sakuʔ '=2SG>1SG' to Proto Atayalic; see cell A of table 2, (11a), and tableau 3 above. Though this portmanteau is absent in Mayrinax, /=sakuʔ/ '=2SG>1SG' is found in the two other dialects of Atayal proper. An optimality-theoretic tableau comparing *=sakuʔ '=2SG>1SG' in Proto Atayalic to both orders of reconstructed overt clusters (i.e., analogous to tableau 6 above) is not listed here for two reasons. First, whereas *=suʔ '=2SG' (case-syncretic) can safely be reconstructed in Proto Atayalic, the other pronoun is considerably more problematic: *=ku* '=1SG.NOM' in Seediq corresponds to /=kuʔ/, but with the variant /=sakuʔ/ in Squliq and /=ʦu/ in both Ts'uli' subdialects, with the variant /=ʦiʔ/ in Mayrinax.[26] Ross (2006:533, 549) proposes "*=[ca]ku" in Proto Atayalic, apparently combining the disyllabicity of the /=sakuʔ/ '=1SG.NOM' variant in Squliq with the initial affricate of the corresponding Ts'uli' forms.[27] In Squliq /=sakuʔ/

[25] As in Seediq, the three dialects of Atayal proper both (i) attest only inaudible 3.NOM clitic pronouns and (ii) require overt 3.GEN clitic pronouns to be cluster-final. Regarding the relative ordering of two SAP clitic pronouns, we concur with Liao's analysis of Squliq (2004:285–296, 2005) and similar, apparently independent suggestions about Mayrinax (Billings and Kaufman 2004:17; Li 1995:40/2004:403 citing "Mei (p[ersonal].c[ommunication].)"; Liao 2005:60) that prosodic weight is the relevant factor, where a monosyllabic pronoun must be cluster-initial (contra exclusively person-based approaches for Squliq in Huang 1989:124–126, 1993:18–20, 1995b:34–35 and Rau 1992:146–147 as well as a considerably more complex proposal for Mayrinax, based on all of person, number, and semantic roles – in each of Huang 1995a:126–129, 1995b:30–36, 2000:72–74). Only in Plngawan does grammatical person account for the internal ordering of *all* pronominal-clitic clusters, where first precedes second person and any SAP must be cluster-initial (Chang 2012). Given that LIGHT-1ST is reported in the literature on both the Squliq and the Ts'uli' branches of Atayal and that ACTOR-1ST does no work in any of the modern clitic-ordering systems in Atayalic, we can reliably reconstruct LOCAL-1ST » LIGHT-1ST » SUBJECT-1ST » ACTOR-1ST.

[26] In Plngawan /=ʦiniʔ/ '=3SG>1SG' (Chang 2012:180, 189) shows a vestige of a Proto Ts'uli' *=ʦiʔ '=1SG.NOM'. Li (1981:286/2004:677) lists a correspondence set including Plngawan, Mayrinax, and two other Ts'uli' subdialects. In addition to /=misuʔ/ '=1SG>2SG' and /=sakuʔ/ '=2SG>1SG', Plngawan attests five portmanteaux for each combination involving a 3SG Actor (Chang 2012:179–180, 185, 189). Shared innovations in support of a Ts'uli' subgroup include such *ʦ-initial 1SG.NOM forms (Li 1980:351 fn. 3/2004:231 fn. 3, 1985a:712/2004:847, both citing unpublished work by S. Tsuchida) and variants /=siʔ/ and /=tiʔ/ of /=suʔ/ '=2SG' and /=taʔ/ '=INCL1PL' (Li 1985a:713/2004:848), both being case-syncretic.

[27] The literature mainly uses <c> to transcribe a coronal affricate in Atayal dialects. Some works specify no exact place of articulation (Huang 1995a:107; Li 1995:30 fn. 2/2004:393 fn. 2, 1995:41 fn. 6/2004:404 fn. 6, 1998:51/2004:714; Ross 2006:545, 549–550). Huang (1995b:17) describes this affricate as *alveopalatal* in Mayrinax; so does Li, but transcribes it as [tš] (1980:353/2004:232, 1981:238) or [ʧ] ([1981/]2004:628, 1985a:714/2004:849), suggesting *palato-alveolar* articulation. Discussing Squliq, Li (1980:363/2004:242) transcribes <c> as [ʦ], an *alveolar* affricate. Given this somewhat confused state of affairs, we follow the alveolar descriptions of this affricate in Squliq, Plngawan, and Mayrinax by Egerod (1980:1.iv/1999:vi), Chang (2012), and Lu (2005:6, 22), respectively, but allow for allophonic variation.

'=1SG.NOM' and the portmanteau /=sakuʔ/ '=2SG>1SG' are homophonous.[28] In this connection, we are fortunate to learn of /=ʦaku/, a variant of /=ʦu/ '=1SG.NOM' in additional Ts'uli' subdialects (all from Li 1998:51, 54–55, 76–77/2004:714, 718–720, 739–740): Mkgugut and Pyahaw, also possibly Mtlaŋan and Knŋyan.[29] Given the /=ʦaku/ variant of /=ʦu/ '=1SG.NOM' in those Ts'uli' subdialects, it seems that Proto Atayalic must have had both *=sakuʔ '=2SG>1SG', with an initial fricative, and *=ʦaku '=1SG(.NOM)', beginning with an affricate. Unfortunately for the current purposes, we know of no subdialect that reflects both of these phonologically distinct disyllabic forms. These separate protoforms would merge into /s/-initial homophones in Squliq (because Proto Atayalic *ʦ regularly de-affricated into /s/ in Squliq). Pending further field research, we do not wish to draw any conclusions based on these forms.[30] Our second reason not to list the tableau comparing *=sakuʔ '=2SG>1SG' to overt clusters in Proto Atayalic is that, as we have verified, any constraint-ranking statements (regardless of the number of syllables in the 1SG.NOM Proto Atayalic form) would be a strict subset of those from tableau 6 above.

Moving to the fusion of two pronouns into *=sakuʔ '=2SG>1SG' in Pre-Proto Atayalic, similarly to his analysis of /=misuʔ/ '=1SG>2SG' in Squliq quoted above, Egerod analyzes /=sakuʔ/ in Squliq as "combining suʔ and **akuʔ" (1980:2.590/1999:263), where his use of two asterisks apparently implies a historical reconstruction of a 1SG form that is not attested in Squliq today. We adopt and extend Egerod's idea, reconstructing *=suʔ*=iakuʔ '=2SG=1SG(.NOM)' in Pre-Proto Atayalic. A tableau (not shown here) comparing this optimum order to the opposite one demonstrates the same rankings as in tableau 8 above.[31]

Considerably more speculative is our reconstruction of *=makuʔ '=2PL>1SG' in Proto Atayalic (crucially not the modern gloss of =*maku* '=1SG>2PL' in Seediq); see cell O of table 2, (11c), and tableau 5 above. This form is not used as a portmanteau outside of Seediq. How such a pronominal combination came to switch its mapping of person/number features to semantic roles in the resulting portmanteau in Seediq is unclear. Still, the

---

[28] In Squliq the /=sakuʔ/ variant of /=ku/ '=1SG.NOM' combines with either /=mamu/ '=2PL.GEN' (with the only attested example, to our knowledge, found in Hsiao 2004:27) or the 3.GEN pronouns (summarized in Li 2010:37), whereas the portmanteau /=sakuʔ/ '=2SG>1SG' is by far the most common way to encode the combination of a 2SG Actor with a 1SG Undergoer (though /=suʔ=sakuʔ/ '=2SG=1SG.NOM' is also possible but "very rare" according to Huang 1989:124 [cited in in Liao 2004:286, 293, 2005:52, 57], 125 [cited in Holmer 1993:86], 1993:20, 1995b:35). This situation forces a synchronic analysis of Squliq with both of these homophonous forms stored separately: /=sakuʔ/ '=1SG.NOM' and /=sakuʔ/ '=2SG>1SG'.

[29] By contrast, in the Skikun, Mnawyan, and Ryuhiŋ subdialects of Ts'uli' – as in Plngawan and Mayrinax, not discussed in Li (1998/2004) – no disyllabic variant of /=ʦu/ is reported (1998:76–77/2004:739–740). An earlier correspondence set (Li 1981:286/2004:677) combines "s-akuʔ" in Squliq (our /=sakuʔ/ variant of /=kuʔ/ '=1SG.NOM'), a bound from, with "y-akuʔ" in Seediq (our *yaku* '1SG'), a free form. In the same correspondence set no Ts'uli' forms (in four subdialects, including Mayrinax and Plngawan) are listed.

[30] In Plngawan, the only known Ts'uli' subdialect with a reflex of *=sakuʔ '=2SG>1SG', the literature does not report whether a disyllabic variant exists of /=ʦu/ '=1SG.NOM'. (In a preliminary inquiry by phone on our behalf by Haowen Jiang on 5 June 2013, the disyllabic variant of /=ʦu/ was consistently rejected.) Conversely, in the Ts'uli' subdialects where a /=ʦaku/ variant of /=ʦu/ is reported, there is no mention of whether a 2SG>1SG portmanteau is attested. We have been unable to access speakers of those subdialects.

[31] As we already mention above, corresponding to =*su* '=2SG' in Seediq, /=suʔ/ '=2SG' is attested (without variant forms, and where this pronoun does not distinguish between NOM and GEN) in all three Atayal dialects. Therefore, *=suʔ '=2SG' can safely be reconstructed for (Pre-)Proto Atayalic. Ross reconstructs *=Su '=2SG.GEN1' in Proto Austronesian (2006:532, 534); cf. also Blust (1977:5). Next, we reconstruct *=iaku '=1SG(.NOM)' (rather than *=aku) for Pre-Proto Atayalic based mostly on Atayalic-external evidence, following Blust's *i-aku 'NOM-1SG' for Proto Austronesian (1977:7), where the gloss of *i as 'NOM' is deemed merely "possible" (Blust 1977:13 n. 6). Corresponding to *yaku* '1SG' in Seediq are rather divergent forms in the Atayal dialects: /kuziŋ/ (alternating with /kun/) in Squliq, /kuɹiŋ/ in Plngawan, and /kuiŋ/ in Mayrinax (cf. Li 1985a:709/2004:844). There was obviously a lexical innovation in Proto Atayal. Seediq *yaku* '1SG' is therefore the only Atayalic reflex of *i-aku in Proto Austronesian.

reconstructed forms in Pre-Proto Atayalic – i.e., *=mu '=**2PL**.GEN' followed by *=iakuʔ '=1SG(.NOM)' – match up with the resulting modern portmanteau in Seediq: *=maku*. Unlike our discussion of the other two reconstructed portmanteaux discussed so far in this subsection, we necessarily begin with the *earlier* period of prehistory: Pre-Proto Atayalic. We propose that *=makuʔ is a fusion of *=mu '=**2PL**.GEN', reconstructed for Proto Austronesian by Blust (1977:9–10) but not attested with a 2PL.GEN function in any modern Atayalic dialect known to us, followed by *=iaku '=1SG(.NOM)', already mentioned in the preceding paragraph. All known modern Atayalic '=2PL.GEN' forms end in [mu], an open syllable.[32] Corresponding to *=namu* '=2PL.GEN' in Seediq is /=mamu/ in all Atayal dialects, with the variant /=momu/ in one Squliq subdialect (Rau 1992:96 fn. 11; see also Rau 1992:126 fn. 8, cited in Liao 2004:286 fn. 10, 2005:52 fn. 10). In Squliq and Mayrinax /=mamu/ is unambiguously '=2PL.GEN'; in Plngawan, /=mamu/ '=2PL' is case-syncretic. Recall that *=mu has been reconstructed with the '=1SG.GEN' function for both Proto Atayalic and Pre-Proto Atayalic (in the discussion above surrounding tableaux 6 and 8, respectively). The main questions posed by our hypothesis in this paragraph is at what stages (i) *=mu stopped encoding '=**2PL**.GEN' and (ii) *=mu began encoding '=**1SG**.GEN'. We have already proposed above that (ii) took place as early as Pre-Proto Atayalic.[33] For our hypothesis about *=makuʔ '=2PL>1SG' to be correct, change (i) would have to have taken place no earlier than Pre-Proto Atayalic and possibly even after the fusion, in Proto Atayalic, of *=mu '=**2PL**.GEN' plus *=iaku '=1SG(.NOM)' into *=makuʔ '=2PL>1SG'. The change in function from '=2PL>1SG' to '=1SG>2PL' probably also took place after that fusion but still in Proto Atayalic.[34] The only support for our claim in the preceding sentence is the /=makuʔ/ variant of /=mu/ '=1SG.GEN', only in Squliq (already mentioned in this subsection). We know of no examples of /=makuʔ/ in Squliq that have a portmanteau interpretation.[35] Both Seediq *=maku* '=1SG>2PL' and the Squliq /=makuʔ/ '=1SG.GEN' variant share the encoding of a 1SG Actor. If these two forms are cognate, then it stands to reason that there was a single switch in mappings (i.e., from '=2PL>1SG' to '=1SG>2PL') prior to the breakup of Proto Atayalic rather than two separate shifts.[36] Then, at some point

---

[32] Li (1980:355, 357/2004:235–236) reports that in Squliq Atayal word-final open syllables are an exception found only in function words such as demonstratives, so-called particles, and pronouns. As his two examples of pronouns, it is not clear whether he intended /=sami/ '=EXCL1PL.NOM' and /=simu/ '=2PL.NOM' or the case-neutral free forms /sami/ 'EXCL1PL' and /simu/ '2PL'. (His observation applies to all four forms.) The only other personal-pronoun form that we would add to this list is /=mu/ '=2SG.GEN'.

[33] Given that neither two [+me] pronouns nor two [+you] pronouns can co-occur in a clause (as, e.g., the diagonal lines in tables 2 and 4 show), there would have been no ambiguity if one of the homophonous *=mu GEN pronouns were to co-occur with various SAP NOM pronouns. The problem would have occurred if both (i) the 2PL and 1SG functions were present in the lexicon concurrently and (ii) one or the other co-occurred with a noun or free third-person pronoun. Given that people usually make statements (at least if realis) using [+me] pronouns but ask questions using [–me, +you] pronouns, then both functions might have co-existed. Eventually, the 1SG encoding took over and the 2PL one disappeared.

[34] With both *=misuʔ '=1SG>2SG' and *=makuʔ (still encoding '=2PL>1SG') there might also have been chronologically intermediate, more transparent forms with the initial syllable [mu]. The morphosemantic changes in (i) and (ii) having been complete, language learners might have associated this [mu] syllable with a 1SG (rather than a 2PL) Actor. However, the concomitant switch of the Undergoer's encoding from 1SG to 2PL still requires explanation. Perhaps if *=misuʔ '=1SG>2SG' (i.e., with a [–me, +you, –pl] Undergoer) was already taken, then the only remaining available [–me, +you] Undergoer was [+pl].

[35] Examples of Squliq /=makuʔ/ '=1SG.GEN' in unambiguously clitic position are listed in Chen (1992:21, 23); Egerod (1980:1.359–360/1999:151), with no fusion etymology; Huang (1989:127, 1993:31–32, 59, 70–71, 95, 112, 127 [cited in Liao 2004:273], 128, 130–131 [cited in Liao 2004:273], 134); and Rau (1992:77, 99, 148, 164, 166, 203, 205, 212, 213, 215, 216, 218, 219, 220, 221, 234, 235, 237, 238, 241).

[36] Squliq /=mu/ '=1SG.GEN' is consistently reported as being an open syllable, whereas its variant /=makuʔ/ is just as unfailingly listed with the final glottal stop. This fits with the etymology of /=makuʔ/ not

later in prehistory, *=maku? lost its function as a portmanteau (but kept one of the person/number feature sets), either in Proto Atayal or even after Squliq and Ts'uli' broke apart, remaining only in Squliq. Additionally, Ts'uli' lost the form *=maku? altogether.[37]

Before concluding this subsection, it's worth pointing out that Ross (2012:1278) deems only two shared innovations reported in the literature to be probative in demonstrating that Atayalic is a subgroup (of his Nuclear Austronesian, a primary subgroup of Austronesian). The first is a conditioned, and therefore distinctive, sound change: of *ə to *u only in final syllables (citing Li 1981:275[/2004:665], also mentioned in Li 1995:40 fn. 5/2004:403 fn. 5). The other probative innovation is the replacement of Proto Austronesian *=ku with *=mu'=1SG.GEN' (citing Harvey 1982:89). Our proposed innovations above of as many as three portmanteaux in Proto Atayalic constitute crucial support for this Atayalic subgroup.

This subsection has done three things. To begin, in tableaux 6 and 7 we established the ranking statements in (13a–c) about Proto Atayalic. These are consistent with the ranking statements about Seediq in (12a–b) from the preceding subsection. We then explored an earlier period, Pre-Proto Atayalic; tableau 8 shows another ranking statement, continuing to be consistent with those of both Seediq in (12) and Proto Atayalic in (13). The upshot is that a single grammar (namely: LOCAL-1ST » LIGHT-1ST » SUBJECT-1ST » ACTOR-1ST), could have remained unchanged in prehistory as only the forms in the lexicon changed.

To summarize section 3 then, three additional perspectives have been added to the facts presented in section 2 above. To begin (§3.1), we have placed Seediq within the typology of clitic positioning in the Austronesian languages of the area. Very few languages require the subject pronoun to precede the other clitic pronoun. The other property also true of pronominal clitic clusters in Seediq is that they must each begin with an SAP pronoun: also quite rare typologically. Aside from a few Atayalic systems, only one other language, Iranun (of Sabah), belongs to this type. Next (§3.2), we have analyzed the facts using the framework of Optimality Theory. Relatively few conclusions can be drawn from the modern data alone. However, comparing Seediq to three other clitic-ordering systems in Atayalic along with internal reconstruction (§3.3), suggests a ranking with a superordinate constraint requiring pronominal clusters to be SAP-initial. We also reconstructed three pronominal portmanteaux (in Proto Atayalic) and their fusions (in Pre-Proto Atayalic).

4 Conclusion and directions for future research

This study has consolidated the facts about clitic personal pronouns Seediq, an Atayalic language of Taiwan. In addition to the preliminary details (§1) and synchronic pronoun-combinatorial facts (§2), we have discussed typological, theoretical, and historical factors that underlie the occurrence of both overt clitic clusters and portmanteaux (§3).

Though all the non-portmanteau ordering facts in Seediq can be captured using only a constraint requiring a cluster-initial pronoun encoding the grammatical subject (as the first author's previous works have argued), historical reconstruction of (Pre-)Proto Atayalic shows that two other constraints probably remain operative. All known Atayalic dialects

---

involving /=mu/ '=1SG.GEN' in Pre-Proto Atayalic. Rather, all but the initial /m/ of /=maku?/ derives from *=iaku? '=1SG(.NOM)' in Proto Atayalic. See also Egerod's "**aku?" (1980:2.590/1999:263).

[37] A tableau (not shown here) for Proto Atayalic comparing *=maku? '=2PL>1SG' as optimum to both a '=1SG.NOM=2PL' cluster and a '=2PL=1SG.NOM' cluster analogous to tableau 6 above would result in the same ranking statement as in Seediq tableau 5 above: at least one of LOCAL-1ST and SUBJECT-1ST dominates LIGHT-1ST (already established for Proto Atayalic in the bottom row of tableau 6 above). Yet another tableau (also not shown here), for Pre-Proto Atayalic, comparing *=mu*=iaku '=2PL.GEN=1SG(.NOM)' as optimum with its reverse order yields the same rankings as in tableau 8 above.

require the clitic cluster to begin with an SAP pronoun (as argued specifically for Seediq in H. Chang's publications), suggesting that this person-based constraint has been operative since the earliest prehistoric stage, during which another constraint requiring monosyllabic pronouns to be cluster-initial was next-highest in the hierarchy. (That is, if both pronouns in the clitic cluster are SAPs, then a monosyllabic form will precede a prosodically heavier SAP pronoun.) We have also shown that an optimality-theoretic grammar in which the constraints – based in turn on grammatical person, prosodic weight, and subjecthood – captures the facts for three historical stages: from Pre-Proto Atayalic through to Seediq.

Much work remains to be done, to be sure, especially on comparative Atayalic pronouns. Our reconstructions (all in a stage that we call Pre-Proto Atayalic here) of *=mu*=isuʔ '=1SG.GEN=2SG(.NOM)', *=suʔ*=iakuʔ '=2SG=1SG(.NOM)', and *=mu*=iakuʔ '=2PL.GEN=1SG(.NOM)' – leading eventually to =*misu* '=1SG>2SG', =*saku* '=2SG>1SG', and =*maku* '=1SG>2PL' in Seediq, respectively – need to be extended to all of the individual pronominal forms in the two clitic paradigms as well as the case-neutral free paradigm.

References

Aldridge, Edith, 2004, Ergativity and word order in Austronesian languages. PhD dissertation. Ithaca: Cornell University.

Asai, Erin, 1953, *The Sedik language of Formosa*. Kanazawa, Japan: Cercle Linguistique de Kanazawa, Kanazawa University.

Billings, Loren A., 2005, Ordering clitics and postverbal R-expressions in Tagalog: A unified analysis? In Andrew Carnie, Heidi Harley, and Sheila Ann Dooley, eds. *Verb first: On the syntax of verb-initial languages* (Linguistik Aktuell 73), 303–339. Amsterdam: Benjamins.

Billings, Loren, and Daniel Kaufman, 2004, Towards a typology of Austronesian pronominal clisis. *ZAS Papers in Linguistics* 34 [special issue: Paul Law, ed. Proceedings of AFLA 11]:15–29.

Bloomfield, Leonard, 1917, Tagalog texts and grammatical analysis. *University of Illinois Studies in Language and Literature* 3(2–4):1–408 [157–564].

Blust, Robert A., 1977, The Proto-Austronesian pronouns and Austronesian subgrouping: A preliminary report. *Working Papers* [Honolulu: Department of Linguistics, University of Hawai'i at Mānoa] 9(2):1–15.

— 1999, Subgrouping, circularity and extinction: Some issues in Austronesian comparative linguistics. In Elizabeth Zeitoun and Paul Jen-kuei Li, eds. *Selected papers from the eighth International Conference on Austronesian Linguistics* (Symposium Series of the Institute of Linguistics 1), 31–94. Taipei: Institute of Linguistics (Preparatory Office), Academia Sinica.

Brichoux, Robert, and Felicia Brichoux, 1977, A sketch of Ilianen Manobo inflection. *Studies in Philippine Linguistics* 1(1):166–171.

Chang, Henry Y. [Yung-Li Chang], 1997, Voice, case and agreement in Seediq and Kavalan. PhD dissertation. Hsinchu: National Tsing Hua University.

— [Yung-Li Chang], 1999a, Actor/non-actor focus asymmetries in Seediq. In Yuen-mei Yin, I-li Yang, and Hui-chen Chan, eds. *Chinese languages and linguistics*, vol. 5: *Interactions in language* (Symposium Series of the Institute of Linguistics 2), 615–637. Taipei: Institute of Linguistics (Preparatory Office), Academia Sinica.

— [Yung-Li Chang], 1999b, Seediq bound pronouns: Pronominal clitics or agreement affixes? In Elizabeth Zeitoun and Paul Jen-kuei Li, eds. *Selected papers from the eighth International Conference on Austronesian Linguistics* (Symposium Series of the Institute of Linguistics 1), 355–370. Taipei: Institute of Linguistics (Preparatory Office), Academia Sinica.

— [張永利], 2000, *賽德克語參考語法* [A reference grammar of Seediq] (Introduction to Formosan Languages 6). Taipei: Yuanliu.

— 2010, On the syntax of Formosan adverbial verb constructions. In Raphael Mercado, Eric Potsdam, and Lisa deMena Travis, eds. *Austronesian and theoretical linguistics* (Linguistik Aktuell 167), 183–211. Amsterdam: Benjamins.

Chang, Yiying Ann, 2012, Pronominal ordering in Plngawan Atayal. In Enrico Boone, Kathrin Linke, and Maartje Schulpen, eds. *Proceedings of ConSOLE XIX*, 177–190. Leiden: Student Organization of Linguistics in Europe.

Chen, Kang [陈康], 1992, 泰耶尔语群 [The Atayalic languages]. In Chen Kang, ed. *台湾高山族语言* [Languages of the Mountain Tribes of Taiwan], 1–44. Beijing: Central Minorities Institute.

Corbett, Greville G., 2000, *Number*. Cambridge: Cambridge University Press.

Dahl, Otto Christian, 1973, *Proto-Austronesian* (Scandinavian Institute of Asian Studies Monograph Series 15). Lund: Studentlitteratur.

Donohue, Mark, 2007, Word order in Austronesian from north to south and west to east. *Linguistic Typology* 11(2):349–391.

DuBois [Dubois], Carl D., 1976, *Sarangani Manobo: An introductory guide* (Special Monograph Issues of the Philippine Journal of Linguistics 6). Manila: Linguistic Society of the Philippines.

Egerod, Søren, 1980, *Atayal-English dictionary*, 2 fascicles (Scandinavian Institute of Asian Studies Monograph Series 35). London: Curzon.

— 1999, *Atayal-English dictionary*, 2nd (posthumous) edn, Jens Østergaard Petersen, ed. (Historisk-filosofiske Skrifter 20). Copenhagen: Kongelige Danske Videnskabernes Selskab.

Erickson, Barbara E., 1965, Patterns of person-number reference in Potawatomi. *International Journal of American Linguistics* 31(3):226–236.

Errington, Ross, 1979, A transition network grammar of Cotabato Manobo. *Studies in Philippine Linguistics* 3(2):105–163.

Grimshaw, Jane, 1990, *Argument structure* (Linguistic Inquiry Monographs 18). Cambridge, MA: MIT Press.

Harvey, Mark, 1982, Subgroups in Austronesian. In Amran Halim, Lois Carrington, and Stephen A. Worm, eds. *Papers from the third International Conference on Austronesian Linguistics*, vol. 2: *Tracking the travellers* (Pacific Linguistics C-75), 47–99. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Henningsson, Lars-Åke, and Arthur Holmer, 2008, The distribution of quantifiers in Seediq. *Working Papers* [Lund: Department of Linguistics and Phonetics, Lund University] 53:23–41.

Holmer, Arthur, 1993, Atayal clitics and sentence structure. *Working Papers* [Lund: Department of Linguistics and Phonetics, Lund University] 40:71–94.

— 1995, Right movement in Seediq. *Working Papers* [Lund: Department of Linguistics and Phonetics, Lund University] 44:71–90.

— 1996a, *A parametric grammar of Seediq* (Travaux de l'Institut de linguistique de Lund 30). Lund: Lund University Press.

— 1996b, Passive typology – the case link. *Working Papers* [Lund: Department of Linguistics and Phonetics, Lund University] 45:63–78.

— 1999, Review of Chang (1997). *Glot International* 4(6):10–12.

— 2002, The morphology and syntax of Seediq focus. In Fay Wouk and Malcolm Ross, eds. *The history and typology of western Austronesian voice systems* (Pacific Linguistics 518), 333–354. Canberra: Research School of Pacific and Asian Studies, Australian National University.

— 2004, Intraposition and Formosan adverbial heads. *ZAS Papers in Linguistics* 34 [special issue: Paul Law, ed. Proceedings of AFLA 11]:120–134.

— 2005, Seediq: Antisymmetry and final particles in a Formosan VOS language. In Andrew Carnie, Heidi Harley, and Sheila Ann Dooley, eds. *Verb first: On the syntax of verb-initial languages* (Linguistik Aktuell 73), 175–201. Amsterdam: Benjamins.

— 2006, Seediq: Adverbial heads in a Formosan language. In Hans Martin Gärtner, Paul Law, and Joachim Sabel, eds. *Clause structure and adjuncts in Austronesian languages* (Studies in Generative Grammar 87), 83–123. Berlin: Mouton de Gruyter.

— 2010, Seediq adverbial verbs: A review of the evidence. In Raphael Mercado, Eric Potsdam, and Lisa deMena Travis, eds. *Austronesian and theoretical linguistics* (Linguistik Aktuell 167), 163–182. Amsterdam: Benjamins.

— 2012, Evidence from Formosan for a unified theory of adverb ordering. *Lingua* 122(8):902–921.

Hsiao, Stella I-Ling, 2004, Adverbials in Squliq Atayal. MA thesis. Hsinchu: National Tsing Hua University.

Huang, Lillian M., 1989, The pronominal system in Atayal. *Studies in English Literature and Linguistics* [Taipei: Department of English, National Taiwan Normal University] 15(May):115–133.

— 1993, *A study of Atayal syntax*. Taipei: Crane.

— 1995a, The pronominal system in Mayrinax, Atayal. *Studies in English Literature and Linguistics* [Taipei: Department of English, National Taiwan Normal University] 21(June):107–146.

— 1995b, *A study of Mayrinax syntax*. Taipei: Crane.

— [黃美金], 2000, *泰雅語參考語法* [A reference grammar of Atayal] (Introduction to Formosan Languages 1). Taipei: Yuanliu.

Huang, Lillian M., Elizabeth Zeitoun, Marie M. Yeh, Anna H. Chang, and Joy J. Wu, 1999, A typological overview of pronominal systems of some Formosan languages. In H. Samuel Wang, Feng-fu Tsao, and Chin-fa Lien, eds. *Selected papers from the fifth International Conference on Chinese Linguistics*, 1.165–198. Taipei: Crane.

Hung, Silvia Yu-ju, and Loren Billings, 2009/2010, Topicality and pronominal ordering in two Manobo languages. Presented at the 11th International Conference on Austronesian Linguistics (Aussois, France; June 2009). Uploaded to the Rutgers Optimality Archive <roa.rutgers.edu> (June 2010).

Johnston, E. Clay, 1975, The verb affixation of Cotabato Manobo. *Philippine Journal of Linguistics* 6(1):25–50.

Kaufman, Daniel, 2010, The grammar of clitics in Maranao. In Loren Billings and Nelleke Goudswaard, eds. *Piakandatu ami Dr. Howard P. McKaughan*, 179–204. Manila: Linguistic Society of the Philippines and SIL Philippines.

Kaufmann, John, 1916, *Principles of Visayan grammar*. Manila: Catholic Trade School.

Kerr, Harland B., 1965, The case-marking and classifying function of Cotabato Manobo voice affixes. *Oceanic Linguistics* 4(1–2):15–47.

— 1988, Cotabato Manobo grammar. *Studies in Philippine Linguistics* 7(1):1–123.

Lee, Amy Pei-jung, 2011, Comitative and coordinate constructions in Truku Seediq. *Language and Linguistics* 12(1) [special issue: Stacy F. Teng, ed. Coordination and comitativity in Austronesian languages]:49–75.

Lee, Celeste Chia-Yen, 2008, Clitic pronouns in Masbatenyo. *Studies in Philippine Languages and Cultures* 17 [special issue: Sue McQuay, ed. 10-ICAL pronoun papers]:121–136.

Lee, Celeste, and Loren Billings, 2008, Clitic-pronoun clusters in Central Philippine. In Wilaiwan Khanittanan and Paul Sidwell, eds. *SEALS XIV volume 1: Papers from the 14th annual meeting of the Southeast Asian Linguistics Society 2004* (Pacific Linguistics E-5), 193–203. Canberra: Research School of Pacific and Asian Studies, Australian National University.

— 2012, Avoiding co-occurrences of monosyllabic pronouns in two Austronesian subgroups. Presented at the 12th International Conference on Austronesian Linguistics (Udayana University, Denpasar, Bali, Indonesia).

Li, Lilian Li-ying, 2010, Clitics in Nantou Isbukun Bunun (Austronesian). MA thesis. Puli, Nantou, Taiwan: National Chi Nan University.

Li, Paul Jen-kuei, 1980, The phonological rules of Atayal dialects. *Bulletin of the Institute of History and Philology* [Taipei: Academia Sinica] 51(2):349–405 (republished in Li 2004:1.229–282).

— 1981, Reconstruction of Proto-Atayalic phonology. *Bulletin of the Institute of History and Philology* [Taipei: Academia Sinica] 52(2):235–301 (republished in Li 2004:1.625–692).

— 1985a, Linguistic criteria for classifying the Atayalic dialect groups. *Bulletin of the Institute of History and Philology* [Taipei: Academia Sinica] 56(4):699–718 (republished in Li 2004:2.835–852).

— 1985b, The position of Atayal in the Austronesian family. In Andrew Pawley and Lois Carrington, eds. *Austronesian linguistics at the 15th Pacific Science Congress* (Pacific Linguistics C-88), 257–280. Canberra: Department of Linguistics, Research School of Pacific Studies, Australian National University (republished in Li 2004:2.805–833).

— 1995, The case-marking system in Mayrinax, Atayal. *Bulletin of the Institute of History and Philology* [Taipei: Academia Sinica] 66(1):23–52 (republished in Li 2004:1.387–414).

— 1998, The dialectal differences in Atayal in I-Lan. In Huang Shuanfan, ed. *Selected papers from the second International Symposium on Languages in Taiwan*, 49–78. Taipei: Crane (republished in Li 2004:1.711–742).

— 2004, *Selected papers on Formosan languages*, 2 fascicles (Language and Linguistics Monograph Series C-3). Taipei: Institute of Linguistics, Academia Sinica.

— 2006, Introduction. In Paul Jen-kuei Li and Shigeru Tsuchida. *Kavalan dictionary* (Language and Linguistics Monograph Series A-19), 1–62. Taipei: Institute of Linguistics, Academia Sinica.

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

— 2005, Another look at the order of clitic pronouns in Wulai Atayal. *Concentric: Studies in Linguistics* 31(1) [special issue: Austronesian linguistics]:47–63.

Lin, Hsiu-hsu, 2005, The grammaticalization of tense/aspect auxiliaries in Seediq. *Concentric: Studies in Linguistics* 31(2):111–132.

Lu, Anne Yu-an, 2005, Mayrinax phonology: With special reference to *UM* affixation. MA thesis. Hsinchu: National Tsing Hua University.

McFarland, Curtis D., 2001, Filipino enclitics. *Philippine Journal of Linguistics* 32(1):1–15.

McKaughan, Howard P., 1959, Semantic components of pronoun systems: Maranao. *Word* 15(1):101–102.

— 1999, An introduction to Iranun verb structure. *Sabah Museum Journal* 1(4):46–91.

Müller, Gereon, 1999/2003, Optionality in optimality-theoretic syntax. *Glot International* 4(5):3–8 (republished, with numerous corrections, in Lisa Cheng and Rint Sybesma, eds. *The second* Glot International *state-of-the-article book: The latest in linguistics* (Studies in Generative Grammar 61), 289–321. Berlin: Mouton de Gruyter).

Ochiai, Izumi, 2009, A study of Seediq pronouns. MA thesis. Taipei: National Taiwan University.

Peng, Adam, and Loren Billings, 2008, Binukid pronominal clisis. *Studies in Philippine Languages and Cultures* 17 [special issue: Sue McQuay, ed. 10-ICAL pronoun papers]:179–212.

Pike, Kenneth L., and Barbara Erickson, 1964, Conflated field structures in Potawatomi and in Arabic. *International Journal of American Linguistics* 30(3):201–212.

Pittman, Richard S., 1965, The fused subject and object pronouns in Red Pheasant Cree. *Linguistics* 3(13):34–38.

Prince, Alan, and Paul Smolensky, 1993, Optimality theory: Constraint interaction in generative grammar. Technical Report TR-2, Rutgers Center for Cognitive Science, Rutgers University, New Brunswick, NJ.

— 2004, *Optimality theory: Constraint interaction in generative grammar*. Malden, MA: Blackwell.

Quakenbush, J. Stephen, and Edward Ruch, 2008, Pronoun ordering and marking in Kalamianic. *Studies in Philippine Languages and Cultures* 17 [special issue: Sue McQuay, ed. 10-ICAL pronoun papers]:213–234.

Rau, Der-Hwa Victoria, 1992/1992/1992, A grammar of Atayal. PhD dissertation. Ithaca: Cornell University (also Ann Arbor: University Microfilms and Taipei: Crane).

Ross, Malcolm, 2006, Reconstructing the case-marking and personal pronoun systems of Proto Austronesian. In Henry Y. Chang, Lilian M. Huang, and Dah-an Ho, eds. *Streams converging into an ocean: Festschrift in honor of Professor Paul Jen-kuei Li on his 70th birthday* (Language and Linguistics Monograph Series W-5), 521–563. Taipei: Institute of Linguistics, Academia Sinica.

— 2009, Proto Austronesian verbal morphology: A reappraisal. In Alexander Adelaar and Andrew Pawley, eds. *Austronesian historical linguistics and culture history: A festschrift for Robert Blust* (Pacific Linguistics 601), 295–326. Canberra: Research School of Pacific and Asian Studies, Australian National University.

— 2012, In defense of Nuclear Austronesian (and against Tsouic). *Language and Linguistics* 13(6):1253–1330.

Schachter, Paul, 1973, Constraints on clitic order in Tagalog. In Andrew B. Gonzalez, ed. *Parangal kay Cecilio Lopez: Essays in honor of Cecilio Lopez on his seventy-fifth birthday* (Philippine Journal of Linguistics, Special Monograph Issue 4), 214–231. Quezon City: Linguistic Society of the Philippines.

Silverstein, Michael, 1976, Hierarchy of features and ergativity. In R.M.W. Dixon, ed. *Grammatical categories in Australian languages*, 112–171. Canberra: Australian Institute of Aboriginal Studies.

Starosta, Stanley, 1974, Causative verbs in Formosan languages. *Oceanic Linguistics* 13(1–2) [special issue: George W. Grace and Elmer Wolfenden, eds. Papers of the first International Conference on Comparative Austronesian Linguistics, 1974: Proto-Austronesian and western Austronesian]:279–369.

Tseng, Temi Nawi C./曾瑞琳, 2009, *Kari pnsltudan rudan* [Traditional tales of the ancestors]/*賽德克族口述傳統文化故事* [Stories from Seediq oral traditions], vol 1. [Puli, Nantou, Taiwan:] 普羅文化出版廣告事業.

— 2010, *Kari pnsltudan rudan* [Traditional tales of the ancestors]/*賽德克族口述傳統文化故事* [Stories from Seediq oral traditions], vol 2. [Puli, Nantou, Taiwan:] 奇詮電腦打字印刷行.

— forthcoming, *Kari pnsltudan rudan* [Traditional tales of the ancestors]/*賽德克族口述傳統文化故事* [Stories from Seediq oral traditions], vol 3.

Tseng, Temi Nawi C. [曾瑞琳], and Li-May Sung [宋麗梅], 2010, 賽德克語 [The Seediq language]. In *住民志* [The residents] vol. 4:*語言篇* [Languages] (南投縣志 [The Annals of Nantou County] 2), 176–205. Nantou City, Taiwan: Cultural Affairs Bureau of Nantou County [南投縣文化局].

Tsukida, Naomi, 2005, Seediq. In Alexander Adelaar and Nikolaus P. Himmelmann, eds. *The Austronesian languages of Asia and Madagascar* (Routledge Language Family Series 7), 291–325. London: Routledge.

Van Valin, Robert D., Jr., 2005, *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press.

Weaver, Dan, and Marilou Weaver, 1964, Ranking of personal pronouns in Agusan Manobo. *Oceanic Linguistics* 3(1) [special issue: George W. Grace and Elmer Wolfenden, eds. Papers in Philippine linguistics]:161–170.

Yen, Doris Ching-jung, 2012, Clitics in Kavalan (East Formosan, Austronesian). MA thesis. Puli, Nantou, Taiwan: National Chi Nan University.

Yen, Doris Ching-jung, and Loren Billings, 2011, Sequences of pronominal clitics in Mantauran Rukai: V-deletion and suppletion. In Lauren Eby Clemens, Gregory Scontras, and Maria Polinsky, eds. *Proceedings of the eighteenth meeting of the Austronesian Formal Linguistics Association (AFLA), Harvard University, Cambridge (MA), March 4–6, 2011*, 168–182. London, Ontario, Canada: Linguistics Program, University of Western Ontario <westernlinguistics.ca/afla/proceedings/afla18/index.htm>.

— forthcoming, The cluster-internal ordering of clitics in Kavalan (East Formosan, Austronesian). To appear in the *Proceedings of the thirty-eighth annual meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society.

Zwicky, Arnold M., and Geoffrey K. Pullum, 1983, Cliticization vs. inflection: English *n't*. *Language* 59(3):502–513.

7 How to Index a Subject in Tolaki

OWEN EDWARDS

1 Introduction[1]

Tolaki is an Austronesian language spoken on mainland South-East Sulawesi by around 280,000 people (Mead 1999:71). In Tolaki, nominal arguments are cross-referenced by clitics on the verb. There are four classes of such clitics: nominative, absolutive, genitive and dative.

While the S/A argument in Tolaki can be shown to be the syntactic subject (Edwards 2012:71-3), many authors (Mead 1998; 2002, Mead and Youngman 2008, Edwards 2012) have noted that Tolaki has a 'fluid-S system' in which the S/A can be indexed with either nominative, absolutive or genitive clitics. However apart from Mead (1998:300-37) and Edwards (2012:36-46) little effort has been made to comprehensively characterise the environments in which each form of indexation is employed.[2] Furthermore, the discussion in Mead (1998) is chiefly from a historical perspective, though there is ample data from modern Bungku-Tolaki languages, including Tolaki, to illustrate the points made.

In this paper I seek to characterise those environments in which each type of indexation is employed. We will see that a complex interaction of semantic, syntactic and morphological environments determines the coding of the S/A. We will furthermore see that nominative clitics have become the default set for subject indexation in Tolaki, though there are also environments in which nominative indexation is required by morphology or syntax. This shows a development from an earlier stage of the language in which absolutive clitics were the "historically unmarked set for indexing subjects on non-transitive verbs" (Mead 1998:332).

---

[1] Unless otherwise cited, data is drawn from my own fieldwork conducted mainly at the beginning of 2012. The variety of Tolaki described in this paper is that spoken in the Unaaha and Wawotobi regencies (within the *Konawe* dialect area), speech that appears specific to Kendari is flagged as 'urban'. I would like to thank my main informants, Darmin, Untung, Sukur Tabara and Omar Pidani.

Glosses follow the list of standard abbreviations in the Leipzig Glossing Rules, with the following exceptions: ACCID 'accidental passive, CN 'common noun', CONF 'confirmative', EX 'exclusive', DESID 'desiderative', IN 'inclusive', NFIN 'non-finite', NSG 'non-singular' and PN'proper noun'.

Example sentences are given in standard Tolaki orthography, with hyphens (-) added to indicate breaks between morphemes. Tolaki letters have the same value as their Indonesian equivalents, with the exception of the apostrophe <'>which represents the glottal stop /ʔ/. Full sentences receive appropriate capitalisation and punctuation. A capital 'N' (i.e. *poN*-) indicates a morpheme after which the morpho-phonemic process of prenasalisation occurs. Under this process the voiceless stops /p, t, k/become the prenasalised stops /mb, nd, ŋg/. Two lines of Tolaki are given in sentences when this morphophonemic process operates. The top line shows the standard orthography, the second line the morpheme breaks.

[2] Mead (2002:156-60) and Mead and Youngman (2008) are concerned primarily with other facts about the language and provide only a brief overview to illustrate that the S/A has multiple encodings.

Asia-Pacific Linguistics, 2014.

1.1 Outline

After a discussion on alignment systems in section 2, including a discussion of split alignment systems, I discuss the data found in Tolaki.

In section 4 I show that nominative clitics are used to index the S/A argument when certain preverbal conjunctions occur, when the verb is negated and in imperative sentences. I also show that nominative clitics are best considered the default indexation strategy for the S/A argument.

In section 5 I show that genitive clitics can be used to index the S/A argument of a verb when this argument is an experiencer as well as indexing the S/A argument of a temporally dependent clause.

In section 6 I show that absolutive clitics are used to index the S/A argument of a verb in an aspectually marked clause.

2 Alignment

'Alignment' is a term that refers to the morphological or syntactic coding of the grammatical relations S, A and P. S is defined as the sole argument of an intransitive verb, A as the most agent-like argument of a transitive verb and P[3] as the most patient-like argument of a transitive verb (Comrie 1978).

For languages which code these grammatical relations differently, there are three systems: (i) *tripartite*: each of S, A and P is encoded differently, (ii) *ergative/absolutive* system: S/P are encoded identically while A is encoded differently and (iii) *nominative/accusative*: S/A are encoded identically while P is encoded differently. The last two of these strategies are the most common among the world's languages (Comrie 2011). They are represented in (1) a. and (1) b. below.

(1) a. **Nominative/Accusative:**

<table>
<tr><td>NOM</td><td>ACC</td></tr>
<tr><td>S</td><td rowspan="2">P</td></tr>
<tr><td>A</td></tr>
</table>

b. **Ergative/Absolutive:**

<table>
<tr><td>ERG</td><td>ABS</td></tr>
<tr><td rowspan="2">A</td><td>S</td></tr>
<tr><td>P</td></tr>
</table>

English is a nominative/accusative language. Thus the P bears accusative case, as shown by the pronoun form '*him*' in sentence (2), while the S and A bear nominative case, as shown by the pronoun form '*he*' in both sentences (2) and (3).

(2) He$_S$ came.
NOM

(3) He$_A$ saw him$_P$
NOM ACC

Dyirbal (Dixon 1979) noun phrases show ergative/absolutive marking. In Dyirbal the A argument takes the suffix *-ŋgu* (for disyllabic stems), like *yabu* 'mother' in sentence (5) while the S and P arguments are unmarked, as is *ŋuma* in both (4) and (5).

[3] 'O' is also sometimes used instead of 'P' (i.e. Dixon 1979).

(4) *ŋuma banaga-ɲu*
father return-NFUT
'Father returned.'

(5) *ŋuma yabu-ŋgu buṛa-n*
father mother-ERG see-NFUT
'Mother saw father.' (Dixon 1979:61)

2.1 Split Alignment

The two systems represented in (1) can be described by the way in which S is encoded. If the S is encoded in the same way as P, the system is ergative/absolutive, if the S is encoded in the same way as A, the system is nominative/accusative. However, languages often display so-called 'split systems', in which S is sometimes encoded like a P and sometimes encoded like an A.

In fact, Dyirbal is such a language; thus while the ergative/absolutive system holds for full NPs and 3rd person pronouns, non-3rd person pronouns are coded along nominative/accusative lines.

(6) *ŋana banaga-ɲu*
1PL return-NFUT
'We returned.'

(7) *ŋana ɲura-na buṛa-n*
1PL 2PL-ACC see-NFUT
'We saw you.' (Dixon 1979:64)

In a language with a split system, there are many factors that can determine the split. In Dyirbal 3rd person referents are encoded along ergative/absolutive lines while other referents are coded along nominative/accusative lines. Dixon (1979; 1994) gives five different factors along which a split can occur. These include:

- the semantic nature of the verb
- the semantic nature of the NP
- bound pronominal vs. free pronoun split
- split conditioned by the aspect or tense of a sentence
- split according to main vs. subordinate clauses

I will illustrate one of these systems, a split according to semantic nature of the verb. In Kurripako (Danielsen and Granadillo 2008:398), a North Arawak language, both an A and an agentive S (S_A) are encoded with a verbal prefix. This is shown below, in which the prefix *nu-* encodes a first person S for 'return' in sentence (8) as well as a first person A for 'barbecue' in sentence (9).

(8) ***Nu**-dia-ka-wa panti-liku.*
1SG-return-PROG-INTR house-LOC
'I'm returning to the house.'

(9) ***Nu**-heema srua.*
1SG-barbecue 3SG.F
'I barbecue it.'

However, a non-agentive S (S_P) is encoded in the same way as a P. In Kurripako, a P is encoded with a postverbal pronoun or noun phrase. Thus we find both the P of 'barbecue' in sentence (10) and the S of 'be tired' in sentence (11) realised as post-verbal pronouns:

(10) *Haamaa-ka* ***hnua***.
tired-PROG 1SG
'I'm tired.'

(11) *Nu-heema* ***srua***.
1SG-barbecue 3SG.F
'I barbecue it.'

Kurripako thus cannot be characterised as either wholly nominative/accusative nor ergative/absolutive. Instead of one of the systems represented in (1), Kurripako displays a classic case of Semantic Alignment, represented in (12).

(12) **Semantic Alignment:**

AGT	PAT
S_A	S_P
A	P

Semantic alignment is defined by Donohue (2008:74) as: "A split in the morphological encoding of arguments according to some feature of the lexical semantics of the verb."[4]

Another instance of semantic alignment is found in Icelandic. In Icelandic subjects can bear one of four cases, nominative, accusative, dative or genitive. The case borne by the subject is lexically specified, though it is possible to identify some loose semantic characteristics associated with non-nominative subjects. Thus the normal pattern in Icelandic is for the subject (S/A) to take nominative case, and the object accusative case. Thus the S in sentence (13) and the A in sentence (14) are nominative, while the P in (14) is accusative.

(13) *Ég* *sef.*
1SG.**NOM** sleep:1SG.PRES
'I'm sleeping.'

(14) *Hún* *sá* *mig.*
3SG.F.**NOM** see:3SG.PAST 1SG.**ACC**
'She saw me.'

However, the subject of some verbs takes accusative, dative or genitive case. All non-nominative subjects in Icelandic can be characterised as non-agentive (Andrews 1982:463). For slightly more fine-grained semantic characteristics that can be identified loosely with each subject case marking strategy, see Andrews (1982:463) and Andrews (2001:99). An accusative S is shown in (15), while an accusative A is shown in (16).

[4] According to this definition Dyirbal does not display 'semantic alignment' as it is not the verbal semantics that determines the split, but the person of the arguments.

(15) *Hana* *dreym-di* *(um* *haf-ið.)*
3SG.**ACC** dream-PAST about sea-DEF
'She dreamt (about the sea.)' (Andrews 2001:102)

(16) *Mig* *vant-ar* *hnif.*
1SG.**ACC** lack-3SG:PRES knife:SG.ACC
'I don't have a knife.' (Andrews 2001:100)

A dative subject is shown in (17) below. Dative subjects are typically, though not exclusively experiencers.

(17) *Mér* *kóln-ar.*
1SG.**DAT** get.cold-3SG:PRES
'I'm getting cold.' (Andrews 1982:462)

Having discussed different split alignment systems in a variety of languages I will now turn my attention to the system we find in Tolaki.

3 Pronominals in Tolaki

Before my discussion of the different Tolaki indexation strategies, I will provide an overview of the different pronominal paradigms in Tolaki and some important properties associated with them. Tolaki pronominals have 4 person and 2 number categories.[5]

The Tolaki free pronouns are given in Table 3.1. They occur as independent NPs within the clause and are used when the referent is emphasised, or contrasted with other participants in the discourse.

Table 3.1: Independent Pronouns

	SG	NSG
1EX	*inaku*	*inggami*
1IN		*inggito*
2	*inggo'o*	*inggomiu*
3	*iee, ie'i*	*ihiro*

In addition to these free pronouns there are four pronominal clitic paradigms found in Tolaki; nominative, absolutive, genitive and dative. These are given in Table 3.2.

Table 3.2: Pronominal Clitics

	NOMINATIVE		GENITIVE		ABSOLUTIVE		DATIVE	
	SG	NSG	SG	NSG	SG	NSG	SG	NSG
1EX	*ku-*	*ki-*	*-nggu*	*-mami*	*-'aku*	*-komami*	*-kona*	*-komami*
1IN		*to-*		*-ndo*		*-keito*		*-keito*
2	*u-*	*i-*	*-mu*	*-miu*	*-ko*	*-komiu*	*-ko'o*	*-komiu*
3	*no-*	*ro-*	*-no*	*-ro*	*-'i,* *(-e, -o)*	*-'iro,* *(-ero, -oro)*	*-kee*	*-ke(h)ero*

[5] The number categories are SINGULAR and NON-SINGULAR.NSG refers to two or more entities. PLURAL refers to three or more individuals and is expressed by the dedicated prefix *mbeN-,* as in sentence (57).

Phonologically, these clitics form a word with the stem to which they attach, as revealed by stress placement.[6] Stress in Tolaki falls on the penultimate syllable of a word, and secondary stress on every second preceding syllable.[7] When an enclitic occurs on a word, stress shifts accordingly. In example (18) a. primary stress falls on the penultimate syllable of the word; the first syllable of the verb stem. In example (18) b., with two enclitics, primary stress likewise falls on the penultimate syllable; in this case the absolutive clitic. Primary stress is shown by an acute accent and secondary stress by a grave accent.

(18) a. *No-láko*
3NOM-go
'S/he goes'

b. *Làko-'í-to.*
go-3ABS-PRF
'S/he has gone.'

Syntactically, the genitive, absolutive and dative enclitics attach to the end of the verb. When postverbal adverbial material occurs, the enclitics attach to the end of this material. Sentence (19) a. shows an absolutive enclitic attaching directly to a verb, while sentence (19) b. shows the adverbial *mendua* 'again' occurring between the verb and enclitic.

(19) a. *Ninggiro* *a-ku* *talipo-ko.*
later so-1NOM phone-2ABS
'I'll call you later today.'

b. *Ninggiro* *a-ku* *talipo* *mendua-ko.*
later so-1NOM phone again-2ABS
'I'll call you back later today.'

The nominative clitics typically attach to the beginning of a verb, as in sentence (18) a. When certain preverbal conjunctions occur, however, they are attracted forward as enclitics to these conjunctions, as in the examples in (19). This is discussed in more detail in section 4.2.

Additionally, whenever any morphology, including a nominative proclitic, attaches to the beginning of a vowel initial stem, a glottal stop is automatically inserted at the clitic boundary (this glottal stop is represented orthographically). The process of automatic glottal stop insertion does not occur at enclitic morpheme boundaries. Automatic glottal stop insertion is shown in example (20).

(20) no-ehe-ko → /noʔeheko/ → <*no'eheko*>
3NOM-like-2ABS
'S/he loves you.'

In (21) below I give the Tolaki verbal template. For the sake of simplicity, only the morphology relevant to this paper is included.

---

[6] This fact, along with the lack of evidence to determine whether other morphemes are clitics or affixes, is the reason a hyphen (-) is used to indicate a clitic boundary rather than an equals sign (=).

[7] Nouns provide an exception to this rule. A 5 syllable noun has primary stress assigned to the penultimate syllable, and secondary stress assigned to the first syllable, rather than the expected second syllable.

(21)

–3	–2	–1	0	+1	+2	+3
NOM	DESID	INDF.P	STEM	ABS GEN	DAT	ASPECT

4 Nominative Indexation

I begin my discussion of Tolaki S/A indexation with a discussion of nominative clitics. The full set of nominative clitics found in Tolaki is repeated in Table 4.1.

Table 4.1: Nominative Clitics

	SG	NSG
1EX	*ku-*	*ki-*
1IN		*to-*
2	*u-*	*i-*
3	*no-*	*ro-*

4.1 Default Strategy

Nominative clitics are best considered the default, or unmarked, indexation strategy for indexing the S/A argument in Tolaki. Thus, when a verb occurs without any of the conditions that force genitive or absolutive indexation (see sections 5 and 6), the S/A argument is indexed with nominative clitics. Examples with an S and A are given in (22) and (23) respectively.

(22) *Ku-lako* *i-daoa.*
1NOM-go LOC-market
'I went to the market.'

(23) *Ku-soro-'i* *oto-nggu.*
3NOM-push-3ABS car-1GEN
'I pushed my car.'

I return to the question of markedness in my conclusions in section 7. There are also some situations in which nominative clitics are obligatory, even if other factors would require indexation with a genitive or absolutive clitic. These environments are given in sections 4.2–4.4 below.

4.2 Conjunctions

One environment in which nominative clitics are obligatory is after certain preverbal conjunctions. Mead (1998:131) provides five such conjunctions, of which the most common two are *a-* 'and, so that' and *ke-* 'if'. When these conjunctions occur the clitic is attracted forward and occurs as an enclitic on the conjunction.

An example of an S being indexed in this way after the conjunction *ke-* is given in (24) a. an example of an A being indexed thus after the conjunction *a-* is given in (24) b.

(24) a. ***Ke-u*** *ehe,* *kolopua* ...
if-2NOM like, tortoise
'If you want, O tortoise, …'

b. … ***a-to*** *lako* *mo-hoko* *haka-pundi*
so-1IN.NOM go NFIN.INDF.P-dig.up root-banana.tree
'… then we will go dig up banana tree roots.' (Untung 2009:28)

For some urban speakers the 3rd person nominative clitic *no* has become an obligatory part of the conditional conjunction *ke-*, no matter the person and number of the subject.[8] For these speakers the nominative proclitic indexing the subject occurs before the verb and after the conjunction *keno*, as in sentence (25).

(25) *Mano* ***keno*** *u-taa* *eta'i-ki,* *u'-onggo* *mate* *ikeni* *dowo-mu!*
but if 2NOM-NEG follow-CONF, 2NOM-FUT die here self-2GEN
'But if you don't follow, you will die here by yourself!'

4.3 Negation

Nominative clitics also occur when the verb is negated. A variety of negators can be used to negate a verb in Tolaki including *oki*, *ki'oki* and *taa*. The most common of these in my data is *oki* which occurs preverbally and like the conjunctions described in 3.2 attracts the nominative clitic forwards. An example with an A is given in (26).

(26) ***Oki-no*** *kii-'i* *inono* *o-manu.*
NEG-3NOM see-3ABS this CN-chicken
'She didn't see this chicken.'

When a conditional clause is negated, the most common pattern is for the negator *taa* to come between the nominative clitic and the verb. This is shown in (27) below.[9]

(27) ***Ke-u*** *taa* *leu ...*
if-2NOM NEG come
'If you don't come …'

The negator *taa* can also occur as an alternative to *oki*, in which case the nominative clitic occurs either before it, as in sentence (28) or after it, as in sentence (29). No factors conditioning the alternate orderings have been identified so far.

---

[8] Note that the conjunction *mano* 'but' ends in *no* for all speakers. The probable historical origin of this conjunction is likewise *ma* + *no* 'but-3NOM'. This conjunction could be related to the Tolaki interjection *maa*, which is used to signal that the clause following is new information (roughly equivalent to the English interjection 'well') as in example (i):

(i) *Lako-no-to* *inaku.* ***Maa,*** *po'opo* *ku-hori* *me-rapu*
go-3GEN-PRF 1SG Well, not.yet, 1NOM-not.yet NFIN.INTR-marry
'Then [there's] me. Well, I'm not married yet.'

[9] Alternately, the negators *oki* or *ki'oki* can precede the conditional conjunction *ke-*, in which case the sentence has a negative irrealis meaning, as in (ii).

(ii) *Maa,* ***ki'oki*** ***ke****-ku,* *tule-'i* *ni'ino.*
well, NEG if-1NOM can-3ABS this
'Well, I would not be able to do this.' (Untung 2009:28)

(28) *Mbaako-'i* *ni'ino* *motoro-nggu* ***no-taa*** *ehe* *ruru?*
why-3ABS this motorbike-1GEN 3NOM-NEG like start
'Why doesn't this motorbike of mine like to start?'

(29) *No-rodo-'aku* *i'-ama* *i'-ina* *melewelewe-'aku-to*
3NOM-crush-1ABS PN-father PN-mother flat-1ABS-PRF

taa-ku *mate.*
NEG-1NOM die
'My parents crushed me, I became flat but didn't die.' (Untung 2009:32)
(This is the tortoise's response when certain monkeys threaten to crush him. Eventually the tortoise convinces them that the only way to kill him is by throwing him into the water, upon which he makes his escape. See sentence (49).)

4.4 Imperatives

A final context in which nominative clitics occur obligatorily is in imperative sentences. While singular S/A arguments are unindexed in imperatives, non-singular arguments are indexed with a nominative clitic. This is shown in sentence(30) below.

(30) *Iamo-to* ***i****-po'-oli* ***i'****-ale-'i* *tokaa.*
NEG-PRF 2NSG.NOM-INDF.P-buy 2NSG.NOM-take-3ABS just
'Don't buy [it], just take it.'

5 Genitive Clitics

In this section I will discuss the circumstances under which the S/A argument of a verb are indexed with genitive clitics. The genitive clitics found in Tolaki are given in Table 5.1.

Table 5.1: Genitive Clitics

	SG	NSG
1EX	*-nggu*	*-mami*
1IN		*-ndo*
2	*-mu*	*-miu*
3	*-no*	*-ro*

The most common function of these clitics is to indicate the possessor of a noun phrase. This is shown in (31) and (32) below.

(31) *Inono* *gondi-**mu**.*
this scissors-2GEN
'These are your scissors.'

(32) *Inono* *banggona-**no*** *ama-**nggu**.*
this friend-3GEN father-1GEN
'This is my father's friend.'

However, under certain circumstances they can also be used to index the S/A argument of a verb. This occurs with certain verbs that take an experiencer and with temporally dependent clauses.

5.1 Experiencers

Some verbs which take an experiencer S/A role in Tolaki index this experiencer with a genitive clitic.

The first type of such verbs is those that take the desiderative prefix *moko-*. Of these, some are transparent derivations from other verbs, such as *mokombo'iso* 'feel sleepy' from *mokoN-* + *po'iso*; DESID + sleep, while other predicates do not occur without this prefix, such as *moko'au* 'miss', *mokorai* 'be cold' and *moko'uo* 'be thirsty'.[10]

An example of a derived desiderative predicate is given in (33) and an example of a predicate which does not occur without this prefix in (34).

(33) *Mokombo'inunggu* *o-kopi.*
mokoN-po'-inu-**nggu** o-kopi.
DESID-INDF.P-drink-1GEN CN-coffee
'I feel like drinking (some) coffee.'

(34) *Moko'-au-**nggu.***
DESID-miss-1GEN
'I'm homesick/I miss [my family].'

When a clitic which occupies the same morphological slot as the genitive clitic occurs, the experiencer S/A is indexed with a nominative clitic. Thus, the 'missed' participant in (34) can be included with a dative suffix, as in sentence (35), in which case the S/A must be indexed with a nominative clitic. Alternately, the 'missed' person can be included as an adjunct, as in sentence (36), in which case the S/A is indexed with a genitive clitic, as expected.

(35) *Nomoko'aunggo'oto* *anamu.*
no-moko'-auN-ko'o-to ana-mu.
3NOM-DESID-miss-2DAT-PRF child-2GEN
'Your daughter's been missing you.'

(36) *Moko'-au-**ro*** *ana* *motu'o-nggu* *kei-inaku.*
DESID-miss-3NSG.GEN child old-1GEN ADJCT.PN-1SG
'My parents [*lit.* 'old children'] miss me.'

Note that while genitive and absolutive clitics occupy the same morphological slot on the verb, absolutive and dative clitics do not (see the verbal template (21) in section 3), thus the so far unattested (37) would be morphologically possible as an alternate rendering of (35).

---

[10] Both *moko'au* and *mokorai* have alternate forms beginning with *mokoko-* i.e. *mokoko'au* and *mokokorai*. (I do not know whether *moko'uo* likewise has a similar alternate form.) These forms appear to consist of *moko-* + *ko-* + VERB; DESID + intensive + VERB. That another morphological element can occur between *moko-* and the root provides evidence that the *moko-* on these verbs is analysed synchronically, by at least some speakers, as a separate prefix.

(37) ? *moko'-au-'i-ko'o-to* *ana-mu*
DESID-miss-3ABS-2DAT-PRF child-2GEN
'Your daughter's been missing you.'

That the S/A is indexed with nominative proclitic, as in sentence (35), rather than with an absolutive clitic provides more evidence that nominative indexation is the default indexation strategy (see section 7.1. for more discussion of the idea of default indexation).

There are also a small number of predicates with a genitive indexed experiencer S/A that do not take the desiderative prefix. For these predicates, genitive indexation is optional and is found in variation with nominative indexation. Such predicates include *mokula* 'be hot', *motaku* 'be afraid' and, for some speakers, *me'aro* 'be hungry'. Two examples are given in (38) and (39) below.

(38) *Mokula-**nggu*** *pake-'i* *inono* *o-babu.*
hot-1GEN use-3ABS this CN-shirt
'I feel hot wearing this shirt.'

(39) *Motaku-**nggu**.*
fear-1GEN
'I'm afraid.'

For both of these predicates the S/A role can also be indexed with nominative clitics. When *mokula* takes a nominative subject the meaning is 'be angry', with *motaku* there is no clear semantic difference. Examples are given in (40) and (41).

(40) ***Ku**-mokula* *podea-'i* *nggiro'o* *tulura-no* *taa* *me'ambo.*
1NOM-hot hear-3ABS that speech-3GEN NEG good
'I'm angry hearing the things he said [*lit.* that speech of his] that aren't good.'

(41) ***Ku**-motaku.*
1NOM-fear
'I'm afraid.'

Finally, when a conjunction or negator occurs (see section 4.2), a 3rd person nominative clitic occurs attached to this conjunction or negator, in addition to the expected genitive enclitic on the verb. This is shown with the conditional conjunction *ke-* in sentence (42) and with the negator *oki* in sentence (43).

(42) *Keno* *mokombo'isomu,* *lako* *po'iso.*
ke-**no** mokoN-po'iso-**mu** lako po'iso
if-3NOM DESID-sleep-2GEN go sleep
'If you're feeling sleepy; go and sleep.'

(43) *Okino* *mokombo'isonggu.*
oki-**no** mokoN-po'iso-**nggu**
NEG-3NOM DESID-sleep-1GEN
'I don't feel sleepy.'

Such constructions need to be investigated in more detail to discover whether the 3rd person nominative clitic indexes the syntactic subject (similar to the expletive 'it' in

English 'It's raining'.) or whether it occurs simply to fulfil the morphological requirements of Tolaki conjunctions.[11]

5.2 Temporally Dependent Clauses

Another situation in which the S/A argument of a verb is indexed with genitive clitics is when there is a dependent temporal relationship between two clauses; the clause with genitive indexation being temporally dependent on the other for its time reference. Often such clauses co-occur with the prefix *sa-*, though this is not obligatory.

Examples (44) and (45) illustrate the indexation of an S and A respectively when the *sa-* prefix occurs. Example (46) illustrates a temporally dependent clause without this prefix.

(44) *Saponggiimami* *nggitu'o* *laika,* *note'eni* *i Omar: ...*
sa-poN-kii-**mami** nggitu'o laika, no-te'eni i-Omar
when-INDF.P-see-1EX.GEN that house 3NOM-say PN-Omar
'When we saw that house, Omar said: …'

(45) ***Sa**-mbule-**nggu*** *ari* *sikola,* *pe-rerehu-'aku-to.*
when-return.home-1GEN from school, INTR-sit-1ABS-PRF
'When I returned home from school, I sat down.'

(46) *Laa-**nggu*** *k‹um›ii~kii-'i* *inono* *uwato,* *no-leu* *o'-ule.*
PROG-1GEN ‹NFIN›REDUP~see-3ABS this grub 3NOM-come CN-snake
'While I was having a look at this grub, a snake arrived.'

Such temporally dependent clauses can also occur after the clause on which they are temporally dependent. In this case the perfective clitic *-to* often occurs, and the prefix *sa-* is often lacking. An example can be seen in the second clause of sentence (47).

(47) *Sa-mbule-mami,* *timba-**mami**-to* *i-Kandari.*
when-return.home-1EX.GEN enter-1EX.GEN-PRF LOC-Kendari
'When we returned home, we then entered Kendari.'

I analyse such clauses as a particular type of temporal dependent clause. This is contrary to Mead (1998:306), who analyses these and similar constructions in Bungku-Tolaki languages as separate 'heightened action clauses' which are used to describe "[…] salient events which move the story forward in time."

Part of Mead's motivation for not analysing such clauses as temporally dependent is that he analyses such clauses as syntactically subordinate, which would force him to identify the initial clause in sentences such as (47) as the main clause. However, as Mead (1998:306) suggests in footnote 158, it is possible to analyse such constructions as correlative, or co-dependent. This is the analysis I pursue. Both clauses are dependent on one another for time reference; the event in the first clause is specified as occurring before the event in the second clause, while the second clause specifies the state that arises after the completion of the event in the first clause.

The particular combination *lako*-GEN-*to* 'go-GEN-PRF' has become semantically bleached, and is often best translated simply as 'then' (Mead 1998:307). Mead illustrates this with sentence (48) in which *lako* occurs twice.

---

[11] Note that (42) and (43) were provided by a speaker who does not have *no* as an obligatory part of the conditional conjunction, as discussed in section 4.2 and illustrated in sentence (25).

(48) *Lako-no-to* *lako* *mo-salei.*
go-3GEN-PRF go NFIN.INDF.P-clear
'Then he went and cleared [it].' (Mead 1998:307)

Thus we have seen that the temporal relation of one clause to another can be specified by indexing the A/S role with genitive clitics. Many such sequential events can be encoded in this way. Thus, in sentence (49) three of the four clauses have their S/A indexed with genitive suffixes, while the other clause (*aro lako tumuha'i*) begins with the resultative conjunction *a-*.

(49) *Lakoroto* *mendene'i* *aro* *lako*
lako-**ro**-to meN-tene-'i a-ro lako
go-3NSG.GEN-PRF NFIN.PL-pick.up-3ABS so-3NSG.NOM go

tumuha'i *i'aalaa,* *tuduno* *kolopua* *i'aiwoi* *pototaono.*
t‹um›uha-'i i'-aalaa tudu-**no** kolopua i'-aiwoi pototao-**no**
‹NFIN›drop-3ABS LOC-river land-3GEN tortoise LOC-water laugh-3GEN
'Then they [the monkeys] went [and] picked him [the tortoise] up, so then they went [and] dropped him in the river, [and] then the tortoise landed in the water [while] laughing.' (Untung 2009:32)

6 Absolutive Clitics

Finally, we find that absolutive clitics[12] index the S/A when the verb is aspectually marked. The full set of Tolaki absolutive clitics is given in Table 6.1.

Table 6.1: Absolutive Clitics

	SG	NSG
1EX	*-'aku*	*-komami*
1IN		*-keito*
2	*-ko*	*-komiu*
3	*-'i*	*-'iro*
	(-e, -o)	*(-ero, -oro)*

The 3rd person absolutive clitic is realised as *-'i* or *-e* when the verb ends in a non-high vowel (/a/, /e/ and /o/) and as *-'i* or *-o* after the high vowels /i/ and /u/. The *-e* and *-o* variants are extremely rare in my data, occurring in only a handful of sentences. I know of the 3NSG variants *-ero* and *-oro* only from Mead (1998:140).

The normal use of these clitics is to index the definite P of a verb. This is shown in sentence (50) in which the patient of the verb, *nggitu'o omanu* 'that chicken', is indexed thus.

[12] As these clitics do not refer exclusively to the grouping S/P (they can encode all of A/S/P) the name 'absolutive' is perhaps misleading. Because this name has been traditionally used to refer to this clitic set in Bungku-Tolaki languages (i.e. Mead 1998) and because no better terminological candidate presents itself, I follow this tradition.

(50) *Iamo* *sumbele-'i* *nggitu'o* *o-manu!*
NEG slaughter-3ABS that CN-chicken
'Don't kill that chicken!'

Note that only definite P's are indexed on the verb. When the P of a clause is indefinite, it is unindexed and the INDF.P prefix *poN-* occurs. An example is given in sentence (51):

(51) *Ku-po'-inu* *o-tee.*
1NOM-INDF.P-drink CN- tea
'I drank some tea.'

Absolutive clitics are also used to index the S/A argument of a verb in aspectually marked clauses.

6.1 Aspectually Marked Verbs

In Tolaki there is a set of enclitics that encode the aspect of the event described. These clitics are *-to* 'perfect', *-'ikaa* 'durative', *-ki* 'confirmative'[13] and *-po* 'irrealis'. Of these clitics *-to* and *-ki* are the most common. When a verb occurs with one of these clitics the S/A argument is indexed with an absolutive clitic. An S is shown indexed thus with a following *-to* and *-ki* in sentences (52) and (53) respectively.

(52) *Inaku* *pundi-nggu* *mate-**'i-to**!*
1SG banana.tree-1GEN die-3ABS-PRF
'[As for] me, my banana tree has died!'

(53) *Soso-**'i-ki**?*
suitable-3ABS-CONF
'Is it suitable?'

An A can likewise be indexed with an absolutive clitic. This is usually only possible when the P is indefinite and thus unindexed, as otherwise the absolutive clitic indexes the P (see example (50)). This is shown in sentences (54) and (55).

(54) *Mo-dapa-**keito-to*** *o-sala!*
NFIN.INDF.P-find-1IN.ABS-PRF CN-path
'We've found the path!'

(55) *Ndee* *monggaakoki* *sanggara?*
ndee moN-kaa-**ko-ki** sanggara
habitually NFIN.INDF.P-eat-2ABS-CONF fried.banana
'Do you usually eat fried banana?'

In all the examples shown so far, an aspectual clitic occurs. However, in the Tolaki text that Mead (1998:390–4) provides, there is one example of an absolutive A with no following aspectual clitic:

---

[13] The confirmative is used when the speaker asserts or checks whether a particular state of affairs (still) holds.

(56) *Mombehawe'i* *akala.*
moN-pehawe-**'i** akala
NFIN.INDF.P-think-3ABS plan
'He began to think of an idea.'

The absolutive clitic in this sentence cannot refer to the P *akala* 'tactic' as it is ungrammatical in Tolaki to index a P when the INDEF.P prefix *poN-* occurs (Edwards 2012:62).This is the only example I have so far encountered in which the S/A is indexed with an absolutive clitic without a following aspectual clitic.[14] Note, however, that this sentence still has an inchoative aspectual meaning.

When the verb is aspectually marked and the absolutive clitic slot is occupied by an indexed P, the S/A is indexed with a nominative clitic, as shown in (57).

(57) *Rombenggii'ito* *kolopua.*
ro-mbeN-kii-'i-to kolopua
3NSG.NOM-PL-see-3ABS-PRF tortoise
'They saw the tortoise.' (Untung 2009:31)

When an auxiliary verb, such as *ari* 'completive' or *laa* 'progressive' occurs, the clause contains two enclitic slots and the A can thus be indexed on the auxiliary, in addition to a definite P on the main verb. This is shown in sentence (58) below.

(58) *Ari-**ko**-to* *k‹um›aa-'i.*
COMPL-2ABS-PRF ‹NFIN›eat-3ABS
'Have you finished eating it yet?'

There is one exception to this rule; in the case of the auxiliary *onggo* 'future', nominative indexation is the norm. This is shown in sentences (59) and (60) below, which would otherwise fulfil the criteria for absolutive indexation of the S/A.

(59) ***Ku'**-onggo-ki* *po'ahi-'i-ko'o.*
1NOM-FUT-CONF carry.on.shoulder-3ABS-2DAT
'[don't worry] I'll carry it for you.' (Untung 2009:29)

(60) *U'onggopo* *monggaa,* *au* *leu* *ilaika.*
u-onggo-po moN-kaa a-u leu i-laika
2NOM-FUT-IRR NFIN.INDF.P-eat so-2NOM come LOC-house
'If you want to eat, come to [my] house.'

However, even with this auxiliary, absolutive indexation can still occur. This is shown in sentence (61). The factors motivating the difference in indexation are not clear.

(61) *Onggokopo* *monggaa,* *au* *leu* *ilaika,* *ieepo* *timba.*
Onggo-**ko**-po moN-kaa a-u leu i-laika iee-po timba
FUT-2ABS-IRR NFIN.INDF.P-eat so-2NOM come LOC-house 3SG-IRR enter
'You only come to [my] house if you want to eat.'
lit. 'You would eat, so you come to [my] house, that would [be when you] come.'

---

[14] With the exception of interrogative predicates, which I argue constitute a special case (see section 6.2).

6.2 Interrogative Predicates

Finally with the interrogative predicates *mbaako* 'why' and *hawo* 'what, what about' absolutive indexation of the S/A is the norm. This is illustrated in sentences (62) and (63).

(62) *Maa, hawoe kutaa onggo monduhako'o*
maa, hawo-**e** ku-taa onggo moN-tuha-ko'o
well what-3ABS 1NOM-NEG FUT NFIN.INDF.P-throw.down-2DAT
'Well, what about if I won't throw down any for you?' (Untung 2009:29)

(63) *Mbaako-**ko** Sandima?*
why-2ABS Sandima
'What's up with you, Sandima?' (Untung2011:64)

I analyse these interrogative predicates as lexically specifying for an absolutive S/A, similar to the way quirky case functions in Icelandic (see section 2.1).

7 Conclusion

7.1 'Markedness'

I argued in section 4.1, with corroborating evidence from sentences (35)–(37) in section 5.1, that the nominative clitics are synchronically the default, or unmarked, set for subject indexation in Tolaki; a development from an earlier stage in which absolutive clitics were probably the unmarked set (Mead 1998:332).[15]

If nominative clitics are the unmarked set for indexing a subject, then we expect to find they occur most frequently. This is indeed the case. In the story *Ohada ronga kolopua* 'The monkey and the tortoise' (Untung 2009:28–35) we find a total of 110 indexed subjects. Of these 61 are nominative, 35 genitive[16] and 14 absolutive.

Table 7.1: Indexation Strategies in *Ohada ronga kolopua*

Set	Construction	no.	total
NOM	Unmarked	14	61
	Conjunctions	44	
	Negated	3	
	Imperatives	0	
GEN	Experiencers	0	35
	Temporally Dependent	35	
ABS	Aspectually Marked	13	14
	Interrogative Predicates	1	

When we examine the different constructions which condition the different indexation strategies, however, a curious fact emerges. Of the 61 instances of a nominative subject, 44 occur after a conjunction and only 14 occur in an 'unmarked' environment. Furthermore, of the 14 instances of verbs with a nominative subject all except 2 also occur with an

[15] Mead (1998:325) notes that a similar development has occurred in the related language Kulisusu.

[16] Of the 35 genitive subjects, 17 are the semantically bleached *lakonoto* (see section 5.2), and 7 the non-singular equivalent *lakoroto.*

indexed object on the same verbal complex, and it would thus be impossible to index the subject with absolutive clitic. These facts are summarised in Table 7.1.

That the 'unmarked' strategy is used in only about 13% of cases calls into question the notion of '(un)markedness' as an appropriate label for the Tolaki phenomena. Instead 'default' is a better label; a kind of last resort strategy to be employed only when a subject occurs in a clause which does not fit one of six constructions.

7.2 Non-subjects and the notion of 'S'

A Tolaki subject can be indexed in one of three ways; with nominative, genitive or absolutive clitics. The constructions conditioning each strategy are summarised in Table 7.2.

Table 7.2: Indexation of the Subject in Tolaki

NOM	a. Unmarked b. Conjunctions c. Negated d. Imperatives
GEN	e. Experiencers f. Temporally Dependent
ABS	g. Aspectually Marked h. Interrogative Predicates

However, subjects are not the only roles which display a range of different indexation strategies. We also find there are two different indexation options for the P. Recall from section 6 that P's are normally indexed with absolutive clitics, as in sentence (64).

(64) *No-langgu-**'aku**.*
3NOM-hit-1ABS
'He hits me.'

While this is the normal pattern, there is a small class of verbs which indexes their P with dative clitics. An example is given in sentence(65).

(65) *U-to'ori-**kona**?*
2NOM-know-1DAT
'Do you know me?'

Verbs which index their P with dative clitics all have non-patient P's. However, dative clitics are not used to index all non-patient P's. Instead, it must be stipulated at the lexical level which verbs take a dative P.A sample of such verbs includes: *watu* 'go along with', *to'ori* 'know', *te'eni* 'say, tell', *teposua* 'meet' and *tealo* 'fetch, pick up'.The indexation of P's is given in Table 7.3.

Table 7.3: Indexation of P in Tolaki

ABS	α. Unmarked
DAT	β. Lexically Specified (*to'ori, watu,* etc.)

Figure 7.1 represents graphically the way the different indexation strategies in Table 7.2 and Table 7.3 map onto the notions of S, A and P. When a line is dashed, this indicates that the indexation strategy in question is used only in certain constructions.

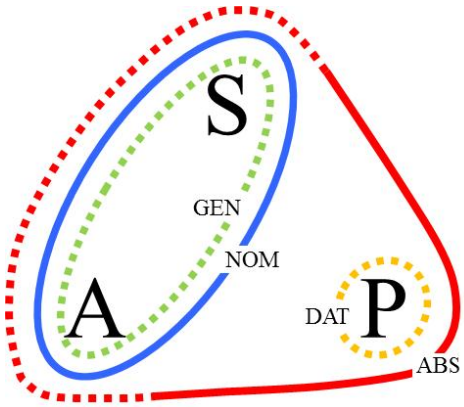


Figure 7.1: Indexation of S, A and P Tolaki

Figure 7.1 shows explicitly a fact about Tolaki morphology that has been implicit throughout this paper: in Tolaki A and S are encoded identically, while A and P are encoded differently.

Syntactically, as noted by Edwards (2012:87), we find a split in the behaviour of S and A. Under certain syntactic tests (plural agreement and relativisation) S and A have the same behaviour (Edwards 2012:71–72), while under certain other syntactic tests (internal relativisation and secondary predication) we find that S has the same behaviour as P (Edwards 2012:81–87).

Andrews (2001:108), on the basis of the behaviour of S in Icelandic, put forward the hypothesis that, "[…] S is not actually a primitive concept of syntactic structure." Instead, he proposes that A and P are primitives which function as "poles of attraction" for the encoding of arguments that are neither A nor P. The data from Tolaki, a language geographically distant from Icelandic and typologically different to it, corroborates this hypothesis; S displays behaviour identical with either A or P.

While in Icelandic both the syntactic *and* morphological behaviour of S displays a mix of properties in common with either A or P, in Tolaki, only the syntactic behaviour of S is split between A-like and P-like properties; morphologically the behaviour of S is completely identical to the behaviour of A.

References

Andrews, Avery. 1982. Case in Modern Icelandic. In *The mental representation of grammatical relations*, Joan Bresnan (ed.), 427–503. The MIT Press.

Andrews, Avery. 2001. Noncanonical A/S marking in Icelandic. In *Non-canonical marking of subjects and objects*, Alexandra Aikhenvald, Robert Dixon, and Masayuki Onishi(ed.), 85–111. John Benjamins.

Comrie, Bernard. 1978. Ergativity. In *Syntactic typology: Studies in the phenomenology of language*, ed. Winfred P. Lehmann, 329–394. University of Texas Press.

Comrie, Bernard. 2011. Alignment of case marking of full noun phrases. In *The world atlas of language structures online*, Matthew S. Dryer and Martin Haspelmath (ed.). Munich: Max Planck Digital Library. URL http://wals.info/chapter/98.

Danielsen, Swintha, and Tania Granadillo. 2008. Agreement in two Arawak languages: Baure and Kurripako. In *The typology of semantic alignment*, Mark Donohue and Søren Wichmann (ed.), 396–411. Oxford University Press.

Dixon, Robert. 1979. Ergativity. *Language* 55:59–138.

Dixon, Robert. 1994. *Ergativity*.

Donohue, Mark. 2008. Semantic alignment systems: what's what and what's not. In *The typology of semantic alignment*, Mark Donohue and Søren Wichmann (ed.), 24–75. Oxford University Press.

Edwards, Owen. 2012. Grammatical functions in Tolaki. Honours Thesis, Australian National University.

Mead, David. 1998. Proto-Bungku-Tolaki: Reconstruction of its phonology and aspects of its morphosyntax. Doctoral Dissertation, Rice University.

Mead, David. 1999. *The Bungku-Tolaki languages of South-Eastern Sulawesi, Indonesia*. Pacific Linguistics.

Mead, David. 2002. Proto Celebic focus revisited. In *The history and typology of western Austronesian voice systems*, Fay Wouk and Malcolm Ross (ed.), 143–177. Pacific Linguistics.

Mead, David, and Scott Youngman. 2008. Verb serialisation in Tolaki. In *Serial verb constructions in Austronesian and Papuan languages*, Gunter Senft (ed.),113–139. Pacific Linguistics.

Untung. 2009. *Cerita Rakyat Tolaki* [Tolaki Folklore]. Bintang.

Untung. 2011. *Mandara Ndolaki: Mombetolaki nggo sekolah dasar* [Tolaki Dexterity: Speaking Tolaki for Primary School]. Bintang.

8 *Event existentials in Tagalog: A Role and Reference Grammar account*

ANJA LATROUITE AND ROBERT D. VAN VALIN, JR.

1 Introduction[1]

This paper investigates an interesting type of existential construction in Tagalog and proposes an analysis of it in terms of Role and Reference Grammar. The construction in question is known as an 'event existential' and is exemplified in (1), and it contrasts with a 'plain' or 'nominal' existential construction in (2). Example (1) is from Aldridge (2011), who presents a Minimalist analysis of the construction; (2) is from Sabbagh (2009).

(1) *May* *b<in>ili-ng* *libro* *ang* *babae.*
EXIST <PERF.UV>buy-LNK book NOM woman
'The woman bought a book.'

(2) *May* *libro-ng* *b<in>ili* *ng* *babae.*
EXIST book-LNK <PERF.UV>buy GEN woman
'There is a book which a/the woman bought.'

A brief introduction to the morphological markers for voice and case of Tagalog which will play a role in the following discussion is in order. Tagalog clauses are typically predicate-initial, and a verbal predicate carries voice marking which indicates the semantic function of the nominative argument, as illustrated in (3)-(4).

(3) *B<in>ili* *ng* *babae* *ang* *libro.*
<PERF.UV>buy GEN woman NOM book[2]
'A/the woman bought the book.'

(4) *B<um>ili* *ang* *babae* *ng* *libro.*
<PERF.AV>buy NOM woman GEN book
'The woman bought a book.'

[1] This research has been supported in part by the Deutsche Forschungsgemeinschaft through SFB 991; the research of the second author has also been supported by a Fellowship from the Max Planck Society. We would like to thank Redemto Batul, Jeruen Dery, Reyal Panotes Palmero and Maureen Saclot for answering our queries about Tagalog, and Dery and Saclot for comments on an earlier version. We would also like to thank an anonymous reviewer for comments. Responsibility for errors and misinterpretations lies with us alone.

[2] Abbreviations: AV 'actor voice', CsP 'case particle', DAT 'dative', GEN 'genitive', LNK 'linker', NOM 'nominative', NUC 'nucleus', PERF 'perfective', PRED 'predicate', RP 'reference phrase', UV 'undergoer voice'.

The two argument case markers are *ang* 'nominative' and *ng* 'genitive' (pronounced [naŋ]), and any of the arguments of a verb like *bili* 'buy' can function as the nominative argument: in (3) the undergoer is nominative, and in (4) the actor is nominative.[3] The voice of the verb signals the semantic function of the *ang*-marked argument: undergoer voice in (3), actor voice in (4). Adjunct modifiers of argument expressions occur with what is called a linker, which is *ng* (pronounced [ŋ]) or *na*, depending on the phonological environment; it is distinct from the case marker *ng*. In (2), the relative clause modifying the noun *libro* 'book' is connected to it by the linker *ng*.

The construction in (1) has interesting properties. First, (1) and (2) differ semantically: (1) asserts the existence of an event, i.e. 'there was buying of a book by the woman', while (2) asserts the existence of an entity. Both may be used to introduce a new referent. This distinguishes (1) from (4), in which the genitive undergoer *libro* 'book' is normally non-referential. Second, the actor of *bili* 'buy' is nominative in (1), despite the verb having undergoer voice, in contrast to (2) with an externally-headed restrictive relative clause in which the actor of *bili* is genitive, as expected. If the actor in (1) were genitive, as in (5), this results in ungrammaticality, and if the actor were nominative in the nominal existential, as in (6), the result is likewise ungrammatical.

(5) **May b<in>ili-ng libro ng babae.*
EXIST <PERF.UV>buy-LNK book GEN woman
'The woman bought a book.'

(6) **May libro-ng b<in>ili ang babae.*[4]
EXIST book-LNK <PERF.UV>buy NOM woman
'There is a book which the woman bought.'

Third, the complement predicate in the event existential must be in undergoer voice, if transitive, whereas this is not true of the plain existential, as (7) and (8) show.

(7) **May b<um>ili-ng libro ang babae.*
EXIST <PERF.AV>buy-LNK book NOM woman
'The woman bought a book.'

(8) *May babae-ng b<um>ili ng libro.*
EXIST woman-LNK <PERF.AV>buy GEN book
'There is a woman who bought a book.'

Fourth, the undergoer of *bili* is not nominative, as its voice normally requires; rather, it is marked by the linker *na*/*ng*, a fact which is left unexplained in Aldridge (2011). Fifth, extraction is possible out of the unit headed by *bili* in (1) but not (2); (9) is an event existential (Aldridge 2011:2), while (10), a plain existential, is from Sabbagh (2009:698).

(9) *Saan may ni-luto-ng isda ang guro?*
where EXIST PERF.UV-cook-LNK fish NOM teacher
'Where was a fish cooked by the teacher?'

---

[3] Their semantic functions are usually described in terms of thematic roles; for a discussion, see Latrouite (2011).

[4] This sentence is grammatical with the meaning 'The woman has a book which was bought'; see §3.

(10) **Saan may bata-ng i<ni>ligtas ni Huan?*
where EXIST child-LNK <PERF.UV>rescue GEN Juan
'Where$_i$ is there a child which$_j$ Juan rescued $\_\__j$ $\_\_\__i$?'[5]

Thus, the event existential differs substantially morphosyntactically and semantically from the nominal existential, and these differences need to be accounted for.

The goal of this paper is to present a Role and Reference Grammar [RRG] (Foley & Van Valin 1984, Van Valin & LaPolla 1997, Van Valin 2005) analysis of the event existential in (1), which will account for the phenomena detailed above. The discussion will proceed as follows. In §2 a brief introduction to RRG will be given, and in §3 the account of event existentials in Tagalog will be presented. Conclusions follow in §4.

2 Role and Reference Grammar

RRG posits a single syntactic representation for a sentence, and it is concrete, not abstract, in the sense that it should represent the actual form of the sentence. No phonologically null elements are permitted in syntactic representations. There is a direct mapping between the semantic representation and the syntactic representation, unmediated by abstract syntactic representations. Discourse-pragmatics (information structure) may play a role in this mapping. This is sketched in Figure 2.1.

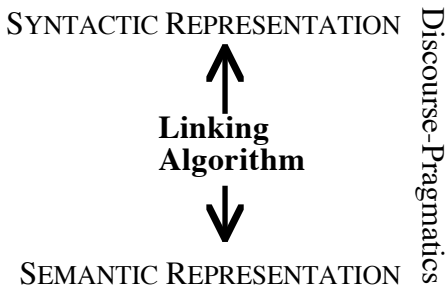


Figure 2.1: The organization of Role and Reference Grammar (Van Valin 2005)

Clauses and referring expressions, termed 'reference phrases' [RPs] in RRG, have a layered structure: the nucleus [NUC] of the clause contains the predicate and the nucleus$_R$ of the RP contains the head. The core of the clause contains the nucleus plus the arguments of the predicate (default), and the core$_R$ contains the nucleus$_R$ plus the arguments of relational nouns and deverbal nominals. Clause and sentence nodes dominate core, while RP dominates core$_R$. Clauses differ from cores in two important ways: first, they may contain optional constituents like a special position for displaced WH-expressions, and they can be modified by tense operators. Sentences are clauses plus optional positions for dislocated phrases. The layered constituent structure of the Tagalog sentence in (3) is given in Figure 2.2.

[5] The intended (impossible) meaning is a question about where the rescue took place, not about the location of the child. It is acceptable if it is a question about the location of the child who was rescued.

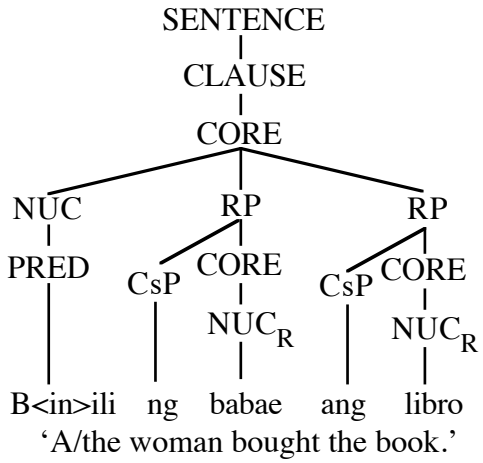


Figure 2.2: Aspects of the constituent structure of the Tagalog sentence in (3)

The case markers *ng* 'genitive' and *ang* 'nominative' occur before the head of the RP but are not prepositions (Himmelmann 2008); they are therefore represented as 'case particles' [CsP] in the RP structure.

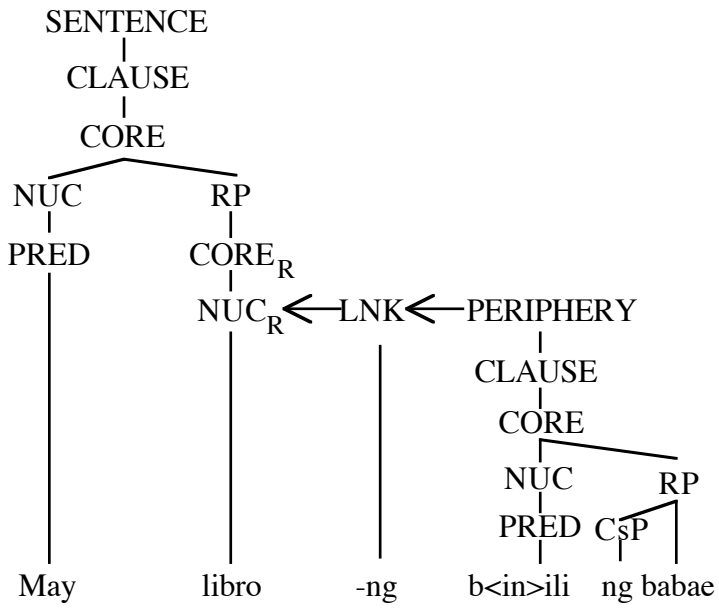
Syntactic categories need not be projections of lexical categories in RRG; the two most important syntactic categories, nucleus and reference phrase, are not projections of verb and noun, respectively (Van Valin 2008). This is particularly important for the analysis of Philippine languages, where there appear to be no constraints on which lexical categories can serve as predicate (nucleus) or argument (RP) (Himmelmann 2008). In (11), the nucleus is a lexical verb and the RP contains a noun, whereas in (12) the predicate in the nucleus is a noun and the RP contains a lexical verb (Schachter 1985).

(11) [$_{NUC}$ *Nag-trabaho*] [$_{RP}$ *ang* *lalaki*]
PERF.AV-work NOM man
'The man worked.'

(12) [$_{NUC}$ *Lalaki*] [$_{RP}$ *ang* *nag-trabaho*].
man NOM PERF.AV-work
'The one who worked is a man.'

Accordingly, the lexical category of nuclei and RPs will not be specified in the syntactic representations.

There is a periphery containing optional adjunct modifiers at each level of the clause and RP, and a linker occurs between each such modifier and the RP layer modified. Of particular relevance to this discussion is the fact that externally-headed restrictive relative clauses are modifiers in the nuclear periphery of the RP, as illustrated in Figure 2.3, which represents the structure of (2), the nominal existential. Note the lack of a case particle on the 'pivot' of the existential construction, *libro* 'book', in contrast to (3) and (4).



'There is a book which a/the woman bought.'

Figure 2.3: Structure of nominal existential in (2)

The semantic representations in RRG are decompositional in nature; the 'logical structure' of the examples in (3)-(4) is given in (13).

(13) **do´** (babae, [**buy´** (babae, libro)]) = (3)-(4)

(14) involves an attributive modifier, which will be relevant to the analysis of restrictive relative clauses as in (2).

(14) *B<um>ili* *ng* *mabuti-ng* *kotse* *si* *Pedro.*
<PERF.AV>buy GEN good-LNK car NOM
'Pedro bought a good car.' = **do´** (Pedro, [**buy´** (Pedro, [**be´** (<u>kotse</u>, [**good´**])])])

The abstract predicate **be´** is a marker of an attributive predication in the semantic representation and **good´** is the attribute. The word *kotse* 'car' is underlined to indicate that it functions both in the embedded attributive predication and simultaneously in the matrix predication; it is both the second argument of **buy´** and the bearer of the attribute **good´**.

RRG distinguishes lexical from constructional meaning and represents them both using the same system of lexical decomposition. In this paper we represent lexical meaning as in (13) and (14), using small boldface letters, e.g. **pred´**, whereas constructional meaning will be represented in small caps, e.g. PRED´. In a relative clause the attribute is a whole logical structure, and the interpretation of the clause as an attribute is constructional, and accordingly the marker of the attributive predication is BE´ instead of **be´** as in (14). This is illustrated in (15) with the English relative clause *the car John bought* in which the interpretation of *John bought* as an attribute in an attributive predication is entirely constructional. The first argument of BE´ is a participant in the logical structure functioning as the attribute.

(15) *Mary likes the car John bought.*
like´ (Mary, [BE´ (car$_i$, [**do´** (John, [**buy´** (John, x$_i$)])])])

The car is the head of the attributive predication and has the attribute 'John bought', which is indicated by the cointexed variable in the logical structure of *buy*. It is simultaneously the thing liked in the logical structure of *like*, and its dual role is indicated by the underlining: the logical structure in (15) is a combination of **like´** (Mary, car) (= *Mary likes the car*) and BE´ (car$_i$, [**do´** (John, [**buy´** (John, x$_i$)])]) (= *the car John bought*).

3 The event existential construction in Tagalog

The sentence in (1) is not only interpretable as an event existential; according to Naylor (2005:430), it can also be interpreted as a possessive construction with the meaning 'the woman has a bought book'(see also Schachter & Otanes 1972:279). Naylor gives the example in (16) of a basic possessive construction, which also has *may* 'exist' as the nucleus, and she also gives the example in (17) (Naylor 2005:430), which seems to contain the same ultimate constituents in the same order as in (1) and has an event existential reading.

(16) *May* *pera* *ang* *bata.*
EXIST money NOM child
'The child has money.'

(17) *May* *b<in>ili-ng* *laruan* *ang* *bata.*
EXIST <PERF.UV>buy-LNK toy NOM child
'The child has a bought toy,' or 'The child bought a toy.'

(18) *May* *laruan* *na* *b<in>ili* *ang* *bata.*
EXIST toy LNK <PERF.UV>buy NOM child
'The child has a bought toy.'

In (17), on a possession interpretation, *binili-ng laruan* 'toy which was bought' is an internally-headed relative clause, realizing the possessed argument.[6] It has the same meaning as (18), which has an externally-headed relative clause; (18) does not have an event existential reading, however. Not all possessive constructions like (17) have an event existential reading; the existence of such a reading depends on the plausibility of the event, as M. Saclot (pers. comm.) points out. The possessive construction in (19) is unlikely to have an event existential interpretation, due to the impossibility of a cat peeling a potato.[7]

(19) *May* *b<in>alata-ng* *patatas* *ang* *pusa.*
EXIST <PERF.UV>peel-LNK potato NOM cat
'The cat has a peeled potato.'/#'The cat peeled the potato.'

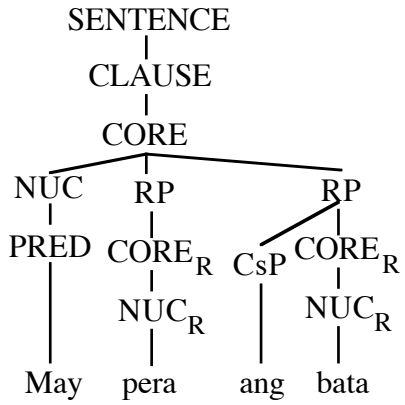
In our view, the superficial identity between the possessive construction in (17) and the event existential in (1) is the key to explaining the properties of the event existential con-

[6] As pointed out by an anonymous reviewer, (17) is, strictly speaking, ambiguous between a structure with an internally-headed relative clause and one with a pre-head externally-headed relative clause. Placement of adjuncts such as temporal adverbials, e.g. in (22), which is unambiguously internally headed, can resolve the ambiguity. Crucial for our analysis is the fact that (17) has as one of its structural analyses a possessive construction containing an internally-headed relative clause.

[7] One of our consultants did in fact get an eventive reading for (19), which he found humorous.

struction given in §1. Our hypothesis, then, is that event existentials are related to possessive constructions such as the one in (17). This involves both the semantic and syntactic properties of the two constructions, and we will show that sentences like (1) and (17) are structurally ambiguous, with the same string of words having two different structures, one corresponding to the event existential interpretation and one corresponding to the possessive interpretation. Moreover, event existentials developed via a diachronic reanalysis from possessive constructions containing an internally-headed relative clause into the event existential construction.

The starting point for this account is the semantic and syntactic representations of possessive constructions such as (16). The semantic representation for (16) is **exist´ (**pera) ∧ HAVE´ (bata, pera); *pera* 'money' is simultaneously the entity of which existence is predicated (the first proposition, **exist´** (pera)) and the possessed entity in the second proposition (HAVE´ (bata, pera)). There is no lexical verb of possession, unlike in English and many other languages; hence the interpretation of *bata* 'child' as the possessor is purely constructional, and therefore the possessive proposition is represented by HAVE´ rather than **have´**. The syntactic structure for (16) is a simple predication, as in Figure 3.1.



'The child has money.'

Figure 3.1: Structure of simple Tagalog possessive predication in (16)

The semantic representation for (17) is more complex, as it includes an attributive predication which is realized as an internally-headed relative clause.

(20) Attributive semantic representation: BE´ (x, [**attribute´**], where **attribute´** is a full predication in the case of a relative clause (see (15)).

(21) **exist´** (laruan$_i$) ∧ HAVE´ (bata, [BE´ (x$_i$, [**do´** (y, [**buy´** (y, $\underline{\text{laruan}_i}$)])])])

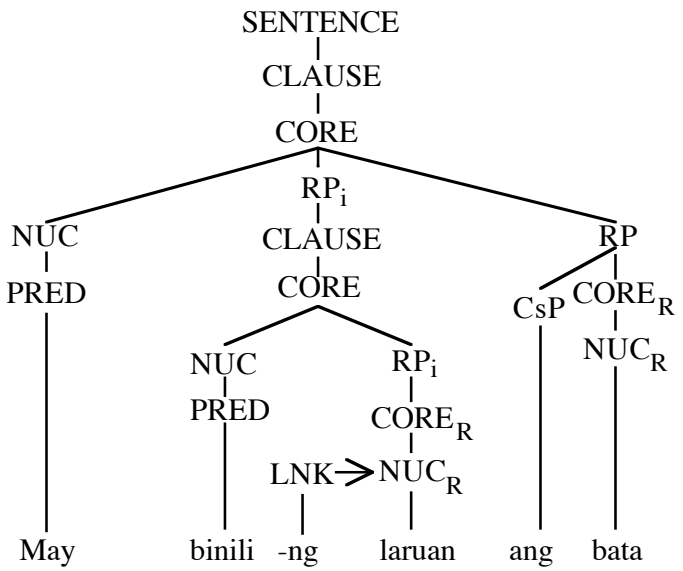
The basic possessive predication in (17) is represented the same way as for (16), i.e. **exist´ (**laruan) ∧ HAVE´ (bata, laruan). In order to represent the predication in the internally-headed relative clause in (17), an attributive predication containing *laruan* replaces it in the possessive representation; [BE´ (x$_i$, [**do´** (y, [**buy´** (y, laruan$_i$)])]) means 'toy which *y* bought', where it is an attribute of *x* that *y* bought *x*, *x* being *laruan*.[8] In the final represen-

[8] In the semantic representation for an externally-headed relative clause, the external *x* would be lexically filled and the corresponding variable in the embedded proposition would be coindexed and left lexically unfilled, as in (15). Thus the semantic representation for (18) would be:

(i) **exist´** (laruan) ∧ HAVE´ (bata, [BE´ ($\underline{\text{laruan}_i}$, [**do´** (y, [**buy´** (y, x$_i$)])])])]).

tation in (21), *laruan* functions as an argument in four predications: the second argument of **buy′**, the first argument of BE′, the second argument of HAVE′ **(**which is signaled by the underlining), and the single argument of **exist′**.

The syntactic structure of (17) is given in Figure 3.2. The head noun *laruan* 'toy' occurs inside the relative clause, and it is coindexed with the RP node dominating the internally-headed relative clause, signaling that it functions in both the main and embedded clauses, following Van Valin & LaPolla (1997). The linker *–ng* signals that *laruan* 'toy' is the modified noun within the relative clause. This is related to its function of connecting adjunct modifiers to the element modified, as in (2) and (18), but there is no adjunct modifier in (17); rather, *laruan* 'toy' is internal to the modifying clause, and the linker indicates that it is the modified element.



'The child has a toy which was bought.'

Figure 3.2: Structure of possessive construction with internally-headed relative clause in (17)

That it is a full clause can be seen in the possibility of the main and relative clauses having conflicting temporal adverbs, as shown in (22); (23) gives the externally-headed relative clause version. (Both examples are from J. Dery, pers. comm.)

(22) *Ngayon* *may* [*b<in>ili-ng* *laruan* *kahapon*] *ang* *bata.*
today EXIST <PERF.UV>buy-LNK toy yesterday NOM child
'Today the child has a toy that was bought yesterday'

(23) *Ngayon* *may* *laruan* *na* [*b<in>ili* *kahapon*] *ang* *bata.*
today EXIST toy LNK <PERF.UV>buy yesterday NOM child
'Today the child has a toy that was bought yesterday.'

The meaning of (21) and (i) is the same propositionally but not necessarily information structurally. The difference reflects a technical point in the RRG linking theory, namely, that an argument or adjunct must be represented in the semantic representation of the clause in which it appears (except for extraction phenomena). Since *laruan* 'toy' appears in the relative clause in (17), it must be linked from the semantic representation of the relative clause, whereas in (18) it appears in the main clause, and therefore it must be linked from the semantic representation of the main clause. See Van Valin (2012) for detailed discussion of linking in relative clauses.

An important property of this construction is that the possessor need not be interpreted as the actor of the relative clause predicate. In the previous examples there has been no explicit actor in the relative clause, either internally-headed, as in (17) or externally-headed, as in (18); in the following pairs of examples, there is an explicit actor in both types of relative clauses: internally-headed in (24) and (26) (Law 2010:309, 315) and externally-headed in (25) and (27) (J. Dery, pers. comm.).

(24) *May* *ni-luto-ng* *isda* *ni* *Pedro* *ang* *guro.*
EXIST PERF.UV-cook-LNK fish GEN NOM teacher
'The teacher has fish that was cooked by Pedro.'

(25) *May* *isda-ng* *ni-luto* *ni* *Pedro* *ang* *guro.*
EXIST fish-LNK PERF.UV-cook GEN NOM teacher
'The teacher has fish that was cooked by Pedro.'

(26) *May* *g<in>awa* *ng* *pamahalaan* *na* *bahay* *ang* *guro.*
EXIST <PERF.UV>build GEN government LNK house NOM teacher
'The teacher has a house that was built by the government.'

(27) *May* *bahay* *ang* *guro* *na* *g<in>awa* *ng* *pamahalaan.*
EXIST house NOM teacher LNK <PERF.UV>build GEN government
'The teacher has a house that was built by the government.'

The transition from a possessive construction to an event evidential can be described in terms of a series of changes:

1. The possessor is analyzed as the actor of the embedded predicate. This creates a control relationship between the possessor (*bata* 'child') and the embedded predicate *bili* 'buy', as in (28), in which the control relationship is indicated by the coindexing of *bata* with the actor argument (*y*) of **do´** (y, [**buy´** (y, z)]).
2. The establishment of a control relationship affects the constructions and their meanings:
   a. As part of a restrictive relative clause, the embedded proposition is in a clause and is presupposed; control relations in RRG can only exist across core boundaries, and consequently the embedded unit must be reanalyzed as a core and is no longer necessarily presupposed. Therefore the embedded proposition is no longer a restrictive modifier, and the attributive constructional meaning is lost. The resulting syntactic structure is given in Figure 3.3 below.
   b. The interpretation of the erstwhile possessor as the actor of the embedded predicate strongly disfavors the constructional possession interpretation, leaving the RP stranded in the matrix core, and as a result it is constructionally interpreted as being the doer of an unspecified action. This loss of the possession reading is caused by the incompatibility of obligatory control relations and possessive predications; possession predicates do not take propositional complements, just RP complements, and obligatory control relations are always between two propositions which obligatorily share an argument. Hence the establishment of an obligatory control relationship between the erstwhile possessor and the actor of the embedded predicate is incompatible with the interpretation of '*may ... ang RP'* as a possession construction, restricting *may* to a purely existential reading and causing the *ang RP* to be interpreted as the actor of an unspecified action, as given in (29).

3. The scope of the existential predicate extends to include the entire event description, yielding an event existential: (30)

(28) **exist´** (laruan) ∧ HAVE´ (bata, laruan) ∧ [BE´ (x_i, [**do´** (y, [**buy´** (y, $laruan_i$)])] =>

exist´ (laruan) ∧ HAVE´ ($bata_j$, laruan) ∧ [BE´ (x_i, [**do´** (y_j, [**buy´** (y_j, $laruan_i$)])]

(29) **exist´** (laruan) ∧ DO´ ($bata_j$, Ø) ∧ **do´** (y_j, [**buy´** (y_j, laruan)])

(30) **exist´** (laruan ∧ DO´ ($bata_j$, Ø) ∧ **do´** (y_j, [**buy´** (y_j, laruan)]))

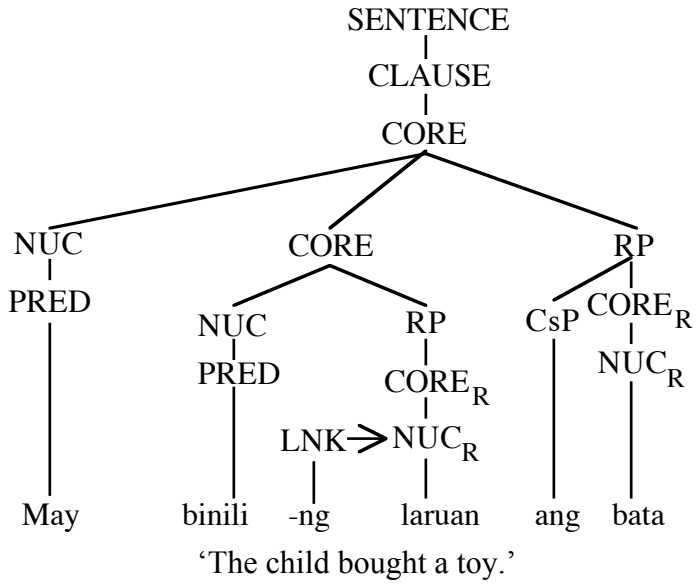


Figure 3.3: The structure of the event existential reading of (17)

The linker is a remnant of the structure in Figure 3.2, but it no longer marks that *laruan* 'toy' is modified.

Evidence that the embedded unit is a core rather than a clause as in Figure 3.2 comes from the impossibility of there being conflicting temporal adverbs in this construction, as shown in (31) and (32), from Aldridge (2011), in contrast to (22) and (23).[9]

(31) *May* [*is<in>ulat-ng* *love letter* *kahapon*] *ang* *babae.*
EXIST <PERF.UV>write-LNK yesterday NOM woman
'The woman wrote a love letter yesterday.'

(32) **Ngayon* *may* [*is<in>ulat-ng* *love letter* *kahapon*] *ang* *babae.*
today EXIST <PERF.UV>write-LNK yesterday NOM woman
'Today, the woman wrote a love letter yesterday.'

Thus, we have arrived at a semantic and syntactic account of event existentials, starting from possessive constructions which look superficially the same but are structurally and semantically distinct.

[9] If these examples are interpreted as possessive constructions, then both are acceptable. The first would mean 'He has a love letter (which was) written yesterday' (M. Saclot, pers. comm.), and the second would be 'Today the woman has a love letter (which was) written yesterday'.

4 Conclusion

The crucial question facing the above analysis is, how well does it account for the distinctive properties of event existential constructions mentioned in §1? Each of the points raised in §1 will be addressed below.

The event existential interpretation follows from the reanalysis of the (constructional) possessive semantic representation in (28)-(30). Because much of the meaning is constructional, syntactic reanalysis has profound semantic consequences. Moreover, the semantic representation in (30) accounts both for the assertion of the existence of an event and also the existence of an entity, the undergoer of the embedded verb, yielding the required specific interpretation. Given the ambiguity inherent in (1) and (17), it is important to not only account for the event existential reading but to relate the two interpretations to each other.

The actor in this construction, *babae* 'woman' in (1) and *bata* 'child' in (17), is marked by *ang*, because it is an argument in the matrix core (see Figure 3.3) and therefore its case is not determined by the voice of the embedded predicate.

That the predicate in the embedded core must be in undergoer voice is due to the reanalysis of the possessor as the actor of the embedded predicate. Kroeger (1993) argues that the controllee in a control construction in Tagalog must be the actor argument, regardless of the voice of the verb,[10] and this means that the only role open to the other argument of the verb is undergoer. Since the construction derives from a restrictive relative clause and the modified RP is the trigger for the voice of the verb, this means that in the event existential the verb must always be in undergoer voice, if transitive. This predicts that in the possessive construction, in which the possessor is not necessarily the actor of the embedded verb, actor voice should be possible, and this is in fact the case, as (33) shows (from M. Saclot, personal communication).

(33) *May* *t<um>akas* *na* *bilanggo* *ang* *sundalo*.
EXIST <PERF.AV>escape LNK prisoner NOM soldier
'The soldier has an escaped prisoner'/'The soldier has a prisoner that escaped.'

Hence the restriction to undergoer voice follows from the origin of the construction in a restrictive relative clause plus the reanalysis of the possessor as the actor of the embedded verb.

The undergoer of *bili* is marked not with the expected nominative, based on the voice of the embedded predicate, but rather by the linker *na*/*ng*, which does not normally function as a case marker but rather links modifiers to heads normally. This reflects in part the relationship of the event existential to the possessive construction, since in the possessive construction the modified noun in the relative clause is signaled by the linker. The embedded core in Figure 3.3 serves as the pivot of the existential construction, and therefore *ang* is ruled out because of the Definiteness Effect (Sabbagh 2009). This is reinforced by the semantic representation in (30), in which *laruan* 'toy' is the argument of **exist´**. The case marker *ng* is incompatible with the voice marking on the predicate. The occurrence of the linker is a remnant of the origin of the embedded core as an internally-headed relative clause, and it cannot be replaced by a true case marker due to the incompatibilities given above.

Extraction is predicted to be possible out of the embedded core by the RRG theory of extraction constraints (Van Valin 1995, 2005). These restrictions apply to embedded clauses, not to embedded cores. The structure in Figure 3.3 therefore satisfies the con-

[10] There are some exceptional circumstances in which this is not the case, as Kroeger points out, but none of these are relevant to this construction.

straints governing extraction in complex sentences, while the structure in (2), which is the same as (34), does not satisfy these conditions.

(34) *May* [RP *libro-ng* [CLAUSE *ib<in>igay* *ng* *babae* *sa* *lalaki*]].
EXIST book-LNK <PERF.UV>give GEN woman DAT man
'There is a book which a/the woman gave to the man.'

(35) **Sa* *lalaki* *may* [RP *libro-ng* [CLAUSE *ib<in>igay* *ng* *babae*]].
DAT man EXIST book-LNK <PERF.UV>give GEN woman
*'To the man there is a book that a/the woman gave.'

(36) *May* [CORE *ib<in>igay* *na* *libro* *sa* *lalaki*] *ang* *babae*.
EXIST <PERF.UV>give LNK book DAT man NOM woman
'The woman gave a book to the man.' ('There was the woman's giving a book to the man.')

(37) *Sa* *lalaki* *may* [CORE *ib<in>igay* *na* *libro*] *ang* *babae*.
DAT man EXIST <PERF.UV>give LNK book NOM woman
'To the man, the woman gave a book.' ('To the man, there was the woman's giving a book'.)

Thus, the RRG analysis accounts for the distinctive attributes of the event existential construction specified in §1, and the key to the explanation is the relationship of the event existential to the possessive construction, both syntactically and semantically. We have hypothesized that the event existential derives from or arose out of a reanalysis of the possessive construction with a restrictive relative clause, and this seems to be true in a diachronic sense: the event existential appears to be an innovation that not all speakers have. Indeed, one of the four native speakers with whom we consulted does not accept the event existential readings for these forms, while another treats them as secondary and as inferences, and one somewhat surprisingly does not get the possession interpretation (see Appendix).

Appendix on Native Speaker Judgments

We have argued that the event existential developed from the possessive *may*-construction. Interestingly, the diverging grammaticality judgments of our consultants (one who left the Philippines a long time ago and presumably shows a more conservative pattern, and one who just recently left the Philippines and may exhibit innovative structures) seem to support this idea.

Table 1: Consultant judgments

	Interpretation of sentence a. May N_{LK} V ANG N b. May V_{LK} N ANG N	Allows for extraction	Allows for double Actors
Consultant 1	a. possessive b. possessive	a. no b. no	(a) yes **(b) no**
Consultants 2 and 3	a. eventive b. eventive	a. yes b. yes	(a) no **(b)no**
Consultant 4	a. possessive b. possessive/eventive	a. no b. yes	(a) yes (b) no

Consultant 1 only accepts *may*-sentences (possessive and eventive) as answers to questions regarding possession, while Consultants 2 and 3 treat both, eventive and possessive existentials, as answers to questions regarding someone's actions. Consultant 4 strongly favors the possessive reading for all of these constructions, but acknowledges the eventive interpretations as a possible inference in some cases.

Their judgments with respect to extraction are in line with their respective interpretations of the sentence structures: Consultant 1 does not accept, while Consultants 2 and 3 do accept, extractions out of both sentences, as would be expected according to our analysis. Consultant 4 accepts extraction out of the structures with an eventive reading.

Their judgments with respect to double actor-sentences are also in line with their respective interpretations and at the same time show one common feature. Consultant 1 accepts double actors (i.e. a possessor and an actor), but only in sentences that exhibit possessive existential word order. The reasons given for this are two-fold: a) the *ni*-Actor is only unambiguously interpreted as an Actor if it follows the verb, otherwise it is interpreted as the Possessor of the object (therefore noun-verb-Actor order is easier to parse in the intended sense) b) longer restrictive relative clauses are hard to parse if they are realized internally (so they get extraposed resulting in the structure '*May* RP *ang* RP *na* RC'). Consultants 2 and 3 do not accept double Actor sentences. Consultant 4 accepts some double actor sentences.

References

Aldridge, Edith, 2011, Event existentials in Tagalog. Presented at 18th AFLA Conference.

Foley, William and Robert Van Valin, Jr., 1984, Functional syntax and universal grammar. Cambridge: Cambridge University Press.

Himmelmann, Nikolaus, 2008, Lexical categories and voice in Tagalog. In P. Austin and S. Musgrave, eds., Voice and grammatical relations in Austronesian languages, 247-93. Stanford: CSLI.

Kroeger, Paul, 1993, Phrase structure and grammatical relations in Tagalog. Stanford: CSLI.

Latrouite, Anja, 2011, Voice and case in Tagalog. Dissertation, Heinrich Heine University.

Law, P., 2010, The impersonal construction in Tagalog. In R. Mercado, E. Potsdam, and . Travis, eds., Austronesian and theoretical linguistics, 297-325. Amsterdam: John Benjamins.

Naylor, Paz B,. 2005, On the stative predicate: Tagalog "existentials" revisited. Hsiu-chuan Liao and Carl R. Galvez Rubino, eds. Current issues in Philippine linguistics and anthropology: Parangal kay Lawrence A. Reid, 419-435. Manila: LSP and SIL.

Sabbagh, Joseph, 2009, Existential sentences in Tagalog. Natural language and linguistic theory 27:675-719.

Schachter, Paul, 1985, Parts-of-speech systems. In T. Shopen, ed., Language, typology and syntactic description, vol. 1, 3-61. Cambridge: Cambridge University Press.

Schachter, Paul and Fe Otanes, 1972, Tagalog reference grammar. Berkeley: UC Press.

Van Valin, Robert, Jr., 1995, Toward a functionalist account of so-called 'extraction constraints'. In B. Divriendt et al., eds., Complex structures: A functionalist perspective, 29-60. Berlin: Mouton de Gruyter.

Van Valin, Robert, Jr., 2005, Exploring the syntax-semantics interface. Cambridge: Cambridge University Press.

Van Valin, Robert, Jr., 2008, RPs and the nature of lexical and syntactic categories in Role and Reference Grammar. In R. Van Valin, ed., Investigations of the syntax-semantics-pragmatics interface, 161-78. Amsterdam/Philadelphia: John Benjamins.

Van Valin, Robert, Jr., 2012, Some issues in the linking between syntax and semantics in relative clauses. In B. Comrie and Z. Estrada Fernández, eds., A typological overview of relative clauses in languages of the Americas, 47-64. Amsterdam: John Benjamins.

Van Valin, Robert, Jr., and Randy LaPolla, 1997, Syntax: structure, meaning and function. Cambridge: Cambridge University Press.

9 Neutral and Imperfective Forms in Kanakanavu

DORINDA TSAI-HSIU LIU

1 Introduction

This paper aims at exploring how the tense/aspect information is grammatically realized in the neutral and imperfective forms of Kanakanavu[1], an Austronesian language spoken in the southern mountainous area of Taiwan. Kanakanavu belongs to the Tsouic subgroup[2], one of the ten primary branches that split directly from the Proto-Austronesian (PAN) language[3] (Blust 1999). Among the aboriginal languages of Taiwan (commonly known as Formosan languages), Kanakanavu is one of the three most critically endangered languages, being nearly extinct with less than ten fluent native speakers left today. There exist only a few studies on Kanakanavu (Tsuchida 1973, 1976; Mei 1982; Chang 2006; and Wu 2006). Tsuchida (1973, 1976) and Mei (1982) provide some account for the tense and aspect system of Kanakanavu though their works did not concentrate on the Kanakanavu temporal/aspectual issue. The Kanakanavu data in this study are derived from the existing literature and my fieldwork[4] in 2010–2011. The previous studies will be reviewed where relevant to my discussion. The goals of this study are two-fold. First, this study is designed to serve as a base for other further synchronic and diachronic studies related to the Kanakanavu temporal/aspectual system due to the scarcity of the existing literature on this moribund language. Second, this study may shed some light on linguistic typological studies by presenting some language-specific features in the grammatical realizations of tense/aspect information in Kanakanavu focus system.

Kanakanavu manifests some language-specific properties in its temporal/aspectual

[1] This research is supported by the project 'The Most Critically Endangered Formosan Languages' (NSC 97-2410-H-001-065-MY3), funded by the National Science Council of Taiwan. I would like to thank the project director Prof. Paul Jen-kuei Li for his support and guidance on this Kanakanavu study.

[2] According to Blust (1999), there are three languages in the Tsouic subgroup: Tsou, Kanakanavu, and Saaroa.

[3] Based on some syntactic evidence, Chang (2006) proposes that Tsou alone instead of the so-called three Tsouic languages (i.e. Tsou, Kanakanavu, and Saaroa) might constitute a subgroup directly split from Proto-Austronesian language.

[4] The Kanakanavu people mainly distribute over three villages of Namasia (那瑪夏) District in the northeastern part of Kaohsiung City—Nansalu (南沙魯), Maya (瑪雅), and Takanua (達卡努瓦). My language consultants are all from Takanua. I would like to show my heart-felt thanks for my language consultants, includingKa'angaina Mu'u (Chinese name: 翁坤), Ka'angaina Paic~~u~~ (Chinese name: 范秀香), and Ka'angaina Angai (Chinese name: 翁博學).

Asia-Pacific Linguistics, 2014.

expressions, which are distinct from how the non-future (also known as 'realis' in the literature) information is grammatically realized in other Formosan languages. The temporal frame in most Formosan languages is generally understood to demonstrate a two-way contrast of realis vs. irrealis, e.g. Ogawa & Asai 1935; Tsuchida 1976; Huang 1995; Zeitoun 1996, 2007; Zeitoun et al. 1996; among many others.[5] To be specific, the non-future information is normally reflected in the verbal paradigms of focus markings in most Formosan languages when no temporal expression is available, whereas the future (also known as 'irrealis' in the literature) information is morphologically or lexically marked in these languages. A verb affixed with a focus marker is traditionally called a "neutral" form in the Formosan literature (e.g. Zeitoun et al. 1996). Unlike other Formosan languages, the occurrences of the neutral verbs are rather limited and less frequent. In contrast to the neutral verbs, the morphologically marked verbs (known as the "imperfective" forms in the literature) occur more frequently and have a wider distribution. This paper will elaborate on how the tense/aspect information is grammatically realized in the Kanakanavu focus markings as well as how the Kanakanavu aspectual expressions are integrated with the neutral/imperfective forms.

This paper is organized as follows. Section 2 gives some background information about Kanakanavu and some grammatical information relevant to my discussion. Section 3 introduces the Kanakanavu focus system, as the focus markings in Kanakanavu carry crucial tense and/or aspect information. Section 4 explores how the aspect information is overtly indicated by the grammatical markers in Kanakanavu and how the Kanakanavu neutral/imperfective forms interact with these aspectual markers. The concluding remarks of this paper are given in Section 5.

2 Background knowledge

This section provides a brief sketch of Kanakanavu grammar relevant to my discussion of the Kanakanavu temporal and aspectual system. Like most other Formosan languages, Kanakanavu is a predicate-initial language. The sentences normally begin with a nominal predicate or a verbal element, such as *kanakanavu* 'Kanakanavu' in (1a) and *t<um>aŋi* 'cry (AF[6])'in (1b) respectively.

(1) a. ***kanakanavu*** *iikasu.*
Kanakanavu 2S.NOM
'<u>You</u>[7] are a Kanakanavu.'

---

[5] The two terms 'realis' and 'irrealis' are normally involved with the markings of grammatical mood. The realis mood, by definition, is used for speakers to indicate that something is actually the case (or not the case). Meanwhile, the irrealis mood indicates that a certain situation or action is not known to have happened as the speaker is talking. Instead of marking the dichotomy of realis/irrealis mood, the two terms realis and irrealis are mostly used to distinguish something that has happened or is happening (realis) and something that has not happened (irrealis) in the Formosan literature. To avoid possible confusion, this study adopts 'non-future' and 'future' to replace 'realis' and 'non-realis' respectively.

[6] Abbreviations used in this paper are as follows: 1, first person; 2, second person; 3, third person; AF, actor focus; BF, beneficiary focus; CAUS, causative; COS, change of state; DEL, delimitative; FUT, future tense; GEN, genitive; IF, instrument focus; IPFV, imperfective; LF, location focus; LOC, location; NOM, nominative; OBL, oblique; P, plural; PE, plural exclusive; PERF, perfective; PF, patient focus; PI, plural inclusive; PROG, progressive; RED, reduplication; S, singular; and TOP, topic.

[7] The underline here is used to indicate the focused NP *iikasu* 'you (NOM)' of the clause in the English translation. This is because there is no direct correspondence in English to express the most prominent argument (i.e. focused NP) of the clause in Formosan languages.

b. ***t<um>aŋi**-cu* *sua* *maanu* *iisi.*
cry<AF>-COS NOM child this
'This child cried.'

In the dyadic constructions, the actor focus (AF) and non-actor focus (NAF) sentences have different word orders. Consider:

(2) a. *AF sentence: Verb-Actor-Patient*
l<~~u~~m>araisi maanu tanali.
bite<AF> child peanut
'This child bit/bites peanuts.'

b. *AF sentence: Verb-Patient-Actor*
l<~~u~~m>araisi tanali maanu.
bite<AF> peanut child
'This child bit/bites peanuts.'

The word order of the arguments in an AF construction is flexible, allowing Verb (AF)-Actor-Patient as well as Verb (AF)-Patient-Actor, as shown in (2). But the preferred word order for an AF construction is Verb (AF)-Actor-Patient like (2a). In contrast, the NAF sentences have a fixed word order. Normally the NAF verbs are directly followed by an actor, as the PF sentences in (3) illustrate.

(3) a. *tia* *aliapiniŋ-in* *maanu* *tuaŋiŋi* *iisi.*
IRR hook-PF.IPFV child mouse this
'A child will hook this mouse.'

b. **tia* *aliapiniŋ-in* *tuaŋiŋi* *iisi* *maanu.*
IRR hook-PF.IPFV mouse this child

A Kanakanavu verbal complex can be decomposed into a stem verb, a focus affix, a post-verbal bound personal pronoun, a causative marker, and/or a temporal/aspectual marker.

(4) a. *c<in>m>~~u~~ʔ~~u~~ra-ku* *ʔavia.*
see<PERF>AF>-1S.NOM 'avia[8]
'I have seen 'avia.'

b. *apa-tupun-un-aku* *maanu* *tacau* *ʔurupaca* *vatu.*
CAUS-throw-PF.IPFV-1S.GEN child dog AF.use stone
'I (will) make the child throw a stone at a dog.'

As shown in the examples (4a) and (4b), both verb complexes are composed of a verb stem, a focus marker, and a bound personal suffix. The difference is that the verb complex in (4a) has a perfective marker <*in*>, and the verb complex in (4b) has a causative marker *apa-*.

[8] The romanization for the male name *ʔavia* is *'avia*, in which the glottal stop is represented by a modified letter apostrophe (').

3 Focus system

This section discusses the Kanakanavu focus system in detail, as the focus system is indicative of the tense, aspect, and mood (TAM) information when the temporal information is not expressed by a TAM marker or by a time adverb.

Like other Formosan and Philippine languages, Kanakanavu has a set of verbal affixes, which permits a range of arguments to serve as the syntactically/pragmatically prominent noun phrase (NP) of the clause. The prominent NP can bear any of the following semantic roles: actor, patient (or theme), location, and instrument. In the literature, the prominent NP has been called focused NP, trigger, pivot, topic, and so on. In the previous Kanakanavu studies, the verbal morphology has been treated as focus marking (e.g. Tsuchida 1973, 1976; Mei 1982; and Chang 2006) or voice inflection (e.g. Wu 2006). Along with the traditional line, the prominent NP is referred to as "focus" in this study; and the verbal morphology encodes the thematic role of the focus NP.

In Kanakanavu, the focus system of the verbal morphology is composed of four types: Actor Focus (AF), Patient Focus (PF)[9], Location Focus (LF)[10], and Instrument Focus (IF)[11], as shown in the following table:

Table 3.1: Kanakanavu focus system (based onTsuchida1973, 1976; Mei 1982; Chang 2006; and Wu 2006).

Focus	Actor Focus (AF)	Patient Focus (PF)	Location Focus(LF)	Instrument Focus (IF)
Focus marker(s) for neutral forms	<um> (variants: mu-, um-, and <ʉm>), m-,Ø	-ai	-a(n), ta-...-a(n)	si-, sia-

When a verb is attached only with a focus marker, such a verb is referred to as a "neutral" form in the Kanakanavu literature (Tsuchida 1976 and Mei 1982). Some examples are given below:

(5) a. AF sentence:
*k<**um**>aʉn (sua) ŋiau tapianaŋai.*
eat<AF> NOM cat bird
'A catate/eats a bird.'

---

[9] The term "Patient Focus" is mainly referred to as "Object Focus" in the previous studies (i. e. Tsuchida 1976 and Mei 1982). Note that the usage of "Object Focus" may have an implication that the language in question is an accusative language. However, this grammatical analysis has not come to that far to conclude that Kanakanavu is an accusative language or an ergative language or any other. In order to avoid possible confusion, we adopt the neutral term "Patient Focus" instead.

[10] In a LF sentence, the focused NP refers to a location, a patient perceived as a spot, a key point of the event, and so on.

[11] In other Formosan languages (e.g. Mayrinax Atayal), the focus marker *si-* attached to a verb is used to indicate a focused instrument or beneficiary. However, my data corpus shows that the verbs prefixed with *si(a)-* in Kanakanavu refer to focused instruments in almost all cases with one suspicious exception *siacʉʔʉra* 'see (BF.IPFV)', as in (i).

(i) ? ***sia**-cʉʔʉra maku naamu (sua) maanu maku.*
BF-IPFV-see 1S.GEN grandchild NOM child 1S.GEN
'I looked/(often) look after (my) grandchild for my child.'

However, one of the three language consultants indicates that he does not accept the beneficiary usage of *siacʉʔʉra*. It is questionable to treat it as the example of beneficiary focus (BF). Thus, my Kanakanavu focus system does not include BF, as in other Formosan languages (e.g. Atayal and Tsou).

b. PF sentence:
*riucuucu-**ai*** *cina* *(sua)* *maanu* *iisi.*
kiss-PF mother NOM child this
'Mother kissed/kisses this child.'

c. LF sentence:
*riucuucu-**an*** *muʔu* *(sua)* *paicʉ.*
kiss-LF Mu'u NOM PaicU[12]
'Mu'u kissed/kisses PaicU.'

d. IF sentence:
***si**-patupun* *maanu* *tacau* *(sua)* *vatu* *iisi.*
IF-throw child dog NOM stone this
'A child threw/throws this stone at a dog.'

The focused NP is optionally preceded by a nominative marker *sua*, such as the actor *ŋiau* 'cat' in the AF sentence (5a), the patient *maanu iisi* 'this child' in the PF sentence (5b), the "perceived" location *paicʉ* 'PaicU' in the LF sentence (5c), and the instrument *vatu iisi* 'this stone' in the IF sentence (5d). It is noteworthy that some verbs can only be a PF form or an LF form but not both, such as **tavalaʔ-ai* 'know (PF)'/*tavalaʔ-an* 'know (LF)'[13]. Meanwhile, both PF and LF forms co-exist in other cases, as evidenced in *riucuucu-ai* 'kiss (PF)' in (5b) and *riucuucu-an* 'kiss (LF)' in (5c). There may still exist certain minute semantic/syntactic differences between *riucuucu-ai* and *riucuucu-an*. Thus, both the PF and LF markings are remained in this study.

In Table 3.1, there are three AF markers: <*um*>, *m-*, and Ø. The three AF markers are lexically conditioned allomorphs. Following Tsuchida (1976:47–48), Kanakanavu AF verbs are divided into three types (i.e. <*um*>-type, *m*-type, Ø-type) depending on which AF marker they can co-occur with. The examples of the three AF types are as follows:

(6) <*um*>-type:
a. *k<**um**>aʉn* *vutukulu* *maanu* *iisi.*
<AF>-eat fish child this
'This child ate/eats a fish.'

b. ***um**-ala/**m**-ala-ku* *tavʉnʉvʉnʉ.*
AF-take/AF-take-1S.NOM green.banana
'I took/take green bananas.'

c. ***mu**-kusa-ku* *na* *kasampu* .
AF-go-1S.NOM LOC Jiaxian[14]
'I am going/go to Kasampu.'

d. *c<**ʉm**>ʉʔʉra-cu-ku* *ŋuain.*
see<AF>-COS-1S.NOM 3.NOM
'I saw/see him/them.'

---

[12] The romanization spelling for the female name *paicʉ* is "'PaicU", where the high back unrounded /ʉ/ is represented by the capital "U". In this study we follow this conventional usage in our glossing and translation.

[13] Thus, PF and LF markings are merged as PF (that is, OF) in Tsuchida (1973, 1976).

[14] The Mandarin Chinese for the place name *kasampu* is 甲仙 'Jiaxian', which is the biggest town near Namasia.

(7) *m*-type:

***m**-itar~~u~~-kim* *maanu-mia.*
AF-wait-1PE.NOM child-1PE.GEN
'We waited/wait/are waiting for our child.'

(8) Ø-type:

tavala?~~u~~-cu *caau* *iisi.*
AF.know-COS person this
'This person knew (it).'

It is noteworthy that there are four variations of the AF marker <*um*>: <*um*>, *um-*, *mu-*, and <~~*u*~~*m*>, as shown in (6). These variations are phonologically conditioned (Tsuchida 1976:47–48).

Like the AF marker <*um*>, there are two variations of LF marking in Kanakanavu: *-a* and *-an*. The two variations are phonologically conditioned (Tsuchida 1976:49). When the LF verb is attached with a suffix, the LF marker is *-an*. Otherwise, the LF marker appears as *-a*. Both variations are exemplified in (9).

(9) a. ~~*u*~~*m*~~*u*~~*n-**a*** *cuma-aku* *kaalu* *iisi.*
tie.up-LF father-1S.GEN wood this
'My father tied/ties up these pieces of wood.'

b. ~~*u*~~*m*~~*u*~~*n-**an**-aku* *kaalu* *iisi.*
tie.up-LF-1S.GEN wood this
'I tied/tie up these pieces of wood.'

It is noted that the LF marker *-a*(*n*) needs to be distinguished from the LF circumfix *ta-...-a*(*n*). Once a verb is attached with the circumfix *ta-...-a*(*n*), it seems to be less verbal since it can also refer to a location/goal involving that verb, such as *ta-pakanaŋul-a*(*n*) 'swimming place' in (10a) and (10b).

(10) a. ***ta**-pakanaŋul-**a*** *ma-maanu* *cak~~u~~lan* *iisi.*
LF-swim-LF RED-child river this
'Children swam/swim in this river.'
or 'This river is where children swim.'

b. ***ta**-pakanaŋul-**an**-cu* *ma-maanu* *cakŋlan* *iisi.*
LF-swim-LF-COS RED-child river this
'Children swam/swim in this river.'
or 'This river is where children swam.'

In contrast, a LF form circumfixed with *ta-...-a(n)*may co-occur with an aspectual marker like *-cu* in (10b). Based on the above, the LF forms carry properties of both nouns and verbs.

Kanakanavu temporal/aspectual system is on a basis of a future/non-future dichotomy (known as realis/irrealis dichotomy) like most other Formosan languages (cf. Ogawa & Asai 1935; Tsuchida 1976; Huang 1995; Zeitoun 1996, 2007; Zeitoun et al. 1996; among many others). In many Formosan languages, a verb affixed with an AF or NAF marking (that is, the so-called "neutral" form) does not specify a fixed time reference when no tense/aspect marker is present in the clause (cf. Zeitoun et al. 1996:24). This phenomenon is also attested for Kanakanavu neutral verbs. Thus, the Kanakanavu neutral forms are claimed to carry an unmarked aspect by Tsuchida (1973, 1976) and Mei (1982). However, such an "unmarked

aspect" analysis for the neutral forms seems to be simplified since the neutral forms are not only related to aspectual reading. To be exact, the examples (5)–(6) show that the neutral forms are restricted to encoding non-future information, including past tense, present tense, present habitual, and present progressive. Despite this fact, the Kanakanavu focus verbal morphology (including both AF and NAF markers) cannot be treated as a certain type of tense and/or aspect marking mainly for two reasons. First, there is no one-to-one correspondence between a focus marker and a specific tense and/or aspect information. Second, the tense/aspect information in Kanakanavu is mainly determined by the combination of several elements, including context, an aspectual marker, and a temporal adverb (cf. Zeitoun 1996:507 and Zeitoun 2007:150).

As introduced above, the focus markings for neutral forms are responsible for non-future interpretation to some extent in Kanakanavu. Unlike other Formosan languages, the case in Kanakanavu is rather complex because another set of verbs, that is, the so-called 'imperfective' forms in the existing literature are also indicative of non-future reading. That is, there are two sets of verbal morphology responsible for the non-future reading in Kanakanavu. The neutral forms in Kanakanavu do not behave like the neutral forms in other Formosan languages. Compared to other Formosan languages, the occurrences of the neutral forms are rather limited and less frequent in Kanakanavu. Rather, the imperfective forms function more like the so-called neutral forms in other Formosan languages. As indicated by Tsuchida (1973:91), "When informants are asked to give verbs in isolation, they always respond with verbs in the imperfective aspect either in the AF or non-actor focus." Compare the two sentence sets of Mayrinax Atayal[15] and Kanakanavu below:

(11) Mayrinax Atayal:

a. ***m-uah*** *ʔiʔ* *limuy.*
AF-come NOM Limuy
'Limuy came/comes/is coming.'

b. ***m-a-(ʔ)uah*** *ʔiʔ* *limuy.*
AF-FUT-come NOM Limuy
'Limuy will come.'

(12) Kanakanavu:

a. ***i-a***[16]***-vatu*** */??**ivatu*** *caau* *iisi.*
come-AF.IPFV/AF.come person this
'This person is coming.' or 'This person (often) comes.'

b. *tia* ***i-a-vatu/*ivatu*** *caau* *iisi.*
FUT come-AF.IPFV/AF.come person this
'This person will come.'

In other Formosan languages, non-future interpretation is mostly morphologically reflected in

[15] Mayrinax Atayal is an archaic dialect of Atayal spoken in the mountainous area of central Taiwan. According to Blust (1999), the Atayalic group is one of the ten major branches that split directly from the PAN language.

[16] The imperfective marker *a-* is not an infix but a prefix since it is prefixed to the verb stem. It occurs right between an AF marker and a verb root, such as *mu-a-kusa* 'go (AF.IPFV)'. In a few cases of the Ø-type AF imperfective verbs, it seems to appear right after the first segment of the verb root, such as *i-a-vatu* 'come (AF.IPFV)'. Paul Jen-kuei Li indicates that the root for this word is *vatu*, not *ivatu*. However, the prefix *i-* in the verb stem *ivatu* might have lost its grammatical function in the historical development. (personal communication, April 21, 2012)

the neutral forms, as illustrated in the Mayrinax Atayal example (11a). In contrast, the non-future reading is mostly carried by the imperfective forms in Kanakanavu. Regarding the non-future usage, the Mayrinax Atayal neutral form *muah* 'come (AF)' in (11a) seems to correspond to the Kanakanavu imperfective form *iavatu* 'come (AF.IPFV)' in (12a) rather than the neutral form *ivatu* 'come (AF)' in (12a). Furthermore, the prefix*a-* in *ma(ʔ)uah* 'will come (AF.FUT)' in (11b) indicates a future event in Mayrinax Atayal, while the same marker *a-* is an imperfective marker in Kanakanavu, such as *a-* in the AF verb *iavatu* 'come (AF.IPFV)' in (12a). That is, the imperfective marking does not carry a future tense reading. In Kanakanavu, the future tense information is grammatically marked by a lexical item *tia*, as in (12b). Though the imperfective form is not responsible for the future tense reading, it is the imperfective forms, not the neutral forms, that can co-occur with the future marker *tia*. In the future tense sentence (12b), the imperfective form *iavatu* 'come (AF.IPFV)' is allowed, but the neutral form *ivatu* 'come (AF)' is not. According to Mei (1982:227–228), it is an important criterion that helps distinguish neutral forms from imperfective forms in Kanakanavu. That is, the imperfective forms can co-occur with the future marker *tia*, but the neutral forms cannot when there is a formal distinction between neutral and imperfective forms.

Below is a table which lists the focus markers for the neutral and imperfective forms.

Table 3.2: Kanakanavu neutral and imperfective markings (Tsuchida 1976; Mei 1982; Chang 2006; and Wu 2006).

Focus	Actor Focus (AF)			Patient Focus (PF)	LocationFocus(LF)	Instrument Focus (IF)	
marker for neutral form	<um>, <ʉm>	mu-, um-, m-, Ø	<um>, <ʉm>, m-, Ø	-ai	-a(n), ta-...-a(n)	si-	---
marker for imperfective form	<um>+Ca-reduplicant, <ʉm>+Ca-reduplicant	mu-a-, um-a-, m-a-, a-	---	-ʉn	---	---	si-a-

Compared to the neutral forms, the imperfective forms are morphologically marked in most AF verbs and some IF verbs since they have extra morphemes Ca- reduplicant or *a-*. As for the PF verbs, there are separate markers, that is, *-ai* and *-ʉn* for the neutral forms and imperfective forms respectively. As for some AF verbs and LF verbs, there are neutral forms only. Note that the imperfective forms are marked with 'IPFV', and neutral forms stay 'unmarked' in the glossing of this paper.

The AF imperfective markings can be divided into two groups: (i) Ca-reduplicant; and (ii) prefix *a-*. The variations of the AF imperfective forms are illustrated below:

(13)

AF marker	Stem	AF neutral form	AF imperfective marker	AF imperfective Form
<um>	ka~~u~~n ‘eat’	k<**um**>a~~u~~n	Ca- reduplicant	**k**<um>**a**-ka~~u~~n
<~~u~~m>	s~~u~~p~~u~~r ‘read’	s<~~**u**~~**m**>s~~u~~p~~u~~r~~u~~		**s**<~~u~~m>**a**-s~~u~~p~~u~~r~~u~~
um-	ala ‘take’	**um**-ala	a-	um-**a**-ala
mu-	panaʔ ‘shoot’	**mu**-panaʔ		mu-**a**-panaʔ
m-	ukusa‘go’	**m**-ukusa		m-**a**-ukusa
Ø	ivatu‘come’	ivatu		i-**a**-vatu
<um>	rac~~u~~k~~u~~c ‘step on’	r<**um**>ac~~u~~k~~u~~c~~u~~	---	---
<~~u~~m>	c~~u~~ʔ~~u~~ra ‘see’	c<~~**u**~~**m**>~~u~~ʔ~~u~~ra		---
m-	atupun‘throw’	**m**-atupun		---
Ø	arapara‘ run’	arapara‘~~u~~		---

The Ca- reduplicant of an AF imperfective form is composed of the first consonant of the verb stem and an invariant low vowel /a/, such as *ka-* of ***k****<um>****a****-ka~~u~~n* ‘eat (AF.IPFV)’in (14a) and *sa-* of ***s****<~~u~~m>****a****-s~~u~~p~~u~~r-an* ‘read (AF.IPFV.DEL)’ in (14b).

(14) AF imperfective marker: Ca- reduplication

a. ***k****<um>****a****-ka~~u~~n* *vutukulu* *maanu* *iisi.*
RED<AF>-eat fish child this
‘This childate/eats/is eating fish.’

b. ***s****<~~u~~m>****a****-s~~u~~p~~u~~r-an* *kamu.*
RED<AF>-read-DEL 2P.NOM
‘You read/read/are reading a little.’

The imperfective forms with a Ca- reduplicant can only be found in the AF forms whose AF markers are *<um>* or *<~~u~~m>*. Regarding the AF imperfective marker *a-*, it occurs immediately between a focus marker and a verb root, such as *m-a-itar~~u~~* ‘wait (AF.IPFV)’ in (15a) and *mu-a-ca* ‘leave (AF.IPFV)’ in (15b).

(15) AF imperfective marker: *a-*

a. *m-****a****-itar~~u~~-ku* *ʔinia* *kaal~~u~~ʔa-maku.*
AF-IPFV-wait-1S.NOM there lover-1S.GEN
‘I waited/wait/am waiting for my lover there.’

b. *mu-****a****-ca* *maanu-maku* *na* *takau.*
AF-IPFV-leave child-1S.GEN LOC Kaohsiung
‘My child left/leaves/is leaving for Kaohsiung.’

Note that some AF verbs do not have imperfective forms, such as *c<~~u~~m>~~u~~ʔ~~u~~ra* ‘see (AF)’ in (16a) and *puʔa* ‘buy (AF)’ in (16b).

(16) a. *c<~~u~~m>~~u~~ʔ~~u~~ra-ku/*****c****<~~u~~m>****a****-c~~u~~ʔ~~u~~ra -ku* *saruanai* *iisa.*
see<AF>-1S.NOM/RED<AF>-see-1S.NOM man that
‘I saw/see that man.’

b. *puʔa/*****a****-puʔa* *cina-maku* *camai.*
AF.buy/IPFV-AF.buy mother-1S.GEN vegetable
‘My mother will buy vegetables.’

The PF imperfective marker is a portmanteau morpheme -*ʉn*, which is indicative of both PF and imperfective markings. There are three variants -*ʉn*, -*un*, and -*in* for the default PF imperfective marker -*ʉn*.[17] These variants are phonologically conditioned (cf. Tsuchida 1973:94–95)[18]. For example:

(17) a. *cʉʔʉr-**ai** /cʉʔʉr-**ʉn*** *maanu* *iisi* *(sua)* *tacau* *iisa.*
see-PF /see-PF.IPFV child this NOM dog this
'This child saw/sees <u>that dog</u>.'

b. *patupun-**ai** /patupun-**un*** *caau* *ŋiau* *iisa.*
throw-PF /throw-PF.IPFV person cat that
'Someone threw/throws away <u>that cat</u>.'

c. *aliapiniŋ-**ai** /aliapiniŋ-**in*** *maku* *tuaŋiŋi* *iisa.*
take.out-PF /take.out-PF.IPFV 1S.GEN mouse that
'I took/take out <u>that mouse</u>.'

As shown in (17), both PF neutral and imperfective forms are indicative of a past or present (progressive) event. However, the PF imperfective forms are more frequently attested than the PF neutral counter parts in my data corpus.

No imperfective forms are attested in the LF verbs. That is, only neutral forms can be found in the LF verbs, as in (18).

(18) a. ***ta**-ima-**a*** *ma-maanu* *avaʉvʉna* *iisa.*
LF-drink-LF RED-child spring that
'Children drank/drink (water) <u>in that spring</u>.'
'<u>That spring</u> was/is where children drank/drink (water).'

b. *tia* ***ta**-ima-**an**-ta* *avaʉvʉna* *iisa.*
FUT LF-drink-LF-1PI.NOM spring that
'We will drink (water) <u>in that spring</u>.'
'<u>That spring</u> is where we will drink (water).'

As indicated before, the neutral forms cannot co-occur with the future marker *tia* when the verbs have imperfective forms (Mei 1982). Because there are no imperfective forms in the LF verbs, the neutral LF verbs are allowed to co-occur with *tia*, such as *ta-ima-an-ta* 'we drink (LF)' in (18b).

The IF verbs can be divided into two groups: (i) some IF verbs have only neutral forms, such as *si-patupun*'throw (IF)'/**si-a-patupun*'throw (IF.IPFV)' in (19) and *si-puʔa* 'buy (IF)'/**si-a-puʔa* 'buy (IF.IPFV)' in (20); and (ii) some IF verbs have only imperfective forms[19], such as *si-a-pauʔucipi* 'cook (AF.IPFV)'/**si-pauʔucipi* 'cook (AF)' in (21) and *si-a-livʉlaʔʉ* 'beat (AF.IPFV)'/**si-livʉlaʔʉ* 'beat (AF)' in (22).

---

[17] The previous studies propose three different analyses for the PF marking in Kanakanavu. In Tsuchida's (1973, 1976) analysis, the PF and LF markings are merged into one marking (that is, Goal/Location focus in his terminology): Ø (for neutral forms) and -*ʉnʉ* (for imperfective forms). Mei (1982) proposes -*ʉ* and -*ai* for the PF neutral forms and -*ʉnʉ* for PF imperfective forms. My analysis is the same as Chang 2006 and Wu 2006: -*ai* is the PF neutral marker, and -*ʉn* the PF imperfective marker.

[18] According to Tsuchida (1973:94–95), the variants of the PF marking include -*ini*, -*ni*, -*unu*, and -*ʉnʉ*.

[19] It is unusual that these IF verbs have no neutral forms but imperfective forms. In most cases, the Kanakanavu verbs have at least neutral forms no matter whether they have imperfective counterparts or not.

(19) a. ***si**-patupun /***si-a**-patunun caau iisi vavulu sua vatu iisa.*
IF-throw/IF-IPFV-throw person this wild.pig NOM stone that
'This person used/uses <u>that stone</u> to throw at a wild pig.'

b. *tia **si**-patupun /***si-a**-patunun caau iisi vavulu sua vatu iisa.*
FUT IF-throw/IF-IPFV-throw person this wild.pig NOM stone that
'This person will throw <u>that stone</u> at a wild pig.'

(20) a. ***si**-puʔa/***si-a**-puʔa maanu-maku ʔ~~ʉ~~nai sua vantuku iisi.*
IF-buy/IF-IPFV-buy child-1S.GEN land NOM money this
'My child used/use <u>this money</u> to buy land.'

b. *tia **si**-puʔa/***si-a**-puʔa maanu-maku ŋ~~ʉ~~nai sua vantuku iisi.*
FUT IF-buy/IF-IPFV-buy child-1S.GEN land NOM money this
'My child will use <u>this money</u> to buy land.'

(21) a. ***si-a**-pauʔucipi /***si**-pauʔucipi cina sua kaalu.*
IF-IPFV-cook/IF-cook mother NOM wood
'<u>The wood</u> is what Mother used/uses to cook (food) with.'

b. *tia **si-a**-pauʔucipi /***si**-pauʔucipi cina sua kaalu.*
FUT IF-IPFV-cook/IF-cook mother NOM wood
'<u>The wood</u> is what Mother will use to cook (food) with.'

(22) a. ***si-a**-liv~~ʉ~~laʔ~~ʉ~~ /***si**-liv~~ʉ~~laʔ~~ʉ~~ cuma maanu sua uriŋ.*
IF-IPFV-beat/IF-beat father child NOM miscanthus
'Father used/uses <u>miscanthus</u> (stems) to beat (his) children.'

b. *tia **si-a**-liv~~ʉ~~laʔ~~ʉ~~ /***si**-liv~~ʉ~~laʔ~~ʉ~~ cuma maanu sua uriŋ.*
FUT IF-IPFV-beat/IF-beat father child NOM miscanthus
'Father will use <u>miscanthus</u> (stems) to beat (his) children.'

Note that most of the IF verbs select the marker for IF imperfective form *sia-* instead of the IF neutral marker *si-*. Therefore, in Wu's (2006) analysis[20], *sia-* instead of *si-* is treated as the IF marker for both neutral and imperfective forms.

The imperfective markers are traditionally treated as imperfective aspect markers by Tsuchida (1973, 1976) and Mei (1982). According to Tsuchida (1976:52), an imperfective form is used to "express an incomplete action or event whether it is momentary or durative, past, present, or future". Thus, it is named "imperfective" in the previous studies. This study proposes that the imperfective forms in Kanakanavu cannot be treated as a type of aspect, as evidenced mainly in two facts. First, the imperfective forms in Kanakanavu function like the neutral forms in other Formosan languages. Like the neutral forms in other Formosan languages, the imperfective forms are not indicative of a specific tense/aspect marking though they are restricted to a non-future domain. Second, the Kanakanavu imperfective forms can indicate incomplete actions as well as complete actions, depending on the context or the intention of the speaker, as in (14)–(17), (18a), (19a), (20a), (21a), and (22a). Maslov (2010) indicates that "the imperfective aspect provides no indication of the completeness or finality of an action." The Kanakanavu imperfective forms in question are commonly used to refer to an on-going action, a

[20] It is noted that *sia-* is written as *se-* in Wu (2006). The difference between *sia-* and *se-* is because the Takanua speakers tend to co-articulate /iɟa/ as [e].

repeated action, and a habitual action, all of which are categorized as carrying imperfective aspect. Meanwhile, the same imperfective forms can also refer to completed actions in some context. For example:

(23) *mu-a-ran/mu-ran* *ikua* *kaalʉʔa-maku* *m-ala* *savuana.*
AF-IPFV-help/AF-help 1S.OBL lover-1S.GEN AF-take medicine
'My lover helped/helps me get medicine.'

(24) *cʉʔʉra-ʉn-aku* *sua* *taniarʉ.*
see-PF.IPFV-1S.GEN NOM sun
'I saw/see the sun.'

As illustrated in (23) and (24), the imperfective forms *mu-a-ran* 'help (AF.IPFV)' and *cʉʔʉra-ʉn* 'see (PF.IPFV)' do not show whether the actions in question are completed or not. Thus, the so-called "imperfective" in Kanakanavu cannot be analyzed as simply marking the imperfective aspect. Based on the above discussion, we can only conclude that the neutral and imperfective forms in Kanakanavu are both confined to encoding non-future information when there is no tense/aspect marker or time adverb in the clause, such as *muaran*/*muran* 'help (AF.IPFV)'/'help (AF)' in (23).

4 Aspectual markers and the neutral/imperfective forms

This section explores how the aspect information is overtly marked by the grammatical markers in Kanakanavu and how the Kanakanavu neutral/imperfective forms interact with these aspectual markers. I shall introduce three aspectual markings in Kanakanavu, including: (i) perfective aspect: *ni-*/<*in*>; (ii) change-of-state aspect: *-cu*; and (iii) durative aspect: *ʔaisi* and CVCV- reduplication.

4.1 Perfective aspect

The perfective affixes <*in*>/*ni-* are used to describe a situation viewed as a past action or completed event with present/past relevance, or a present state resulting from a past situation (cf. Tsuchida 1976; Mei 1982; Chang 2006; and Wu 2006). For example:

(25) a. *c<**in**>m>ʉʔʉra-ku* *paʔʉ.*
see<PERF>AF>-1S.NOM Pa'U
'I have seen Pa'U.'

b. ***ni**-kaʉn* *maanu* *iisi* *vutukulu.*
PERF.PF-eat child this fish
'This child has eaten the fish.'

As shown in (25), there is no clear-cut distinction between past tense and present perfective. As long as the action has been finished, it is consistently marked with the perfective marking no matter whether the action has been done just now, recently or in the past. A further clarification relies on context and/or time adverbs available in the sentence, such as *sauna* 'now; just now' and *miʔʉra* 'yesterday'.

There are two perfective forms *ni-* and <*in*> in Kanakanavu. The perfective marker *ni-* is the default form. The other perfective marker <*in*> is attested only in some of the IF verbs and the <*um*>-type AF verbs whose stems begin with /t/, /c/, or /s/ (Tsuchida

1976:54–55).The examples of the perfective forms are given below:

(26) AF perfective forms:

a. ***ni***-*k<um>aʉn-kasu* *savuana.*
PERF-eat<AF>-2S.NOM medicine
'You have taken medicine.'

b. *c<**in**>m>ʉʔʉra-ku* *ʔavia.*
see<PERF>AF>-1S.NOM 'avia
'I have seen 'avia.'

c. ***ni***-*vatu* *maanu* *iisi.*
PERF-AF.come child this
'This child has come.'

(27) NAF perfective forms:

a. ***ni***-*tʉa* *caau* *maanu* *maku.*
PERF.PF-find person child 1S.GEN
'Someone has found my child.'

b. ***ni***-*tʉa-an-cu* *caau* *maanu* *maku.*
PERF-find-LF-COS person child 1S.GEN
'Someone has found my child.'

c. *s<**in**>i-sʉʔʉ/**ni**-si-sʉʔʉ-cu-maku* *sua* *maʔain* *camai* *iisi.*
IF<PERF>-put/PERF-IF-put -COS-1S.GEN NOM salt vegetable this
'I have putsalt in this vegetable.'

As shown in (26) and (27), the perfective marking (i.e. *ni-* or<*in*>) is affixed to a neutral verb to indicate that the event or action has been completed.

The perfective marking can also be attached to imperfective verbs to signal that the action/state has been done for some time with reference to the speech time. Consider:

(28) ***ni***-*m-a-itarʉ-ku* *ʔinia.*
PERF-AF-IPFV-wait-1S.GEN there
'I have been waiting there.'

(29) a. ***ni***-*m-usa-ku* *lipun* *makaan.*
PERF-AF-go-1S.NOM Japan once
'I have been to Japan once.'

b. ***ni***-*m-a-usa-ku.*
PERF-AF-IPFV-go-1S.NOM
'I have gone (for some time) and am still on the way.'
(lit. 'I have been going.')

For instance, *ni-m-a-itarʉ* 'wait (AF.IPFV)' in (28) implies that the speaker has waited there for some time but will keep waiting until he meets the person with whom he has an appointment. The examples in (29) illustrate how the AF neutral and imperfective forms for 'go' differ when both are affixed with a perfective marker. When *ni-* is attached to the AF neutral verb *m-usa* 'go (AF)' like *ni-m-usa*'go (AF.PERF) 'in (29a), it is used to signal that the action of 'go' has been terminated a long time ago and is referred to as an experience with respect to the speech-time. In contrast, the imperfective verb *ni-m-a-usa* in (29b) suggests that the speaker has set off and is still on the way. This suggests that these

AF imperfective forms like *ni-m-a-usa* 'go (AF.IPFV.PERF)' in (29b) are associated with the concept of "incompleteness", while the AF neutral forms are not. According to my database, such a fine distinction is only found in a few examples of the AF neutral/imperfective verbs. This can account for the phenomenon: the perfective marking tends to co-occur with the neutral forms instead of the imperfective counterparts in most cases of the AF verbs. This constitutes another useful tendency for dividing the AF neutral forms from the AF imperfective forms in addition to the one proposed by Mei 1982. According to Mei (1982), the imperfective form can co-exist with the future tense marker *tia* but the neutral counterpart cannot when there is a formal distinction between neutral and imperfective forms.

The perfective marker (i.e. *ni-* or <*in*>) cannot co-occur with a PF marker (i.e. *-ai* or *-ʉn*).[21] Despite this fact, a verb stem affixed only with a perfective marker functions as a PF verb, such as *ni-kaʉn* 'eat (PF.PERF)' in (30a) and *ni-itarʉ* 'wait (PF.PERF)' in (30b).

(30) PF perfectives:

a. ***ni**-kaʉn/***ni**-kaʉn-**ai**/***ni**-kaʉn-**ʉn** *maanu* *vutukulu* *iisi.*
PERF.PF-eat child fish this
'(Some) child has eaten <u>this fish</u>.'

b. ***ni**-itarʉ/***ni**-itarʉ-**ai** /***ni**-itarʉ-**ʉn** -cu-maku* *kaalʉʔa-maku,* *nakai*
PERF-wait -COS-1S.GEN lover-1S.GEN but
kuu *ivatu.*
NEG AF.come
'I have waited for <u>my lover</u>, but (he) did not come.'

In my database, one exception is attested in that both the perfective marker *ni-* and the PF marker *-ai* are attached to the verb *pakanaŋul* 'swim', as in (32b).

(31) *pakanaŋul-**ai*** *ma-maanu* *cakʉlan* *iisi.*
swim-PF RED-child river this
'Children swam/swim <u>in this river</u>.'

(32) a. ***ni**-pakanaŋulu-cu-maku* *cakʉlan* *iisi.*
PERF-swim-COS-1S.GEN river this
'I have swum <u>in this river</u>.'

b. ***ni**-pakanaŋulu-**ai**-cu-maku* *cakʉlan* *iisi.*
PERF-swim-PF-COS-1S.GEN river this
'I have swum <u>in this river</u> (before).'

My language consultant indicates that the PF perfective *ni-pakanaŋulu-ai* in (32b) shows that the speaker sounds like sort of showing off to imply that he has swum in this river before. However, the PF perfective *ni-pakanaŋulu* in (32a) does not have such a connotation but simply shows that the speaker has finished the action of 'swimming in this river' or has just experienced 'swimming in this river'. At this point, it is uncertain about the exact function of this marker *-ai* in (32b). I will leave this for further study.

Like the AF verbs, the perfective marker *ni-* co-exists with the LF and IF markers

[21] It holds true for many Formosan languages in that the perfective marker cannot co-occur with a PF marker. Blust (1998) indicates that the co-occurrence of the perfective marker and a PF marker is attested in some examples of Thao. However, Li (2011:13–14) indicates that these examples in Thao are just exceptions.

respectively. As introduced in Section 3, there are two variants of the LF neutral/imperfective marking: *-a*(*n*) and *ta-...-a*(*n*). When the two LF forms are affixed with the prefix *ni-*, their verbal morphology consistently appears as a circumfix *ni-...-a*(*n*), such as *ni-pisanuŋ-an-cu-maku* 'I have got (sth.) in (LF.PERF)' in (33a) and *ni-pakanaŋul-a* 'swim (LF.PERF)' in (33b).

(33) a. ***ni**-pisanuŋ-**an**-cu-maku* *nakʉvʉ-su.*
PERF-get.in-LF-COS-1S.GEN clothes-2S.GEN
'I have got <u>your clothes</u> in.'

b. ***ni**-pakanaŋul-**a**/***ni-ta**-pakanaŋul-**a*** *ma-maanu* *cakʉlan* *iisi.*
PERF-swim-LF/PERF-LF-swim-LF RED-child river this
'The children have swum in <u>this river</u>.'

As shown in (33b), the perfective *ni-* replaces the *ta-* of the circumfix*ta-...-an* instead of co-occurring with the prefix **ni-ta-*. That is to say, the two variations of the LF neutral/imperfective marking *-a*(*n*) and *ta-...-a*(*n*) are merged into one perfective LF circumfix *ni-...-a(n)*.

The perfective marker *ni-* co-exists with the IF marker (i.e. *si-* or *sia-*), respectively, as in (34a) and (34b).

(34) a. ***ni**-si-patupun/***ni**-si-a-patupun -cu* *maanu* *tacau* *sua* *vatu* *iisi.*
PERF-IF-throw/PERF-IF-throw -COS child dog NOM stone this
'A child has thrown <u>this stone</u> at a dog.'

b. ***ni**-si-a-sʉʔʉ/* **ni**-si-sʉʔʉ -cu-maku* *sua* *maʔain* *ʔaau.*
PERF-IF-IPFV-place/PERF-IF-place-COS-1S.GEN NOM salt soup
'I have put <u>salt</u> in the soup.'

Notice that the two forms, that is, *si-* marked IF verbs and *sia-* marked IF verbs are not interchangeable. As introduced in Section 3, the IF verbs are divided into two groups: some IF verbs have neutral forms only, while some IF verbs have imperfective forms only. This uninterchange ability also holds for the *si-* marked perfective forms and *sia-* marked perfective forms. That is to say, there is a co-occurrence restriction between the perfective marking and the IF neutral/imperfective forms. The selections of the IF neutral/imperfective perfective forms are lexically conditioned.

4.2 Change-of-state aspect

The aspectual marker *-cu* is used to signal a change of state (cf. Wu 2006). This change-of-state grammatical marker *-cu* is interpreted as carrying perfective aspect or past tense when it occurs in a bound event. As illustrated in (35), the occurrence of *-cu* denotes the event that has terminated or the event that has happened.

(35) a. *ʔacai-**cu*** *ŋuain.*
AT.leave-COS 3.NOM
'<u>He</u> (has) passed away.'
(lit. '<u>He</u> (has) left.')

b. *m-akasi-**cu**-ku* *(na)* *tanasa-musu.*
AT-go.toward-COS-1S.NOM LOC house-2S.GEN
'<u>I</u> (have) arrived at your house.'

To be exact, the basic communicative function of the aspect marker *-cu* is to signal that "a state of affairs has special current relevance with respect to some particular situation" (cf. Li and Thompson 1981:240). For example:

(36) a. *tavalaʔ~~u~~-**cu*** *maanu* *iisi* *ni-ka~~u~~n* *ŋiau* *sua* *tapianaŋai.*
AF.know-COS child this PF.PERF-eat cat NOM bird
'This child knew that some cat had eaten the bird.'

b. *ʔucaŋ-aku* *ia,* *m-an~~u~~n* *caau* *miaʔinan.* *aaka-**cu*** *sauni.*
wife-1S.GEN TOP AF.good person before AF.bad-COS now
'My wife, her health was good before. (Her health) has become bad now.'

As shown in (36), the change-of-state marker *-cu* is not used to indicate that the action has been finished. Also, it is impossible to interpret that the action have been ended for an event like *tavalaʔ~~u~~* 'know (AF)' in (36a) and *aaka* 'bad (AF) in (36b). Instead, both cases emphasize that the speaker has noticed the change of state with some current relevance to some particular (or implied) situation. Like what Li and Thompson (1981:240) proposed for the Chinese sentence-final particle *le* (了), the Kanakanavu aspect marker *-cu* may be better described as a marker signaling a "recent change of state" rather than a perfective aspect which takes the whole event as being bounded temporarily, spatially, or conceptually.

The aspect marker *-cu* is directly suffixed to the sentence-initial element, such as the future marker *tia* in (37a), *cuculu* 'really' in (37b), *maaraŋ* 'old' in (37c), and *ʔacai* 'leave (AF)' in (37d).

(37) a. *tia-**cu**-ku* *ara-~~u~~n* *caau.*
FUT-COS-1S.NOM get-PF person
'I will get married.'
(lit. 'I will be caught by a person.')

b. *cuculu-**cu**-ku* *mu-a-pacai.*
really-cos-1S.NOM AF-IPFV-get.tired
'I am really tired.'

c. *maaraŋ-**cu**-kasu.*
old-COS-2S.NOM
'You are getting old.'

d. *ʔacai-**cu*** *cina* *matu* *maanu* *tanasa.*
AF.leave-COS mother and child house
'Mother and (her) child left for home.'

There is no co-occurrence between the aspectual marker *-cu* and the neutral/perfective forms. For example, the aspect marker *-cu* can be suffixed to either a neutral form or an imperfective counterpart, as in (38).

(38) a. *mu-ran-**cu*** *ikua* *ʔucaŋ-maku.*
AF-help-COS 1S.OBL wife-1S.GEN
'My wife helped me.'

b. *mu-a-ran-**cu*** *ikua* *ʔucaŋ-maku.*
AF-IPFV-help-COS 1S.OBL wife-1S.GEN
'My wife helped me.'

4.3 Durative

In Kanakanavu there are two grammatical markers encoding durative aspect: (i) *ʔaisi* (progressive marker); and (ii) CVCV- reduplication (iterative marker).

The progressive aspect marker *ʔaisi* appears in the sentence-initial position to signal an on-going activity or mental status. When there is a distinction between neutral and imperfective forms, the progressive marker *ʔaisi* is only allowed to co-occur with an imperfective form in most cases, as illustrated in (39a) and (39b).

(39) a. *ʔaisi-ku* ***mu-a-ciri/*mu-ciri.***
PROG-1S.NOM AF-IPFV-stand/AF-stand
'I am standing.'

b. *ʔaisi-ku* ***k<um>a-kaʉn/*k<um>aʉn*** *talukuka.*
PROG-1S.NOM RED<AF>-eat/eat<AF> chicken
'I am eating chicken.'

Still, *ʔaisi* can co-exist with either a neutral or imperfective form in some cases, as in (40).

(40) a. *ʔaisi-kim* ***m-itarʉ/m-a-itarʉ*** ***maanu-mia.***
PROG-1PE.NOM AF-wait/AF-IPFV-wait child-1PE.GEN
'We are waiting for our child.'

b. *ʔaisi-ku* ***mu-raan/mu-a-raan*** *c<ʉm>ʉʔʉra* *maanu-su.*
PROG-1PE.NOM AF-help/AF-IPFV-help see<AF> child-2S.GEN
'I am helping look after your child.'

Nevertheless, it is not clear if there is any semantic/pragmatic difference between neutral and imperfective forms in (40).

The progressive marker *ʔaisi* can occur in the AF and NAF sentences, as illustrated in (41) and (42) respectively.

(41) ***ʔaisi-ku*** *arapanaʔʉ.*
PROG-1S.NOM AF.run
'I am running.'

(42) a. ***ʔaisi*** *ʉmʉn-a* *cuma-aku* *kaalu* *iisi.*
PROG tie.up-LF father-1S.GEN wood this
'My father is tying up these pieces of wood.'

b. ***ʔaisi-maku*** *ʉmʉn-a* *kaalu* *iisi.*
PROG-1S.GEN tie.up-LF wood this
'Iam tying up these pieces of wood.'

The progressive marker *ʔaisi* can be attached with a bound personal pronoun, such as *-ku* 'I (1S.NOM)' in (41) and *-maku* 'I (1S.GEN)' in (42b).

The iterative aspect employs a morphological device of CVCV- reduplication to denote a continuously repeated action or physical disposition of an entity. The reduplicant copies the last two syllables of the base except the coda of the last syllable, such as *taŋi-* of *t<um>aŋi-taŋi* in (43b) and *kili-* of *kili-kilim-in* in (44b).

(43) a. ***t<um>aŋi-cu*** *maanu* *iisi.*
<AF>-cry-COS child this
'This child cried.'

b. ***t<um>aŋi-taŋi*** *maanu* *iisi.*
RED<AF>-cry child this
'The child keeps crying.'

(44) a. *tia-maku* ***kilim-in*** *maanu* *maku.*
FUT-1S.GEN search-PF child 1S.GEN
'I will search for my child.'

b. *tia-maku* ***kili-kilim-in*** *maanu* *maku.*
FUT-1S.GEN RED-search-PF child 1S.GEN
'I will keep searching for my child.'

Note that the base form does not include the word-final echo vowel. For example, the final vowel ***ʉ*** of *arapanaʔʉ* in (45b) is not part of the stem. Instead, it is a supporting vowel. According to Tsuchida (1976:33), a Kanakanavu word which ends in a consonant other than a nasal is normally inserted with a word-final supporting vowel.

(45) a. *tia-ku* ***arapanaʔʉ.***
FUT-1S.NOM AF.run
'I will run.'

b. *tia-pa-ku* ***ara-pana-panaʔʉ.***
FUT-1S.NOM AF.run-RED
'I will run.'

The CVCV- reduplication can be attested in both AF and NAF verbs, as illustrated by (43b)/(44b) and (45b), respectively. It is noteworthy that CVCV- reduplication is also used to indicate a plural agreement with a focused plural NP in a few cases. Consider:

(46) a. *ni-mu-ciri /* ni-mu-**ciri**-ciri* *cau* *iisa.*
PERF-AF-stand/PERF-AF-RED-stand person that
'That person has stood up.'

b. *ni-mu-**ciri**-ciri /*ni-mu-ciri* *cau* *iisa* ***kavaŋvaŋ.***
PERF-AF-RED-stand/PERF-AF-stand person those all
'Those people all have stood up.'

In (46b), the verb is required to be affixed with a CVCV- reduplicant *ciri-* when the focused NP is a plural NP *caau iisa kavaŋvaŋ* 'all those people'. Compared to (46b), the verb in (46a) cannot take a CVCV- reduplicant since the focused NP *caau iisa* is referred to as a singular NP 'that person', not a plural NP 'those people'.

4.4 Recapitulation

Based on the above discussion, the neutral and imperfective forms constitute a complementary distribution in terms of their co-occurrence with the tense/aspect markings, as evidenced in three conditions listed in Table 4.1.

Table 4.1: Co-occurrence between the neutral/imperfective forms and tense/aspect markings in Kanakanavu.

	Neutral forms	**Imperfective forms**
(i)	Tend not to co-occur with the future tense marker *tia*.	Tend to co-occur with the future tense marker *tia*.
(ii)	Tend to co-occur with the perfective marker *ni*-.	Tend not to co-occur with the perfective marker *ni*-.
(iii)	Tend not to co-occur with the progressive marker *ʔaisi*.	Tend to co-occur with the progressive marker *ʔaisi*.

The complementary distribution in Table 4.1 manifests a tendency for the distributions of the Kanakanavu neutral/imperfective forms: the neutral forms tend to co-occur with an aspectual marking involving the concept of completeness (i.e. the perfective marker *ni*-), while the imperfective forms tend to co-occur with an aspectual marking involving the concept of incompleteness (i.e. the future tense marker *tia* and the progressive marker *ʔaisi*). Given this tendency, there is still no clear-cut dividing line between the neutral and imperfective forms. In addition, the neutral/imperfective forms also constitute an overlapping distribution since both can co-exist with the change-of-state marker -*cu*.

5 Conclusion

This study explores two research questions: (i) "how is the temporal/aspectual information encoded in Kanakanavu neutral and imperfective forms?"; and (ii) "how are the Kanakanavu aspectual expressions integrated with the neutral/imperfective forms?" In other Formosan languages, a verb affixed with a focus marker (that is, the so-called neutral verb) is related to a non-future reading. Being unique among the Formosan languages, the neutral verbs in other Formosan languages are split into two sets (i.e. neutral and imperfective forms) in Kanakanavu. Furthermore, the Kanakanavu imperfective forms function like neutral forms in other Formosan languages in most cases. This study elaborates on the fine variations of the verbal morphology of the neutral and imperfective forms. Compared to the neutral forms, the imperfective forms are morphologically marked in most AF verbs and some IF verbs since they have an extra morpheme (i.e. *a*- or Ca-reduplicant). As for the PF verbs, there are separate markers, that is, -*ai* and -*ʉn* for the neutral and imperfective forms, respectively. As for some AF verbs and LF verbs, there are neutral forms only.

In the existing literature, the neutral and imperfective forms are claimed to carry unmarked and imperfective aspect respectively (Tsuchida 1973, 1976; Mei 1982). However, this study shows that such an unmarked/imperfective aspect analysis seems to be simplified and misses some generalizations to account for the Kanakanavu language facts.

Regarding the 'unmarked aspect' analysis for the neutral forms, the neutral forms are treated as denoting no particular mode of action in relation to the passage of time. This claim seems to be correct since the temporal fuzziness is attested in the neutral forms. When there is no temporal marker or time adverb in the sentence, there exist various possible tense/aspect readings for the neutral verbs, including past tense, present tense, present habitual, and present progressive. This language fact suggests that the neutral forms alone be restricted to a non-future interpretation instead of an "unmarked" aspect or tense. Note that the non-future domain categorized by the neutral verbs is not limited to a tense domain, as the name 'non-future' suggested. In Kanakanavu, tense interacts with aspect just

like some Indo-European languages (cf. Comrie 1985 and Bybee et al. 1994).

According to Tsuchida (1976:56), the imperfective markers are treated as carrying imperfective aspect, expressing "an incomplete action or event whether it is momentary or durative, past, present, or future". This study proposes that the 'imperfective aspect' analysis does not hold for the imperfective forms in Kanakanavu mainly based on three reasons. First, the imperfective forms in Kanakanavu function like the neutral forms in other Formosan languages. Like the neutral forms in other Formosan languages, the imperfective forms are not indicative of a specific tense/aspect marking though they are restricted to a non-future domain. Second, the so-called Kanakanavu "imperfective" forms can refer to an incomplete action as well as a complete action depending on the context or the intention of the speaker. Third, the imperfective forms do not always co-occur with an aspectual marker involving the concept of incomplete. But there is no denying that the neutral forms tend to co-occur with an aspectual marking involving the concept of completeness (i.e. the perfective marker *ni*-), while the imperfective forms tend to co-occur with an aspectual marking involving the concept of incompleteness (i.e. the future tense marker *tia* and the progressive marker *ʔaisi*). Despite this tendency, there is still no clear-cut dividing line between the neutral and imperfective forms. Moreover, the neutral/imperfective forms also constitute an overlapping distribution since both can co-exist with the change-of-state marker -*cu*.

References

Blust, Robert A. 1998. Squib: A note on the Thao patient focus perfective. *Oceanic Linguistics* 37.2:346–353.

Blust, Robert A. 1999. Subgrouping, circularity and extinction: Some issues in Austronesian comparative linguistics. In Elizabeth Zeitoun and Paul Jen-kuei Li eds., *Selected Papers from the Eighth International Conference on Austronesian Linguistics*, 31–94. Taipei: Academia Sinica.

Bybee, Joan L., Revere Perkins, and William Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. University of Chicago Press.

Chang, Yung-li. 2006. Rethinking the Tsouic subgroup hypothesis: A morphosyntactic perspective. In Yung-li Chang, Lillian Huang, andDah-an Hoeds., *Streams Converging into an Ocean*: *Festschrift in Honor of Professor Paul Jen-Kuei Li on his 70th Birthday*. Taipei: Institute of Linguistics, Academia Sinica.

Comrie, Bernard. 1985.*Tense*. Cambridge University Press.

Frawley, William. 1992.*Linguistic Semantics*. Hillsdale, NJ, Hove & London: Lawrence Erlbaum.

Huang, Lillian Mei-jin. 1995. *A Study of Mayrinax Syntax*. Taipei: The Crane Publishing Co.

Li, Charles N., and Sandra A. Thompson. 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press.

Li, Paul Jen-kuei. 2011. *Thao Texts and Songs*. Language and Linguistics Monograph Series 44. Taipei: Academia Sinica.

Maslov, IU. S. 2010 (1970–1979). *The Great Soviet Encyclopedia*. 3rd edition. The Gale Group, Inc.

Mei, Kuang. 1982. Pronouns and verb inflection in Kanakanavu. *Tsinghua Journal of Chinese Studies, New Series* 14: 207–32.

Ogawa, Naoyoshi and Asai, Erin. 1935. *The Myths and Traditions of the Formosan Natives Tribes* [in Japanese]. Taihoku: Taihoku Imperial University.

Tsuchida, Shigeru. 1973. Outline of Kanakanavu grammar. Unpublished manuscript.

Tsuchida, Shigeru. 1976. *Reconstruction of Proto-Tsouic Phonology*. Tokyo: Study of Languages & Cultures of Asia & Africa, Monograph Series no.5. Tokyo: Tokyo University of Foreign Studies.

Wu, Chunming. 2006. Verb serialization in Kanakanavu. *University System of TaiwanWorking Papers in Linguistics* 2: 109–39.

Zeitoun, Elizabeth. 1996. The Tsou temporal, aspectual and modal system revisited. *Bulletin of the Institute of History and Philology* 67.3:503–31.

Zeitoun, Elizabeth. 2007. *A Grammar of Mantauran* (*Rukai*). Language and Linguistics Monograph Series A4-2. Taipei: Academia Sinica.

Zeitoun, Elizabeth, Lillian M. Huang, Marie M. Yeh, Anna H. Chang, and Joy J. Wu. 1996. The temporal, aspectual, and modal systems of some Formosan languages: A typological perspective. *Oceanic Linguistics* 35.1:21–56.

10 *A unified account of the Tagalog verb and adjective affix systems*

RESTY M. CENA

1 Adjective as a Word Category in Tagalog

The status of adjective as a distinct word class has been a subject of discussion in many languages (for Austronesian languages, see for example Dixon (1982, 2004), and Pearson (2010)). The matter has not received wide attention among Tagalog linguists. The reason likely has to do with the fact that Tagalog adjectives—in contrast with verbs—clearly undergo different morphosyntactic processes. Below we show some processes that apply to adjectives but not to verbs.[1,2]

First, the superlative affix *pinaka-* takes adjective roots and stems, but not verb roots and stems.

(1) Adjective
a. *pinaka-sikat*
DEG.SUPER-buy
'most popular'

Verb
b. **pinaka-takbo*
DEG.SUPER-run

(2) a. *pinaka-ma-bait*
DEG.SUPER-ADJ-kind
'kindest'

b. **pinaka-t-um-akbo*
DEG.SUPER-ASP.PERF-run

---

1 We prefer to use our data set, rather than the data in de Guzman (1996:311), where she also concluded that adjective forms "still comprise a distinctive paradigm."

2 Abbreviations used in this paper:

3pl = third person, plural
abs = absolutive
accom = accomplish mode
adj = adjective; adjectivalizing affix
agt = agentive voice, affix, or role
appl = applicative voice, affix, or role
apt = aptative
asp = aspect
AspP = aspect phrase
beg = the aspect feature 'begun'
cause = causative mode
contemp = contemplated aspect
cv_r = consonant-vowel redup
deg = degree
delib = deliberate mode
det = determiner
dw_r = derived word redup
fin = the aspect feature 'finished'
gen = genitive
intense = intensive degree
ip = inflection phrase
lnk = linker
obj = objective voice, affix, or role
perf = perfective aspect
pred = predicate
pw_r = partial word redup
rw_r = root-word redup
super = superlative degree
V = verb
vce = voice
vP = small VP
vp = verb phrase
v = small verb

Asia-Pacific Linguistics, 2014.

Similarly the intensive affix *napaka-* takes adjective roots and stems, but not verb roots and stems.

(3) Adjective
a. *napaka*-bait
DEG.INTENSE-kind
'very kind'

Verb
b. **napaka*-takbo
DEG.INTENSE-run

(4) a. *napaka-ma-pag-bigay*
DEG.INTENSE-ADJ-MODE.DELIB-give
'very giving (of oneself)'

--

The special verb *maging* 'to be, to become' takes as its complement an adjective (or noun) stem but not a verb stem.

(5) a. *naging ma-bait*
ASP.PERF-become ADJ-kind
'became kind'

b. **naging t-um-akbo*
ASP.PERF-become ASP.PERF-run

Baker's (2005) third test for adjectivehood—that adjectives serve as resultative predicates—holds as well for Tagalog. Verbs don't serve as resultative predicates.

(6) *p-in-a-laki-ng*
MODE.CAUSE-ASP.PERF-grow-LNK
'reared/raised (to be) kind'

ma-bait (predicate)
ADJ-kind

The following two derivations rely on derived verb and adjective forms as input. The first example shows a derivation through stress shift from a fully formed derived verb (*sibak-ín*) into an adjective (*síbāk-in*). The patient semantics of the derived adjective came from the derived verb because the objective voice affix that gives rise to patient semantics in the adjective form can only be inserted initially into a verb as part of voice morphology.

(7) *sibák* (root) > *sibak-ín* (V) > *síbāk-in* (ADJ)
chop — chop-VCE.OBJ 'to chop up' — chop- ROLE.OBJ 'ready to be chopped up'

In the example below the derivation is from a fully formed adjective stem (*ma-samá`*) into a verb (*m-in-a-samá`;* the grave accent represents the glottal stop).

(8) *samá`* (root) > *ma-samá`* (ADJ) > *m-in-a-samá-ø-`* (V)
bad — ADJ-bad 'bad' — ASP.PERF-bad-VCE.OBJ "to attribute the quality 'bad'"

But whether adjectives are a distinct word class at the level of verbs, nouns, prepositions[3], and the like, or whether adjectives and verbs are subclasses under a super category, they undergo distinct morphosyntactic processes from verbs. We want to show that at an abstract level, such seemingly category-defining processes have the same morphosyntactic functions.

2 Verb affixation

Tagalog verbs carry voice and aspect affixes. These subcategories do not, however, account for all of the affixes that occur in many derived verbs. Occasionally, when providing the significance of a cluster of affixes, mention is made of "mode" but mode as a true subcategory of verbs never attracted mainstream attention. Rather, descriptive work on the verb affix system of Tagalog inevitably relies on listing affix combinations that include voice-mode hybrids. Schachter & Otanes (1972) identified 162 simple affixes and complex voice-mode affixes; examples are shown below. The affix *ipa*-, for example, consists of the applicative voice affix *i*- and the causative affix *pa*-, and *ipa*- itself occurs in combination with other affixes, for example, with the bound requestive affix -*ki*-as in *ipaki-,* or the bound deliberative affix -*ag*- as in *ipa+ag>ipag*-.

Verb affixes

i-*, -in-, -um-, -an, -in, pa-, ipa-, ipag-, ipang-, pa-…-an, pag-…-an, pang-…-an, ka-…-an, pang-…in, mag-, ma-, mang-, maka-, ika-, ipaki-, ipakipag-, ikapag-, ikapagpa-, ikapang-, ikapapang-, ipagpa-, ipapang-, ma...-an, makapag-, makapagpa-, makapagpaka-, makapag-...-um-, makapang-, maki-, makipag-, makipagpa-, makipang-…*

As the (incomplete) list above shows, simple affixes like *i-, pa-,-ka-*, *-ag-*, *-ang-*, *ma-* and others occur in many combinations. Thus -*ag*- is a part of *pag*- and *mag*-, and *pag*- in turn occurs in *ipag*-, *makipag*-, *ipakipag*-, and so forth, suggesting a potentially unfactored category. Wolfenden (1961) classified some of these segments under mode; however, he also maintained a set of voice-mode combinations[4]. The present study completely separates voice and mode. Also contributing to the perceived complexity of the verbal affix system are voice-aspect portmanteau units, for example, the infix -*um*- is treated as a marker of both the agent voice and the initiated aspect; we will show that these forms serve solely as aspect markers. With three independent categories—voice, aspect, and mode—scaffolding the verb affix system, a clean system emerges.

---

[3] Gil (1995) downplayed the role of morphology in determining word classes: "…since in Tagalog there are no distinct syntactic categories, morphological word classes are partly arbitrary and partly motivated by semantic factors" (p. 81).

[4] Wolfenden's (1961) five mode affixes are ø (indicative), *ka*- (inversive, reciprocal), *pa*- (injunctive), *pang*- (iterative, habitual) and *pag*- (comprehensive). The five voice-modes are objective (-*in*), locative (-*an*), implicative (*i*-), aptative (-*a*-), and subjective, which has three forms: -*ag*- (comprehensive), -*um*- (casual), and -*ang*- (iterative, habitual).

2.1 Voice and aspect

Voice refers to the affix in the verb that imposes a thematic role reading on the subject; 'subject' here refers to the absolutive nominal. For example, the suffix *-in* in *bilihin* is patient voice affix, which expects that the absolutively marked argument satisfies a patient, theme, or goal reading. The following table shows the standard analysis of Tagalog voice (Schachter & Otanes 1972, Ramos 1971).

Tagalog voice

Voice affixes	Roles of the subject nominal	Cover term
-um-	agent, experiencer, force	
ma-[5], *mag-, mang-*	agent	agentive, actor
-in	patient, theme, goal	objective, object
-an	location, directional (source, goal)	locative
i-	theme, benefactor, instrument, reason	applicative

As can be seen, the agentive voice affixes are *-um-, ma-, mag-*, and *mang-*. These voice affixes are also said to mark the initiation component of aspect. Initiation is indicated as [±beg(un)], as shown in the table below. The second component of aspect is completion, marked as [±fin(ished)] and expressed as CV_R (reduplication of the first CV of the stem). Presence of CV_R means the action is finished, and its absence indicates the action is unfinished.

Tagalog aspect

Aspect name	Aspect markers in the verb Nasal affix*plus*CV-reduplication of stem	
Perfective	*-um-, -in-, na-, nag-, nang-* [+beg]	+ ø (no CV_R) [+fin]
Imperfective	*-um-, -in-, na-, nag-, nang-* [+beg]	+ CV_R [-fin]
Contemplative	ø [-beg]	+ CV_R [-fin]

For example:

(9) *B-um-ibili* *si Pao* *ng niyog.*
ASP.[+BEG]-ASP.[-FIN]-buy DET.ABS Pao DET.OBJ coconut
'Pao is selling (some) coconuts.'

The nasal affixes occur in agentive voice forms, with the exception of the infix *-in-*, which occurs in non-agentive verb forms (i.e., the objective, locative, and applicative forms), as in *b-in-ili* 'was bought'. We seize this opportunity to propose that the nasal affixes have nothing to do with voice.

We propose that the nasal affixes, represented as the morpheme *m-*,expresses aspect, where *m-*, with allomorph *-um-*, is [-beg] and *n-*, with allomorph *-in-*, is [+beg], as shown

[5] This refers to the *ma-* actor focus affix (Schachter & Otanes 1972:288), for example, *maligo* 'bathe', *matuto* 'learn', *makinig* 'listen'.

in the following table. The underlined segments in *ma-, mag-*, and *mang-* we treat as part of the stem, as will be discussed shortly.

Revised analysis of aspect affixation

Aspect markers					
Nasal affix *plus* CV_R of stem			**Example: Root word *bili* 'buy'**		
Aspect (Non-agentive verbs)			**Objective**	**Locative**	**Applicative**
Command/infinitive	ø [-beg] +	ø [-fin]	bilih-in	bilih-an	i-bili
Perfective	(-i)n- [+beg] +	ø [+fin]	b-*in*-ili	b-*in*-ilh-an	i-b-*in*-ili
Imperfective	(-i)n- [+beg] +	CV_R [-fin]	*b-in-i*-bili	*b-in-i*-bilh-an	i-*b-in-i*-bili
Contemplative	ø [-beg] +	CV_R [-fin]	*bi*-bilihin	*bi*-bilh-an	i-*bi*-bili

Aspect(Agentive*m*-verbs)Example: Root word *bili* 'sell'			
Command/infinitive	(-u)m- [-beg] +	ø [-fin]	*m*-ag-bili
Perfective	(-i)n- [+beg] +	ø [+fin]	*n*-ag-bili
Imperfective	(-i)n- [+beg] +	CV_R [-fin]	*n*-ag-*bi*-bili
Contemplative	(-u)m- [-beg] +	CV_R [-fin]	*m*-ag-*bi*-bili

Aspect(Agentive -*um*-verbs)Example: Root word *bili* 'buy'			
Command/infinitive	(-u)m-, ø [-beg] +	ø [-fin]	b-*um*-ili, bili
Perfective	(-u)m- [+beg] +	ø [+fin]	b-*um*-ili
Imperfective	(-u)m- [+beg] +	CV_R [-fin]	*b-um-i*-bili
Contemplative	ø [-beg] +	CV_R [-fin]	*bi*-bili

The aspect paradigm for non-agentive verbs and the active *mag*-verbs obey the pattern for using the nasal affix and CV_R to express the different aspect instances. The active-*um*-verbs, on the other hand, disobey the alternation *m-/n-* ([-beg]/[+beg]) in that the forms *b-um-li* 'buy, bought' and *b-um-i-bili* 'is/was buying' both express initiated action but yet are marked for *-um-*. These two exceptional forms prevent the construction of a perfect paradigm for aspect in *-um*-verbs with respect to initiation. This imperfection does not obscure the overall tendency to use the nasal affix as aspect marker.

With no overt marker, the active voice is assigned the null voice form[6]. The voice affixes and their corresponding roles are shown below.

Revised list of voice affixes

Voice affixes	**Voice and role labels**	**Semantic roles of subjects**
ø	agent, agentive	agent, experiencer, force
-in	object, objective	patient, theme, goal
-an	location	location, directional (source, goal)
i-	applicative	theme, benefactor, instrument, reason

[6] The null verb voice affix is of course to be distinguished from the null form of the objective voice affix *-in* in the perfective and imperfective aspect, for example: *b-in-ili-ø* 'was bought', *b-in-i-bili-ø* 'is/was being bought'.

We assume that the forms *-um-* and *m-* are allomorphs of the nasal affix. They are attached to the first vowel of the root or stem, likely in observance of the Sonority Hierarchy, as shown below. Parallel processes apply to (*-i*)*n-*.

	Aspect	**Stem**	**(-*u*)*m*-+Stem**
(10)	-um-	biyak	b-*um*-iyak
(11)	m-	-agbiyak	*m*-agbiyak

So far, we have simplified the inventory of voice and aspect affixes into two non-overlapping categories. We will now propose that the rest of the affixes in derived verbs belong to a third category: mode.

2.2 Mode

Verb mode relates to the manner of the action. In the first example below, *mang-(k)uha*, qualifies the action with the meaning of, roughly, 'extensive, intensive.' The affix *-ang-* represents this comprehensive action. The second form has no pronunceable mode representation; this is the basic, 'casual' mode (Wolfenden 1961).

	Voice	**Aspect**	**Mode**	**Root**	**Derived word**
(12)	ø	m-	-ang-	kuha 'get, take'	> m-*ang*-uha
(13)	ø	-um-	ø	kuha'get, take'	> k-*um*-uha

In the pair below, the affix *-ag-* in *m-ag-biyak* calls for deliberateness in the action, as compared to the casual *b-um-iyak* "to split open".

	Voice	**Aspect**	**Mode**	**Root**	**Derived word**
(14)	ø	-um-	ø	biyak 'split open'	> b-*um*-iyak
(15)	ø	m-	-ag-	biyak 'split open'	> m-*ag*-biyak

The following is a list of mode affixes.

Mode affixes

Affixes	**Significance**	**Examples**	**Root**
ø	casual	ø-b-um-ili, ø-b-in-ili	bili 'buy'
-a(ka)-[7]	stative accomplish (possibility, aptative, accidental/non-volitional)	m-*a*-tulog m-*a*-bilí, m-*a*-kúha[8] ma-*ka*-kuha	tulog 'sleep' bilí 'buy', kúha 'get'
-ag-	deliberate	ip~~a~~-*ag*-bili m-*ag*-bigay	bili 'sell' bigay 'give'

[7] The enclosing hyphens in this and other similarly marked affixes in the list indicate that they are bound prefixes, not that they are infixes, of which there are only two in the language: *-um-* and *-in-*.

[8] 'Accomplish' mode may also be expressed with stress shift, as shown in the continuatation of the table below.

-an[9]	comprehensive reciprocal extensive	nagtawan-*an* nagsuntuk-*an* ipagpasikat-*an*	tawa 'laugh' suntok 'hit with fist' sikat 'shine'
-ang-	extensive across time(habitual, reservational, instrumental); or intensive in effort	ip~~a~~-*ang*-tulog m-*ang*-isda	tulog 'sleep' isda 'fish'
-ka-	reason, motivation unexpected	i-*ka*-bili nag-*ka*-sakit	bili 'buy' sakit 'ill'
-kang(da)-	out of control	mag-*kangda*-hulog	hulog 'fall down'
-ki-	deferential, requestive	pa-*ki*-kuha	kuha 'get, take'
-nga-[10]	simultaneous	na-*nga*-ag[11]-uwian	uwi 'go home'
pa-	injunctive, causative 'caused by circumstances'	*pa*-kuha na-*pa*-iyak	kuha 'get, take' iyak 'cry'
-sa-	directional	pa-*sa*-Maynila	sa-Manila 'to Manila'
-si-[12]	one by one, individually	mag-*si*-kuha	kuha 'get, take'
-ti-	resignedly, self-assistive	magpa-*ti*-hulog	hulog 'drop'
-um-[13]	Energetic	magt-*um*-akbo	takbo 'run'

Stress shift with the *-ag-* mode affix expresses intensiveness, as in the example below. Stress shift with the *-a-* mode affix, which is accompanied by compensatory length of the previously stressed vowel, expresses the 'accomplish' mode, as shown in the second row. These stress-shifted accomplish forms are synonymous with the regular forms.

-ag-+Stress Shift	intensive	m-*ag*-kaín	káin 'eat'
-a-+Stress Shift	accomplish (possibility, aptative, accidental/non-volitional)	m-*á*-bilī, m-*á*-kūha	bilí 'buy', kúha 'get'

Some forms of eduplication also express mode.

Affixes	Significance	Examples	Root
DW_r	conditional	*bumili*ng-bumili	bili 'buy'
PW_r	moderative disapprobational occasional	ma*busi*-busisi *b*um*ili*-bili *bali*-balitaan	busisi 'annoy, disturb' bili 'buy' balita 'news'
RW_r	continuative	*salita* nang salita	salita 'speak'
DW_r +Stress Shift	'able to with much effort'	*mábili*ng-mábili	bili 'buy'
PW_r + Stress Shift	regrettably	má*bili*-bili	bilí 'buy'

9 If the suffix *-an* in the reciprocal *naggupit-an* "cut each other's hair" is the mode affix, what is the reciprocal form's voice affix? One point of view is to treat *-an* as a simplification of the complex suffix *-anan* in the putative form **naggupit-an-an* > *naggupitan,* where the first *-an* is mode and the second is voice. This will be the only instance where mode occurs in suffix position. Alternatively, the reciprocal sentence may be a conflation of two instances of the locative-directional sentence, for example: *Ginupitan ni Ben si Og* at *Ginupitan ni Og si Ben* 'Ben cut Og's hair and Og cut Ben's hair'>*Naggupitan sina Ben at Og* 'Ben and Og cut each other's hair' where the reciprocity relation is a constructional meaning.

10 The affix *-nga-* is traditionally treated as a plural marker in the verb.

11 *-nga-* + -ag- > *-ngag-*

12 The affix *-si-* is traditionally treated as a plural marker in the verb.

13 An interesting form, in that the normally aspect affix *-um-* serves as a mode affix.

Abbreviations: DW_R = derived word reduplication, PW_R = partial word reduplication, RW_R = root word reduplication. Partial word reduplication applies on derived words and it duplicates the first CVCV of the root, which may be a partial of the root, as in *bali-balitaan*, or the full root, as in *biling-bili.*

As shown, mode affixes combine to form complex derived forms; the combinations and the order of affixes are still not fully determined. Modes combine selectively with particular verb root classes, again in ways that have not yet been worked out. Mode produces lexemes, in contrast to voice and aspect, which produce word-forms.

2.3 Components of derived verbs

Again, here are the components of derived verbs.

Voice	Aspect	Mode	Root
{ø, -in, -an, i-}	m-+ CV_R	{ø, -a-, pa-, -ag-, -an,-ang-, -ka-, -ki-, -ti, -sa-, -kang(da)-, -si-,… }	bili 'buy', etc.

The voice affixes *-in* and *-an* are suffixes. The applicative voice affix *i-* occupies a prefix position. For example:

(16) *i-bili*
VCE.APPL-buy
'to buy'

(17) *i-k-in-a-bili*
VCE.APPL-ASP.[+BEG]-MODE.REASON-buy
"reason-'ed' to have bought"

The applicative voice *i-* allows the accomplish mode affix *-a-* to its left (the only voice affix so allowed), to which in turn aspect *m-* attaches, as shown in (18)..The applicative voice undergoes CV_R to express incompleted aspect, as in (19).

(18) *n-a-i-pa-bili*
ASP.[+BEG]-MODE.ACCOM-ASP.[+FIN]-VCE.APPL-MODE.CAUSE-bili
'was able to have caused (x) to buy (y)'

(19) *n-a-i-i-pa-bili*
ASP.[+BEG]-MODE.ACCOM-ASP.[-FIN]-VCE.APPL-MODE.CAUSE-bili
'is being able to cause (x) to buy (y)'

Voice, Aspect, and Mode are heads of functional projections. Mode occupies the head position of the shell vP (following Larson 1988). For example:

(20) $[_{vP'}\ v_{Mode}\ VP]$

There can be more than one vP, thus:

(21) $[_{vP'}\ v_{Mode}\ [_{vP'}\ v_{Mode}\ VP]]$

vP is merged under AspP', and AspP' is a component of IP, thus (in X-bar notation; see Jackendoff 1977, Chomsky 1970):

(22) i-pa-*p~~a~~-ag*-bili
VCE.APPL-ASP.CONTEMP-MODE.CAUSE-MODE.DELIB-bili
'will be caused to be bought'

The *pa-* in *pa-ag-* (>*pag*) is an internal (or 'weak') causative, which is not a valence-increasing mode. The external ('strong') causative *pa-*, in contrast, increases valence; it takes an agent nominal argument, for example:

(23) N-a-pa-uwi
ASP.[+BEG].ASP.[+FIN].MODE.ACCOM-MODE.CAUSE.come/go.home

ni Ampi si Ben.
DET.GEN.AGT Ampi DET.ABS.PAT Ben
'Ampi was able to cause Ben to come home.'

Below is a complex derived word on the root *sigaw* 'shout'. The abbreviations A, M, R, and V refer to aspect, mode, root, and voice (in this gloss only), respectively. The *-um-* infix in the stem s-*um-igaw*, normally an aspect affix, appears to have been 'modalized' and given the meaning of 'exerting effort or energy'.

(24) n a i i ka pa ki p~~a~~ ag s um igaw an
A M A V M M M M M M Root M
'is being able to socially comprehensively shout off energetically'

Other roots that may be so derived are *utot* 'fart', *ire* 'groan', *suka* 'throw up', *iyak* 'cry', *dighay* 'belch', and, indeed, roots that in the context of the utterance can be construed to denote an activity that is performed socially and in an energetic matter!

In summary, we have simplified the nasal affixes *-um-, ma-, mag-, mang-* and *-in-* into one form: *m-*, designated it as the initiation marker of aspect, and assigned *-a-, -ag-* and *-ang-* to mode. We suggested that the actor voice affix takes the null variant form, confirmed Wolfenden's analysis of mode as an independent subcategory, separated mode from voice-mode hybrids, and confined verb lexeme formation in mode derivation. We noted that it has not yet been worked out which root classes take which mode affixes. We also showed that a derived verb may have more than one mode affix, but that the allowable

combination of modes and their order with respect to one another have not yet been mapped out.

What are the functions of verb affixation? First, the existential function: a way of creating verbs; this is accomplished by the affix voice. Then an anchoring function which fixes the semantic role of the subject, also performed by the affix voice; a duration function, which describes the extent of the action across time, expressed as aspect, and finally a qualifying function, expressed as manner or mode of the action. We will show that adjective affixation can be similarly analyzed.

3 Adjectives

Derived adjectives consist of a word category-forming affix and affixes that elaborate on the quality expressed in the root, such as plurality, degree, number and, we will propose, mode, thus:

The affix *ma-* is the major adjectivilizer. *Ma-* requires an absolutive/subject, thus:

(25) *Ma*-bait si Ben.
ADJ-kind DET.ABS-Ben
'Ben is kind.'

A number of derived adjectives do not show the *ma-* prefix. They carry affixes that are subcategories of the adjective class, enough credential to appear in adjective positions. For example, the affix *ka-* in (a) below is a degree affix and *sing-* in (b) is a comparative affix. Stress shift also creates an adjective; thus in (c), *gápāsin* 'ready for harvest' is converted from the verb *gapásin* 'to harvest' through stress shift and compensatory length.

(26) a. *Ka-bait* *ni Ben.*
DEG.INTENSE-kind DET.GEN Ben
'Ben is very kind.'

b. *Ka-sing-bait* *ni Ben* *si Pao.*
ADJ-DEG.INTENSE-DEG.COMPARE-kind DET.GEN Ben DET.ABS Pao
'Pao is as kind as Ben.'

c. *Gápas-in* *na* *ang palay.*
harvest-MODE ADV DET.ABS rice
'The rice is ready to harvest.'

Notice the parallel construction between derived verbs and derived adjectives in terms of the requirement for a subject. Thus the presence of the category-forming affix null voice affix ø and adjectivalizer *ma-* in (28-29) is accompanied by a subject.

Verb voice requires a subject

(27) *B-um-ait-ø* *si Ben.*
ASP.PERF-kind-VOICE.AGT DET.ABS Ben
'Ben became kind.'

Adjectivalizer ma- *requires a subject*

(28) *Ma-bait* *si Ben.*
ADJ-kind DET.ABS Ben
'Ben is kind.'

The absence of the word category markers of voice and adjectivalizer allows constructions unmarked for subject, thus:

(29) *Ka-u-uwi*[14] *nila.*
ASP.[+BEG]-ASP.CV_R-arrive.home 3PL.GEN
'They just got home.'

(30) *Ka-bait* *nila.*
DEG.INTENSE-kind 3PL.GEN
'They are very kind.'

But mere absence of the adjectivalizer *ma*-does not guarantee a subjectless construction. In comparative constructions, as in (27b) above, the comparee is subject, and in adjective predicates derived from verbal constructions via stress shift, as in (27c), the subject of the verbal construction is carried over.

When both *ma*- and *ka*- appear in the same adjective derivation, the first affix in sequence controls subjecthood. If *ma*- precedes *ka*-, as in (32) below, a subject is required. Otherwise, no subject is required, as in (33).

(31) *Ma-ka-ama* *si Ben.*
ADJ-MODE-father DET.ABS Ben
'Ben is close to his father.'

(32) *Ka-ma-pag-mahal* *ni Ben*
DEG.INTENSE.-ADJ-MODE-love DET.GEN-Ben
'Ben is very loving.'

Plurality is expressed using CVR, though its presence is optional.

---

[14] *ka-u-uwi* 'just got home' is an example of the so-called recent-perfective verb form (Schachter & Otanes 1972). This is a construction that is highly irregular. The form is used to express completed action, but its use of CV_R formally marks its action as incompleted. A completed action is additionaly marked by the begun affix (*n*-), but this affix is missing in this form. *ka*- hardly qualifies as an aspect affix. There is no voice nor mode marker. We make a distinction between the null voice allomorph in verb forms with fully formed aspect, that requires a subject (as in *bumili-ø*), and its absence in the exceptional recent-perfective form, which does not require a subject.

(33) *Ma-(ba)-bait* *sila.*
ADJ-(CVR)-kind 3PL.ABS-they
'They are kind.'

Degree indicates extent of the quality expressed in the root, as illustrated in the following table.

Adjective Degrees

moderative /dimunitive	ma*bait*-bait	'sort of kind'
positive	*ma*bait	'kind'
equal degree	(*magka)sing*bait	'as kind as'
unequal degree	*mas* mabait *kaysa*	'kinder than'
intensive	(*napa*)*ka*bait	'very kind'
superlative	*pinaka*mabait	'kindest'

A set of affixes elaborates on the quality expressed in the root; we call these mode affixes. (*P*- in *pag*- and *pang*- is unaccounted for.)

Mode Affixes

Affixes	Significance	Example	Root
-al-	'partaking of the nature of'	m*al*atenga	teynga 'ear'
-ka-	'favoring, leaning towards'	ma*ka*-ama	ama 'father'
ka-	'same as'	*ka*mata	mata 'eye'
-ag-	'intensively'	map*ag*mahal	mahal 'love'
-ang-	'habitually'	map*ang*hiram	hiram 'borrow'
stress shift + -*in*	'moderatively', 'tending to'	mahíyā*in*	hiyá 'shy'
stress shift + -*an*	'intensively', 'intensively used for'	málakās*an,* gulay*án*	malakás 'strong', gulay 'vegetable'
(gang)ga-	'same dimension as'	*ga*munggo	munggo 'mongo'
pa-	'purposely for'	*pa*damit	damit 'cloth'

Stress shift on verbs suffixed with -*in* produces adjectives. For example:

(34) *gápas* + *in* > *gapás-in* (V) > *gápās-in* (A)
'harvest' 'to harvest' 'ready to be harvested'

Some forms of reduplication indicate mode:

	Significance	**Example**	**Root**
DWr	'very'	*masagana*ng-masagana 'very abundant'	sagana 'abundant'
RWr	'very'	*sagana*ng-sagana 'very abundant' *bili*ng-bili	sagana 'abundant' bili 'buy'

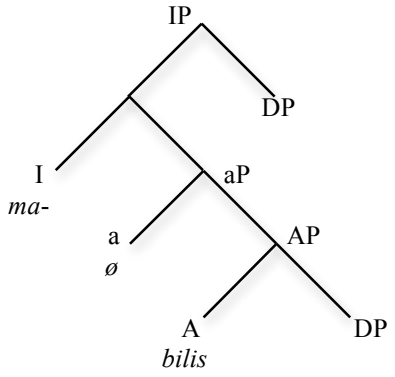
		'eager to buy'	
PWr	'moderately'	ma*saga*-sagana 'moderately abundant' ma*bili*-bili 'moderately saleable'	sagana 'abundant' bili 'buy'

Abbreviations: DWR = full derived word reduplication, reduplication, RWR = full root word reduplication, PWR = partial word reduplication.

3.1 Structure of the derived adjective

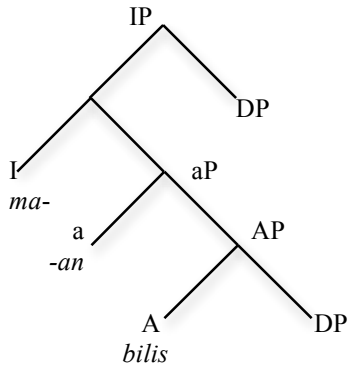
Following our analysis of verb affixes as heads of functional projections, we treat the adjectivalizer, mode, number, and degree affixes as heads of their own projections. As with verb mode, which is a small verb and occupies the head of a vP position, adjective mode is the head of a small adjective aP. Likewise, in simple derived adjectives, as in simple derived verbs, mode is null, thus:

(35) *ma-ø-bilis*
ADJ-MODE.*ø*-fast
'fast'



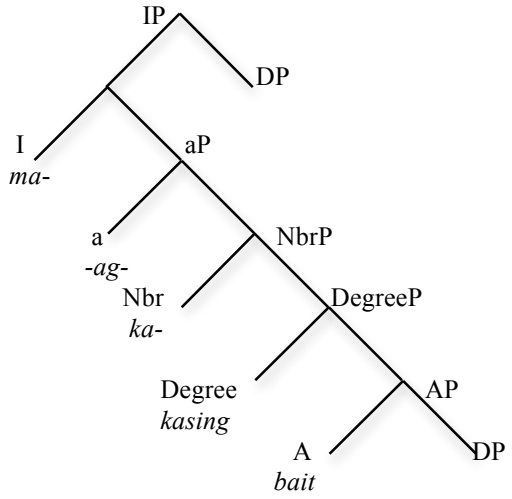
Below is an example of an adjective with a non-null mode, in this case, the intensive mode -*an*:

(36) ma-bilis-*an*
ADJ-fast-MODE.INTENSE
'intensified fast pace'



All the adjective affix components (adjectivalizer, mode, number, and degree affixes) are shown in the following form.

(37) m~~a~~-ag-ka-ka-sing-bait
ADJ-COMPREHENSIVE-NBR.[PLURAL]-DEG.INTENSE-kind
'equally kind'



Below is a summary of the characteristics of verb and adjective affixation.

	Verbs	**Adjectives**
(i)	Voice selects for subjects *Bilih-<u>in</u> mo ang basi.*	Adjectivalizer *ma-* selects for subjects *<u>Ma</u>-bait si Ben.*
(ii)	Voiceless verbs are subjectless *Katatapos-øni Ben.*	Null adjectivalizer requires no subject *ø-Kabait ni Ben.*
(iii)	Verb subcategories: mode, aspect, voice	Adjective subcategories: mode, number, degree

(iv)	Verb mode adds adverb-like property to action	Adjective mode adds adverb-like property to quality
(v)	Verb mode is derivational, aspect and voice are inflectional	Adjective mode is derivational, number and degree are inflectional

On the last point above, we realize that there has been a discussion on whether verb voice is inflectional (de Guzman 1997) or derivational (Starosta 2002), or the possibility that they may be both (Reid 1992). We believe that voice is inflectional. When the complex 'hybrid voice-mode' (Wolfenden 1961) affixes such as *ipa- magpaki-*, *maikapagpaki-* and such are stripped of the mode components, leaving the true voice affixes *-an*, *-in*, *i-* and *ø*, the inflectional nature of these affixes become apparent.

4 Conclusions

Verbs and adjectives exhibit parallel structural properties that perform essentially similar morphosyntactic functions. First, the existential function: a way of creating members of the class verbs and adjectives, accomplished by the voice and adjectivalizer, respectively. Second, an anchoring function, also performed by the voice affix and the adjectivalizer affix, which ties the verb to an argument—the subject. This relation fixes the semantic role of the subject of the verb, or of the attributee of the quality. Third, an "extent" affix expressed in verbs as aspect-duration and in adjectives as degree of the attribute referred to. Finally, a recursive mechanism for lexeme formation, in the form of a set of modification or qualification affixes, expressed as mode of the verb action and of the quality of the adjective.

References

Baker, Mark. 2005. *Lexical categories: Verbs, nouns and adjectives*. Cambridge: Cambridge University Press.

Chomsky, Noam (1970). Remarks on nominalization. In: R. Jacobs and P. Rosenbaum (eds.) *Reading in English Transformational Grammar*, 184-221. Waltham: Ginn.

Jackendoff, Ray (1977). *X-bar-Syntax: A Study of Phrase Structure*. Linguistic Inquiry Monograph 2. Cambridge, MA: MIT Press.

de Guzman, Videa P. 1996. Lexical categories in Tagalog. Pan-Asiatic Linguistics: *Procedings of the Fourth International Symposium on Languages and Linguistics, Vol. 1.* Mahidol University, Bangkok.

de Guzman, Videa P. 1997. Verbal affixes in Tagalog: Inflection or derivation? In *Proceedings of the Seventh International Conference on Austronesian Linguistics,* eds. Cecilia Odé and Wim Stokhof, 303–25. Amsterdam: Rodopi.

Dixon, R.M.W. 2004. Adjective classes in typological perspective. *Adjective classes: A cross-linguistic typology*, ed. R.M.W. Dixon and Alexandra Aikhevald, 1-49. Oxford: Ofxord University Press.

Gil, David. 1995. Parts of speech in Tagalog. *Papers from the Third Annual Meeting of the Southeast Asian Linguistics Society*, ed. Mark Alves, 67-90. Tempe: Arizona State University.

Larson, Richard. 1988. On the double object construction. *Linguistic Inquiry* 19, 335-391.

Pearson, Hazel. 2010. On the presence of adjectives in Fijian and the universality of the category 'Adjective'. Paper read at the 17th conference of the Austronesian Formal Linguistics Association, Stony Brook University.

Ramos, Teresita V. 1971. *Tagalog structures*. PALI Language Texts: Philippines. Honolulu: University of Hawaii Press.

Reid, Lawrence L. 1992. On the development of the aspect system in some Philippine languages. *Oceanic Linguistics*, Vol. 31, No. 1. (Summer, 1992).

Starosta, Stanley. 2002. Austronesian 'focus' as derivation: Evidence from nominalization. *Language and Linguistics* 3 (2).

Schachter, Paul and Fe T. Otanes. 1972. *Tagalog reference grammar*. Los Angeles, Berkeley: UCLA Press.

Wolfenden, Elmer. 1961. *A restatement of Tagalog grammar.*Manila: Bureau of Printing.

11 *The Event Semantic Role of the Nasal Prefix in Jakarta Indonesian*

THOMAS J. CONNERS AND CLAUDIA M. BRUGMAN

1 Introduction

This paper explores the nature of the nasal prefix *N-* in Jakarta Indonesian (JI), in contrast to its function in standard varieties of Indonesian (SI). In this and other varieties of Malay, the nasal prefix has most often been analyzed to be the paradigmatic counterpart of the *di-* prefix. Together, these have been taken to mark complementary, opposing syntactic and semantic functions, variously described as voice (active vs. passive), focus (actor vs. patient), discourse role (topic vs. comment), etc. (see Wouk & Ross (2002) and references therein). In this paper we highlight aspects of those analyses which can be subsumed under a more general framework of transitivity, in the sense of Hopper & Thompson (1980). Both the SI and the JI patterns can be understood as systems expressing a three-way distinction in transitivity, differing on the set of properties associated with each.

Under the view presented here, the nasal prefix's function marking active voice, high agency, or actor focus receives a unified analysis in terms of the expression of a higher position on the transitivity cline. In this paper, we demonstrate that, in contrast to the function of the nasal prefix in SI, the nasal prefix in JI is not in paradigmatic opposition with *di-*, and further, that there has been a reanalysis of *N-* from a largely voice-marking function to a (syntactically restricted) marker of aspect. While in the synchronic language *N-* largely functions on the event semantics, it does however maintain some restricted voice function, detailed below. The aspectual functions of *N-* in JI can broadly be characterized as marking atelicity, and include a relatively broad semantic range of atelic aspectual functions.

The *di-* marker, in contrast, has undergone no such reanalysis, and maintains its primary function as a marker of voice alternations. We adopt an analysis of *di-* much in line with Gil's (2002) generalized voice marker analysis of the cognate form in Riau Indonesian. Gil's analysis of the nasal prefix in Riau is more compatible with the data from SI than that from JI.

In order to understand the shift in function of *N-* in JI, we frame a possible analysis in terms of Hopper & Thompson's discourse-based notion of transitivity, which delimits a range of properties that cross-linguistically tend to associate collectively with binary contrasts encodable on predicates. The particulars of this play out differently in different languages, but there is broad consistency in the set of properties associated with the dichotomous relations such as active vs. passive, ergative vs. absolutive, etc. Hopper & Thompson establish a cline based on a set of binary properties, with different constructions encoding collections of properties which contribute to an overall evaluation of the transitivity of the clause. We demonstrate that the verbal system in JI instantiates three distinct points along this cline of transitivity. We speculate that the reanalysis of *N-* has

Asia-Pacific Linguistics, 2014.

both altered the degree of transitivity encoded on the verb, and also privileged a different set of properties associated with this cline, resulting in a three-way contrast that differs from that of SI, and presumably also from an earlier state of an antecedent Malay variety. In order to demonstrate this opposition, we contrast the situation in JI with that in SI, where the cognate morpheme to JI *N*-, *meN*-, has not undergone this reanalysis and maintains what is largely a discourse function as a marker of voice.

The paper is structured as follows. First we briefly explain how voice works in SI, before discussing the cognate morphemes and constructions in JI. We adduce evidence demonstrating that in JI, the set of parameters determining the acceptable occurrence of *N*- and *di*- are no longer parallel, in contrast to the situation in SI. We then explore in greater detail the heretofore undescribed aspectual functions of *N*- in JI. Given the limitations of scope, the present paper focuses on the presentation and description of some aspectual phenomena in JI. To highlight how fully grammaticalized[1] this function has become, we look at the interaction of *N*- with lexical aspect markers. Finally, we present the analysis of *di*- as a generalized voice marker in the sense of Gil (2002), and place the JI system within the broader typology of transitivity of Hopper & Thompson, suggesting this as a fruitful way for understanding the shift of *N*- from a marker of voice to an encoder of aspect.

2 Data

The data for the current study come from elicitation sessions conducted at the Jakarta Field Station of the Max Planck Institute for Evolutionary Anthropology during the period 2009-2010, originally under the auspices of the Jakarta Valency project. We worked with four native speakers of Jakarta Indonesian, and coded alternations of 110 lexical items for a full range of valency alternations, including all permutations of verbal prefix, applicative suffixes, bare stems, and one-, two-, and three-place predicates.

Further data were collected from the JFS corpus of naturalistic Jakarta Indonesian, where each token of the 110 lexical types in the Jakarta Valency Project was classified. The results reported here therefore account for both naturalistic and elicited data.

3 Voice in Standard Indonesian

Basic voice in Standard Indonesian is characterized by a dual alternation between active and passive constructions, marked through both word order and verbal affixation, and encoding participant semantics. Tense, aspect, and mood are not encoded or marked on the verbal root. With few exceptions, verbal roots must appear with one of the voice markers, as shown in (1):[2]

---

[1] What we mean by *grammaticalization* here is the innovation of an aspect system as part of the inventory of grammatical morphemes in JI. We are not making any claims about diachronic shift from lexical to grammatical status of any morpheme.

[2] We leave aside here discussion of the passive *semu* or object fronting constructions of the type:

Kursi itu sudah Edwin buat
chair that PFCT Edwin make
'Edwin made the chair.'

While at first glance these appear to be bare stems, the preverbal agent is a clitic, demonstrated by its appearance between the otherwise adjacent modal and root, and through the use of the reduced first and second person proclitic forms *ku*- and *kau*- in these constructions. Further, Cole et al. (2006) argue

(1) a. *Edwin membuat kursi.*
Edwin me*N*.make chair
'Edwin made/makes/will make a chair.'

b. *Kursi dibuat Edwin.*
chair DI.make Edwin
'A chair was made/being made/will be made by Edwin.'

c. **Edwin buat kursi.*

d. **Kursi buat Edwin.*

The examples in (1c) and (1d) show that most verbal roots in SI are not acceptable in bare stem form. There is an exceptional and highly restricted class of high frequency verbs which freely occur in the bare stem, that is, without any further morphological marking. This class includes mostly intransitive predicates and predicates with cognate objects, such as *tidur* 'sleep', *mandi* 'take a bath', *minum* 'drink', *makan* 'eat' and *pergi* 'go'. In these few cases where a verb can occur in the bare stem, verbal prefixation is in fact ungrammatical, unless co-occurring with an applicative. The nasal prefix interacts in interesting ways with the applicative marker *–in* in JI; this is beyond the scope of the present paper.

(2) a. *Dalan pergi.*
Dalan go
'Dalan went/goes/will go.'

b. **Dalan mergi.*
Dalan *N*.go
'Dalan went/goes/will go.'

c. **Dalan dipergi.*

So in SI there is a basic voice alternation with verbs occurring marked either by *di-* or by me*N*-. This is not, however, the full story in SI[3].

SI is a symmetrical voice language (see e.g. Himmelman 2005): there is no demotion of agents in passive constructions. SI in fact encodes three points on the transitivity cline, as defined by Hopper & Thompson. We have already seen in detail the functions of the me*N*- and *di-* markers.

SI also has a verbal prefix *ber-*. This marker tends to mark predicates that have only one expressed argument (by that token, they are low on the transitivity scale):

(3) a. *Dalan berjalan.*
Dalan BER.walk
'Dalan walks.'

b. *Ayam bertelur*
chicken BER.egg
'Chickens lay (eggs).'

---

convincingly that there is no *passive semu* in Jakarta Indonesian, and Nomoto (2006) arrives at the same conclusion for basilectal Malay. That situation makes a contrastive analysis of this phenomenon impossible.

[3] There are also voice alternations encoded by *ter-* and *ke-/ke-an* in SI, generally marking an adversative or accidental passive type construction. These are more restricted in their distribution, and while a proper analysis of voice in SI would include a discussion of these markers, they fall outside the narrower scope of the present paper. They will be mentioned later in reference to Hopper & Thompson 1980.

In SI, then, we need reference to three distinct values of transitivity, which are encoded with the three verbal prefixes: *meN-, di-* and *ber-*. The set of properties encoded by these markers extends beyond the number of participants, since both *meN-* and *di-* can mark two-place predicates. In SI, these properties include the level of agentivity and the affectedness of the object. The most parsimonious way to describe this opposition is in terms of Hopper & Thompson-style transitivity, since both the number of arguments and these properties of individual arguments characterize the opposition.

There is also a three-way distinction encoded on the verb in JI; however, it differs crucially in the set of properties encoded. We turn to voice in JI next.

4 Jakarta Indonesian

Jakarta Indonesian is the colloquial language of Jakarta; it is used for most inter-ethnic and ethnically neutral communication. Increasingly it is being used as a language of intra-ethic communication among recent migrants to the capital. In a multilingual setting, it is the one language which is native to most inhabitants of the city and surrounding area, acquired at home and in informal settings. It is further the second language of many millions of more recent migrants. Jakarta Indonesian encompasses a wide range of registers and variation, a state that is facilitated by the fact that is it not affiliated with any dominant ethnic group.[4] Increasingly, Jakarta Indonesian is imitated as a prestige variety throughout many parts of the archipelago, especially western Indonesia.

5 Verbal Alternations in Jakarta Indonesian

Verbal alternations in JI are signaled through two prefixes and the bare stem form, as well as the applicative suffix *-in,* which is not the focus here. The nasal prefix in JI is represented with *N-*; it is cognate with the SI *meN-,* though the particulars of the morphophonological alternations differ somewhat in the two varieties. JI also makes use of the prefix *di-*, which is cognate with SI *di-*.

In JI, SVO is the default preferred order for elements, though it is rare in naturalistic speech for all arguments to be expressed and when they are expressed, other orders are very common.

Verbs in JI undergo a three-way alternation different from that in SI, as demonstrated with the root *cuci* 'wash' in (4).

(4) a. *Adi* *nyuci* *baju.*
Adi *N.*wash clothes
'Adi washes the clothes.'

b. *Baju* *dicuci* *Adi.*
clothes DI.wash Adi
'The clothes were washed by Adi.'

---

[4] See Paauw (2009) on the difference between Betawi, the post-creole Malay variety native to the small indigenous population, and Jakarta Indonesian.

c. *Adi cuci baju.*
Adi wash clothes
'Adi washes the clothes.'

In (4a) the verb 'wash' is marked with the nasal prefix, *N*-, the agent appears preverbally, and the patient appears postverbally. This is the active or actor-focused construction, equivalent in form to the SI example in (1a).

In (4b) the verb is marked with *di-*, the agent appears postverbally, and the patient appears preverbally. This is the passive, or patient-focus construction, equivalent in form to the SI example in (1b).

Although the two examples in (4a and b) are formally equivalent to the SI examples in (1a and b), they are not functionally equivalent to them. The verb forms also differ in frequency between the two varieties.

In (4c) above, the verb appears in the bare stem form. There is no grammatical equivalent construction in SI, despite the fact that in JI this is by far the most frequent construction in naturalistic speech. Bare stems in JI are not defective, may be grammatically intransitive or transitive, and freely license second arguments. They are, further, not limited to a highly restricted class of verbs—as is the case in SI: most roots in JI can occur in the bare stem.

(5) a. *Bolanya gelinding.*
ball.NYA roll
'The ball rolls.'

b. *Nadia gelinding bola.*
Nadia roll *ball*
'Nadia rolls the ball.'

c. *Bola gelinding Nadia.*
ball roll Nadia
'The ball is rolled by Nadia.'

d. *Bola digelinding Nadia.*
ball DI.roll Nadia
'The ball is rolled by Nadia.'

In (5a) above, the intransitive verb 'roll' appears with its single theme argument. There is no marking on the verb. In (5b) the transitive equivalent of 'roll' appears with both agent and patient arguments expressed. Again, there is no marking on the verb. The arguments can appear in almost any order, and the passive English translation provided for (5c) is rather forced. The construction in (5d) with the *di-* marked verb highlights the salience of the patient; this is in fact the primary function of the *di-* prefix in JI. The distinction between (5c) and (5d) is accounted for under Gil's analysis of generalized voice markers, but is not our focus.

The order of elements, while interesting, is also not the main point here. The main point is that in JI verbs freely occur without any verbal affixation. This is true of intransitive, transitive, and even some inherently ditransitive roots:

(6) *Erni kasi gue buku.*
Erni give 1SG book
'Erni gives me a book.'

Such examples would all be ungrammatical in SI, yet are productive and in fact the most frequent constructions in JI.

The vast majority of lexical items that we tested in the current dataset fall into the pattern demonstrated in (4) above, showing a three-way contrast between a bare stem, a form marked with *N*-, and a form marked with *di*-. However, not all verbs are grammatical in all three forms. One pattern that emerged from the dataset includes forms where the bare stem and the form marked with *di*- are both grammatical, but the form marked with the nasal prefix is ungrammatical:

(7) a. *Tessa dorong gerobak.*
Tessa push cart
'Tessa pushes the cart.'

b. **Tessa ndorong gerobak.*
Tessa *N*.push cart
'Tessa pushes the cart.'

c. *Gerobak didorong Tessa.*
cart DI.push Tessa
'The cart is pushed by Tessa.'

This pattern, where the bare stem alternates with forms marked with *di*-, and where forms marked with *N*- are ungrammatical, occurred frequently in our dataset. It is particularly common with verbs such as *bawa* 'carry', *bakar* 'burn', *bangun* 'build', *dorong* 'push', *gelinding* 'roll', *giling* 'grind', *tinggal* 'stay', *bantu* 'help', *gali* 'dig', etc. The analysis presented here does not account for the ungrammaticality of the nasal prefix with this class of verbs.[5]

There are also several cases in the dataset where the bare stem alternates with a form marked with *N*-, and the form marked with *di*- is ungrammatical:

(8) a. *Gue panjat pohon.*
1SG climb tree
'I climb the tree.'

b. *Gue manjat pohon.*
1SG *N*.climb tree
'I climb the tree.'

c. **Pohon dipanjat gue.*
tree DI.climb 1SG
'The tree was climbed by me.'[6]

---

[5] Most of the verbs in this class have voiced stop initial consonants. Prenasalized voiced stops may be phonotactically dispreferred in this variety, because it is perceived by native speakers as Javanese.

[6] It is not the postverbal first person agent that renders this construction ungrammatical. Although such constructions are ungrammatical in SI, they are licit in JI, which as noted above does not have the passive *semu* construction for first and second person agents in passive-type phrases.

A smaller class of verbs shows a pattern where the prenasalized form is grammatical, but both the bare stem and the form marked with *di-* are ungrammatical:

(9) a. **Gue cium baunya.*[7]
1SG smell odor.NYA
'I smell the odor.'

b. *Gue nyium baunya.*
1SG *N*.smell odor.NYA
'I smell the odor.'

c. **Baunya dicium gue.*
odor.NYA DI.smell 1SG
'The odor was smelled by me.'

The roots *ajar* 'teach' and *omong* 'speak' follow a similar pattern.

These three verbal marking patterns show clearly that the parameters for the grammaticality of marking a verb with *di-* are clearly distinct from those determining the grammaticality of *N-*.

In fact, in the entire dataset, there is but a single example which maintains the pattern found in SI, where the bare stem is ungrammatical and the *N-* and *di-* marked forms stand in opposition. This is shown in (10):

(10) a. **Dalan tuang aer.*
Dalan pour water
'Dalan pours water.'

b. *Dalan nuang aer.*
Dalan *N*.pour water
'Dalan pours water.'

c. *Aer dituang Dalan.*
water DI.pour Dalan
'The water is poured by Dalan.'

Taken together, the examples in (3-9) clearly demonstrate that in JI, *N-* and *di-* do not stand in paradigmatic contrast—at least not on a single dimension of contrast. This observation about the variety can shed some light on some puzzling JI data first described in Gil (2006), where in naturalistic speech certain verb forms are found which are marked by both *N-* and *di-* simultaneously.

(11) a. *Adi dimukul Bowo.* (stem *pukul*)
Adi DI.*N*.hit Bowo
'Bowo hit Adi.'

b. *Diminggirin dulu rambutnya.* (stem *pinggir*)
DI.*N*.edge-APPL before hair.NYA
'Move your hair aside.'

If the parameters for the grammaticality of *N-* and *di-*vary independently, there is nothing barring them from co-occurring. Our analysis predicts this possibility.

---

[7] Note that with the meaning of 'kiss' *cium* occurs in the bare stem.

Gil (2002) argues that *di-* and *N-* in Riau Indonesian are 'weak generalized voice markers': they do not affect the syntactic valency of the verb but rather highlight the semantic salience of a particular argument. Modulo other varietal differences, we endorse the semantic level of Gil's analysis for JI as well, but show below how this does not fully account for the range of properties in the JI system.

In this section we have demonstrated that *N-* and *di-* in JI do not co-occur in the same range of constructions. We have further posited that *di-* maintains its largely voice-related function. What remains to be answered is: what role does *N-* play in JI? What we argue here is that JI *N-* has undergone a reanalysis from a voice-marking function to a (syntactically restricted) marker of aspect. In the next section we turn to a discussion of aspect in JI.

6 Aspect in Jakarta Indonesian

Now that we have demonstrated that *N-* and *di-* in JI do not form part of the same voice paradigm, we turn to explicating the role of the *N-* in JI. From the data to be adduced below, it is clear that *N-* has stepped out of its function as a voice marker, and plays an active role in encoding event structure. In particular, when in opposition to the bare stem, it marks a predicate as atelic.

Most of the languages of Western Indonesia do not have explicit non-adverbial markers of tense. They do, however, have aspect markers, most of which are periphrastic: they are independent modals or auxiliary verbs that mark a range of aspectual contrasts.

Most recent work on JI and other colloquial Indonesian and Malay varieties has focused on these periphrastic aspectual markers, including Ginsberg & Paauw (2010) and Grangé (2011). Hidayat (2010, 2011) makes a first attempt to explore the aspectual qualities of *N-* in JI. Specifically, she looks at possible uses of *N-* as a progressive marker. Ultimately, though, her results are inconclusive.

The proper analysis of *N-* in JI is primarily as a grammaticalized aspect marker, as opposed to a voice marker. However, in at least one kind of construction, a subordinate clause following a relativizer, the voice function is apparent, robust, and unambiguous. This is a possible indication of the earlier state of the language and provides the strongest evidence for an actual reanalysis, in the context of comparison of the role of the cognate items in the standard and other languages.

The aspectual role played by *N-* in the synchronic language is a generalized role encompassing a range of atelic interpretation possibilities, including habituality, progressivity, iterativity, and genericity. They are demonstrated in the following examples:

(12) a. *Mak cuci piring.* (telic)
Mom wash plate
'Mom washed the dishes.'

b. *Mak nyuci piring.* (habitual)
Mom *N.*wash plate
'Mom does dishes.'

c. *Piringnya dicuci maknye.* (aspectually unmarked)
plate.NYA DI.wash Mom.NYA
'The dishes were washed by Mom.'
'Mom washes the dishes.'

With the 'wash' class of verbs, the bare stem has a default telic interpretation. The addition of the nasal prefix renders a habitual interpretation, as in (12b) above.

Compare, however, the example in (12) with that in (13), which provides a parallel set, but with different default interpretations:

(13) a. *Nadia pake baju.* (aspectually unmarked)
Nadia use clothes
'Nadia wears clothes.'

b. *Nadia make baju.* (progressive)
Nadia *N.*use clothes
'Nadia is wearing clothes.'

c. *Baju dipake Nadia.* (aspectually unmarked)
clothes DI.use Nadia
'Clothes are worn by Nadia.'
'Nadia wears clothes.'

In example (13a) above, the verb appears as a bare stem: it is both unmarked morphologically, and as a default interpretation, it is also unspecified aspectually. The addition of the nasal prefix renders the default interpretation as progressive, as shown in (13b).

We note here that the *di-* prefix also does not alter the aspectual reading for this class. This is expected under the current account, as the *di-* prefix, as opposed to the nasal prefix, is primarily a marker of voice alternations in JI.

The nasal prefix can also denote iteration:

(14) a. *Gue lempar batu.* (aspectually unmarked)
1SG throw stone
'I throw a stone.'

b. *Gue nglempar batu.* (iterative)
1SG *N.*throw stone
'I throw stones (repeatedly).'

c. *Batunya dilempar gue.* (aspectually unmarked)
stone.NYA DI.throw 1SG
'The stone is thrown by me.'

It is important to point out that the nasal prefix has not lost all of its voice function:

(15) a. *yang nyolong…*
REL *N.*steal
'the one who did the stealing…'

b. *yang dicolong…*
REL DI.steal
'the thing that was stolen…'

c. *yang colong…*
REL steal
'the one who stole/the thing stolen…'

In a relative clause the verbal prefix clearly identifies focus on one or the other of the arguments (i.e. acts as a voice marker). Note that the bare stem is also possible in this construction in JI and in that case either argument is a possible target of relativization. The same pattern of prefixation in a relative clause is seen in SI and other Malay varieties. Given its widespread distribution, it is not unreasonable to hypothesize that this is a vestige of an earlier state of the language, when both *N*- and *di*- operated primarily as voice markers. In JI, the *N*- has undergone a reanalysis, possibly motivated partially by the acceptability and high frequency of the bare stem in JI.

7 Interaction with Aspectual Auxiliaries

As noted above, JI encodes most of its aspectual distinctions through lexical modals and auxiliaries. We might expect, therefore, that the shift of the *N*- prefix from a marker largely encoding voice distinctions to one largely encoding aspectual distinctions would result in potential co-occurrence restrictions. Specifically, we have posited that the *N*- plays a role as a broad a marker of atelicity. This would predict therefore that it should not co-occur with modals that are telic. This expectation is generally confirmed.

(16) a. ??*Gue* *lagi* *cari* *tas* *gue*. (progressive)
1SG PROG look.for bag 1SG
'I am looking for my bag.'

b. *Gue* *lagi* *nyari* *tas* *gue*.
1SG PROG *N*.look.for bag 1SG
'I am looking for my bag.'

In (16) the verb appears with the progressive modal *lagi*. We would expect the progressive marker to co-occur with the nasal prefix—a marker of atelicity—and indeed it does. In (16b) above the verb occurs marked with *N*-, and is strongly preferred over (16a), where the verb appears with the bare stem.

Similarly when combined with another inherently atelic modal, such as the habitual marker *suka*, there is a preference for the form of the predicate marked with *N*-:

(17) a. ?*Kucing* *suka* *tangkap* *tikus*. (habitual)
cat HAB catch rat
'Cats like to catch rats.'

b. *Kucing* *suka* *nangkap* *tikus*.
cat HAB *N*.catch rat
'Cats like to catch rats.'

In contrast, when co-occurring with the perfect marker, *(s)(u)dah,* an inherently telic modal, the bare stem form is strongly preferred, as in (18) below:

(18) a. *Erni* *dah* *kasi* *dia* *kado.* (perfect)
Erni PFCT give 3SG gift
'Erni already gave her a gift.'

b. ??*Erni dah ngasi dia kado.*
Erni PFCT *N*.give 3SG gift
'Erni already gave her a gift.'

8 Transitivity: A Shift in Alignment

Hopper & Thompson (1980) explore the notion of transitivity cross-linguistically. They find that certain clause-level properties cluster together to mark a clause as being of relatively higher or lower transitivity, and these patterns hold across languages. The basic pattern is presented below:

	HIGH TRANSITIVITY	**LOW TRANSITIVITY**
Participants	two or more	one
Kinesis	action	non-action [e.g. *like*]
Aspect	telic	atelic
Punctuality	punctual	non-punctual
Volitionality	volitional [of A]	non-volitional
Affirmation	affirmative	negative [e.g. *I didn't kick him*]
Mode	realis	irrealis
Agency	A high in potency [e.g. *Charles startled me* vs. *The picture startled me*]	A low in potency
Affectedness of O	O totally affected	O not affected
Individuation of O	O highly individuated	O non-individuated

Hopper & Thompson (1980:252)

These are the factors that are involved in determining the actual encoding of clauses as one or the other member in a binary distinction such as active vs. passive or ergative vs. absolutive, in languages that differentiate clauses binarily. Active and ergative clauses cross-linguistically tend to share properties of high transitivity, whereas those marked passive or absolutive tend to share properties of lower transitivity. This can be seen in the English active-passive alternation:

(19) a. John hit the ball.
b. The ball was hit (by John).
c. John hit balls.
d. ?Balls were hit (by John).

(19a), the active example, encodes most of the properties Hopper & Thompson identify with high transitivity: two arguments, telic action, punctuality, volitionality of the agent, etc. The passive counterpart of (19a), shown in (19b), is lower in transitivity in expressing only a single argument. While the active sentence is equally acceptable with either a highly individuated or a less individuated object (as shown in (19c)), there is something odd about (19d), the passive sentence with the less individuated passive subject, suggesting that it would require more loading of the discourse context to render it more acceptable.

Hopper & Thompson surveyed a number of languages to arrive at the clusters listed above. Critically, languages differ in the particular properties they encode in these contrasts, as well as in the form in which the binary contrast is expressed.

SI and JI both diverge from the Hopper & Thompson standard model in overtly differentiating clauses into three rather than two classes, based on the transitivity properties noted above. However the properties themselves cluster differently in the two varieties: in

SI, the relevant properties are those associated with A and O, while JI has apparently innovated attention to aspectuality as the privileged property.

As noted in Hopper & Thompson, SI *ber-* marks clauses of various low-transitivity types, including nominalizations, statives, and reflexives (see Hopper & Thompson 1980:278). What is notable is that these constructions strongly tend to have the low-transitivity values associated with participants, including a lack of referential specificity of participants, and a low differentiation between A and O. SI marks two additional points along the transitivity cline: the *di-* marker marks clauses of higher transitivity in comparison to *ber-*: there is greater differentiation of A and O; there are one or two arguments instead of only one—these are both associated with participants. SI *meN-* marks yet a third distinct point in the transitivity scale, preferentially encoding two arguments and functioning largely as a marker of voice. For each of these three markers, the most salient properties crucially involve the participants, and none indexes aspectual properties per se.

SI's *ber-* is lost in JI, but this does not result in a reduction of distinctions; three points in the transitivity scale are still distinguished, one of which is associated with the bare stem verb. Unlike in SI, which treats transitivity as a linear scale and distinguishes three classes in that it references and privileges the same set of properties, JI has added a dimension of aspectual profile to the properties of transitivity being attended to, and specifically privileges that property in only one of the markers. We speculate that the loss of *ber-* and the reorganization of the system are not unrelated: the loss of one marker allowed for the transitivity space to be reapportioned on two dimensions rather than one. We regard this as a topic ripe for historical and synchronic investigation. It is also intriguing in this context to consider the general shift from voice-marking systems in Western Austronesian languages to aspect-marking systems in Eastern Austronesian languages, having undergone at some point a similar reanalysis.

In order to understand verbal prefixation in JI, one must make reference to aspect; this is the major point of contrast between the systems of encoding in JI and SI. In JI, the *N-* prefix significantly encodes aspectual differences on the verb, while the *di-* marker focuses on the patient and the bare stem form represents a default or unmarked case.

The reanalysis that *N-* has undergone is not one of a marker of high to a marker of low transitivity. Rather, it has undergone a reordering of the salience of the properties associated with marking a clause as of higher transitivity.

9 Summary

There is a three-way transitivity distinction that is expressed in both SI and in JI. The two varieties of Indonesian show this distinction in different characters. The SI system operates on one dimension of distinction, namely number and discourse status of the participants. By contrast, JI has added a dimension of distinction: that of telicity within the event semantics. We have viewed this varietal distinction through the lens of Hopper & Thompson's notion of transitivity of clauses: this provides a unifying theme for each system internally, as well as a direction for the future more detailed analysis of the possible reanalysis of *N-* in JI. Our future work will explore in greater detail the classes of verbs with respect to acceptable derived forms and their interpretation, as well as a look at the discourse environment that would favor one or another of these forms.

References

Cole, Peter, Gabriella Hermon, and Yassir Tjung. 2006. Is there a Pasif Semu in Jakarta Indonesian? *Oceanic Linguistics* 45:1, 64-90.

Gil, David. 2002. The Prefixes *di-* and *N-* in Malay / Indonesian Dialects. In *The History and Typology of Western Austronesian Voice Systems*, (ed.) F. Wouk and M. Ross, 241-283. Pacific Linguistics.

Gil, David. 2006. The Acquisition of Voice Morphology in Jakarta Indonesian. In *The Acquisition of Verbs and Their Grammar: The Effect of Particular Languages* (ed.) N. Gagarina and I. Gülzow, 201-227. Springer.

Ginsberg, Elisabeth, and Scott Paauw. 2010. Verbal Aspect in Colloquial Jakarta Indonesian. Paper presented at ISLOJ 2: Lombok, Indonesia.

Grangé, Philippe. 2011. Aspect in Indonesian: free marker versus clitic markers. In Proceedings of the Workshop on Tense, Aspect, Mood, and Evidentiality in Indonesian Languages, RILCAA, Tokyo.
URL http://lingdy.aacore.jp/doc/indonesiatam/philippe_grange_p.pdf

Hidayat, Lanny. 2010. The Acquisition of Verb Argument Structure in Basilectal Jakarta Indonesian. PhD Dissertation, University of Delaware.

Hidayat, Lanny. 2011. The Analysis of the Semantic Function of the Prefix *N-* in Basilectal Jakarta Indonesian. In Proceedings of the Workshop on Tense, Aspect, Mood and Evidentiality in Indonesian Languages. RILCAA, Tokyo.
URL http://lingdy.aacore.jp/doc/indonesiatam/lanny_hidajat_p.pdf

Himmelmann, Nikolaus. 2005. The Austronesian languages of Asia and Madagascar: Typological overview. In *The Austronesian languages of South East Asia and Madagascar*, (ed.) K.A. Adelaar and N.P. Himmelmann, 110-181. Routledge.

Hopper, Paul J., and Sandra Thompson.1980. Transitivity in Grammar and Discourse. *Language* 56:2, 251-299.

Nomoto, Hiroki. 2006. Voice in Colloquial Malay Relatives. Ms. Tokyo University of Foreign Studies.

Paauw, Scott. 2009. The Malay Contact Varieties of Eastern Indonesia. PhD Dissertation, University of Buffalo.

Wouk, Fay, and Malcolm Ross. 2002. *The History and Typology of Western Austronesian Voice Systems.* Pacific Linguistics.

12 *Ditransitives and benefactives in Lamaholot*

NAONORI NAGAYA

1 Introduction[1]

Lamaholot, a Central Malayo-Polynesian language spoken on Flores Island of eastern Indonesia, has two similar yet distinct types of verb alternations: the ditransitive and the benefactive alternations. In these alternations, one and the same verb is used in two different constructions without a noticeable difference in meaning. See (1) and (2).[2]

(1) **Ditransitive alternation:**

a. **Prepositional-recipient construction:**

go neĩ buku ia Nia.
1SG give book LOC Nia
'I gave a book to Nia.'

b. **Ditransitive construction:**

go neĩ Nia buku.
1SG give Nia book
'I gave Nia a book.'

[1] An earlier version of this paper was presented at the 12th International Conference on Austronesian Linguistics held at Udayana University in Bali, Indonesia, on July 5, 2012. I am grateful to the audience, especially, Robert D. Van Valin, Jr. and Naomi Tsukida, for their comments and criticism that have helped in improving the manuscript. My thanks also go to Kazuya Inagaki, Ritsuko Kikusawa, Kunio Nishiyama, Hiroki Nomoto, Asako Shiohara, and Atsuko Utsumi, who have each provided constructive comments and suggestions on the manuscript. Of course, any errors that remain are my responsibility. This work was supported by the Japan Society for the Promotion of Science (Grant-in-Aid #25884018).

[2] Abbreviations used in this paper are: A-more agent-like argument in a mono- or di-transitive construction, CONJ-conjunction, DEM-demonstrative, DIR-directional, DIS-distal, EXC-exclusive, INC-inclusive, LOC-locative, NEG-negator, NMLZ-nominalization, P-more patient-like argument in a transitive construction, PL-plural, PO-primary object, PROS-prospective, PROX-proximal, REL-relativizer, S-single argument in an intransitive construction, SEA-seaward, SG-singular, SUBJ-subject, 1-first person, 2-second person, 3-third person, and "="-cliticization.

Asia-Pacific Linguistics, 2014.

(2) **Benefactive alternation:**

a. **Benefactive serial verb construction:**

go biho lama neĩ Ika.
1SG cook rice give Ika
'I cooked rice for Ika.'

b. **Benefactive construction:**

go biho Ika lama.
1SG cook Ika rice
'I cooked Ika rice.'

The examples in (1) represent a ditransitive alternation, where the verb of giving *neĩ* 'give' appears in two different constructions, each bearing the roughly equivalent meaning that the speaker transferred the ownership of a book to Nia. In the prepositional-recipient construction in (1a), the recipient participant *Nia* is introduced by the locative *ia*, while the theme participant *buku* 'book' appears directly after the verb. By contrast, in the ditransitive construction in (1b), both the recipient and the theme participants directly follow the verb, in that order, yielding a double object construction.

The examples in (2) demonstrate a benefactive alternation, where the verb of cooking *biho* 'cook' is used in two distinct syntactic environments. On the one hand, (2a) is a serial verb construction using *neĩ* 'give', which introduces a recipient-beneficiary participant. On the other hand, the benefactive construction in (2b) has the recipient-beneficiary and the patient participants appear, in that order, directly after the verb, while expressing the same situation that (2a) does.

In this paper, I present a description and analysis of these two kinds of verb alternations in Lamaholot. After providing a basic description of these alternations, I argue that these two alternations have the shared syntactic function of realigning the primary object grammatical relation from one argument to another. It is also shown that these two alternations do differ in many ways; in particular, in the ditransitive construction, a non-primary object NP still counts as a grammatical argument, but in the benefactive construction, a non-primary object NP does not work as such.

The remainder of this paper is arranged as follows. In Section 2, I provide preliminary information on the dialect of Lamaholot examined in this paper. Sections 3 and 4 introduce and analyze the ditransitive and the benefactive alternations, respectively. In Section 5, similarities and differences between the two alternations are discussed. I conclude in Section 6.

2 Preliminaries

Lamaholot is a Central Malayo-Polynesian language of the Austronesian language family. It is spoken in the eastern part of Flores Island and neighboring islands in eastern Indonesia, serving as the lingua franca of the region (Grimes et al. 1997). Lamaholot is best understood as a dialect chain, with enough substantial differences between some of the dialects so as to make them mutually incomprehensible (Keraf 1978; Bowden 2008; see also Nagaya 2010). In this description, I focus on the Lewotobi dialect, the most westerly dialect in the chain. This dialect is spoken by approximately 6,000 speakers in Kecamatan Ile Bura (Lewis et al. 2013).[3]

[3] Lewis et al. (2013) treat the Lewotobi dialect of Lamaholot as a distinct language named Lewotobi.

Lamaholot has two important typological characteristics. First, Lamaholot is nearly an isolating language, like other Flores languages (Himmelmann 2005; Arka 2007; Donohue 2007). Its grammatical formatives include: S/A-agreement prefixes (Table 1.1); S-agreement enclitics (Table 1.2); the third person singular primary object enclitic pronoun *=roʔ* (see Section 5); the nominalization/possessive markers *-N* and *=kə̃*; and the pronominal possessive marker *=ə̃ʔ*.

Table 1.1: S/A-agreement prefixes

	SG	PL
1	k-	m- (EXC)
		t- (INC)
2	m-	m-
3	n-	r-

Table 1.2: S-agreement enclitics

	SG	PL
1	=əʔ	=kə (EXC)
		=kə (INC)
2	=ko	=kə
3	=aʔ	=ka

Second, although transitional languages are predominant on Flores Island, Lamaholot is a typical instance of a preposed possessor language, showing a series of typological features characteristic of languages in eastern Indonesia (see Himmelmann 2005 for preposed possessor and transitional languages; cf. Klamer 2002; Donohue 2007; Musgrave 2008). The basic word order in Lamaholot is subject-verb-object, as shown in (3).

(3) **SVO word order:**
Hugo belo kadʒoʔ.
Hugo cut tree
'Hugo cut down a tree.'

There is person marking for S and A arguments, as illustrated in (4) and (5). S/A-agreement prefixes (Table 1.1) mark the person and number of either an intransitive or a transitive subject argument, while S-agreement enclitics (Table 1.2) indicate the same agreement information with an intransitive subject argument only. The former are used obligatorily with a certain number of verbs, whereas the latter go with most intransitive verbs optionally.

(4) **S/A-agreement prefixes:**
a. *go k-aʔi lau skola kia.*
1SG 1SG-go DIR.SEA school PROS
'I will go seawards to school.'
b. *Hugo n-enũ tua.*
Hugo 3SG-drink tuak
'Hugo drank tuak.'

(5) **S-agreement enclitics:**
Wato turu =aʔ kaeʔ.
Wato sleep =3SG PFV
'Wato already slept.'

A (lexical) possessor precedes its possessum, as in (6) and (7). Inalienable and alienable possessions take different constructions, as in (6) and (7), being marked by *-N* and *=kə̃* on

the possessum, respectively.[4] The suffix -*N* is realized as nasalization of the final vowel of a preceding word (Nagaya 2010).

(6) **Inalienable possession:**
Hugo *kotə̃ʔ*
Hugo kotəʔ -N
Hugo head -NMLZ
'Hugo's head'

(7) **Alienable possession:**
Hugo *laŋoʔ* *=kə̃*
Hugo house =NMLZ
'Hugo's house'

Nouns precede numerals, as in (8). The negator and other TAM markers occur in the clause-final position, as in (9).

(8) **Noun + numeral:**
ata *rua*
person two
'two persons'

(9) **Clause-final TAM markers:**
go isə kbako həlaʔ.
1SG suck tobacco NEG
'I don't smoke.'

3 Ditransitive alternation

A handful of verbs referring to the physical transfer of possession or the mental transfer of experience can appear in both the prepositional-recipient and the ditransitive constructions. In **prepositional-recipient constructions**, a recipient NP is introduced by the locative *ia*, while a theme NP appears in object position. See (10).

(10) **Prepositional-recipient construction [NP V NP PP]:**
ə̃mə̃ *neĩ* *doi* *ia* *go.*
mother give money LOC 1SG
'Mother gave money to me.'

In **ditransitive constructions**, both recipient and theme NPs occur in object position without a preposition-like element, resulting in a double object construction-like structure. See (11).

---

[4] The enclitic =*kə̃* and the suffix -*N*, both glossed as NMLZ, are used either as nominalizers or as possessive markers.

(11) **Ditransitive construction [NP V NP NP]:**
əmə̃ neĩ go doi.
mother give 1SG money
'Mother gave me money.'

The prepositional-recipient and ditransitive constructions are paraphrases of each other. They have the same truth-conditional meaning: *AGENT transfers (the ownership of) THEME to RECIPIENT*. For instance, both (10) and (11) mean that the mother transferred the ownership of money to the speaker.

Such an alternation without additional verbal morphology is familiar from English, which is known for the alternation between the double object construction (e.g., *Mother gave me money*) and the prepositional-dative construction (e.g., *Mother gave money to me*). But typologically speaking, this kind of alternation may not be common. For example, Dryer (2007:254) says "[m]any other languages employ constructions which are similar to one or the other of these two constructions [NN-double object and prepositional-dative constructions] in English, though it is less common to have both constructions, the way English does." See also Haspelmath (2004, 2005), Kittilä (2005:278ff), and Malchukov, Haspelmath & Comrie (2010:18).

Importantly, the ditransitive alternation in Lamaholot is not productive: only a few verbs can be involved in this alternation, and such verbs are divided into a small number of semantic classes: GIVE-verbs (*neĩ* 'give' and *sorõ* 'give'), TELL-verbs (*tutu* 'tell' and *marĩ* 'say'), a TEACH-verb (*adʒa* 'teach'[5]), and SHOW-verbs (*nonĩ* 'show' and *pro* 'show something to eat or drink'). These verbs express the physical transfer of possession, or the mental transfer of experience or knowledge. See ditransitive alternations with TELL-, TEACH-, and SHOW-verbs in (12), (13), and (14), respectively.

(12) **TELL-verb *tutu* 'tell (a story)':**
a. *Hugo tutu koda ia go.*
Hugo tell story LOC 1SG
'Hugo told a story to me.'
b. *Hugo tutu go koda.*
Hugo tell 1SG story
'Hugo told me a story.'

(13) **TEACH-verb *adʒa* 'teach':**
a. *Hugo adʒa koda kiwə̃ ia go.*
Hugo teach language forestLOC 1SG
'Hugo teaches the language of the forest (i.e., Lamaholot) to me.'
b. *Hugo adʒa go koda kiwə̃.*
Hugo teach 1SG language forest
'Hugo teaches me the language of the forest (i.e., Lamaholot).'

[5] The verb *adʒa* 'teach' is borrowed from the verb *ajar* 'teach' in Indonesian.

(14) **SHOW-verb *nonĩ* 'show':**

a. *Hugo nonĩ gambar ia go.*
Hugo show picture LOC 1SG
'Hugo showed a picture to me.'

b. *Hugo nonĩ go gambar.*
Hugo show 1SG picture
'Hugo showed me a picture.'

From a typological perspective, the fact that only these verbs can occur in the Lamaholot ditransitive alternation is not surprising. Malchukov, Haspelmath & Comrie (2010: 50ff) note that "when a language has a closed class of ditransitive verbs, the same lexemes tend to recur in this class in language after language, most frequently verbs like 'give', 'show', 'teach', sometimes also 'tell', 'send', and 'ask'."

As will be discussed in Section 5, there are clear syntactic differences between the two alternating constructions in question. In contrast, semantic differences, if any, between the two constructions of the alternation are not immediately obvious. Malchukov, Haspelmath & Comrie (2010:20ff) point out several factors, such as animacy, definiteness, and affectedness, that may condition the choice between the alternating constructions in a language with a verb alternation of the kind that Lamaholot has. However, none of these factors seem to determine the choice of a construction in the Lamaholot ditransitive alternation.[6] For instance, observe in (15) and (16) that the "person effect" is not observed in the ditransitive alternation in Lamaholot. An inanimate theme NP is always allowed in both constructions, which may not be the case in other languages such as English.

(15) **Prepositional-recipient construction + inanimate theme:**
go neĩ peʔẽ ia mo.
1SG give DEM.DIS.NMLZ LOC 2SG
'I will give that/it to you.'

(16) **Ditransitive construction + inanimate theme:**
go neĩ mo peʔẽ.
1SG give 2SG DEM.DIS.NMLZ
'I will give you that/it.' (Not acceptable in English; cf. Haspelmath 2007)

Before going on to discuss the benefactive alternation, it is important to note that Lamaholot is a so-called pro-drop language, and therefore, sentences like (17) are grammatical. In such cases, these sentences are ambiguous between the prepositional-recipient and the ditransitive clause structures.

(17) *ə̃mə̃ neĩ doi.*
mother give money
'Mother gave money (to you/me/etc.).'

---

[6] However, these factors might affect preferences for one construction over another, as argued by Bresnan et al. (2007) with regard to the English dative alternation.

4 Benefactive alternation

Lamaholot employs various kinds of serial verb constructions. This section is concerned with benefactive serial verb constructions using the verb of giving *neĩ* 'give'. In **benefactive serial verb constructions**, a verb of giving introduces a participant who receives something from an action and on behalf of whom the action is done (i.e., **recipient-beneficiary**; see Kittilä 2005), while another verb, either intransitive or transitive, expresses an action done for the recipient-beneficiary participant.[7,8] In (18) below, an intransitive verb is used in the serial construction, while in (19), a transitive verb is used.

(18) *Nia kriə̃ neĩ ba nəʔẽ.*
Nia work give father3SG.NMLZ
'Nia works for her father (i.e., in order to give some help to her father).'

(19) *Ika hope gula neĩ Wato.*
Ika buy candy give Wato
'Ika bought candy for Wato (i.e., in order to give candy to Wato).'

In this language, a number of transitive verbs that profile an action done on behalf of a recipient-beneficiary participant can appear in both the benefactive serial verb and the benefactive constructions. In benefactive serial verb constructions, as described above, a recipient-beneficiary NP is introduced by a verb of giving, while a patient NP occurs in object position. See (20).

(20) **Benefactive serial verb construction [NP V NP give NP]:**
go buka knaweʔ neĩ Besa.
1SG open door give Besa
'I opened a door for Besa.'

In **benefactive constructions**, by contrast, both beneficiary-recipient and patient NPs occur in object position without a preposition-like element, creating a double object construction-like structure. See (21).

(21) **Benefactive construction [NP V NP NP]:**
go buka Besa knaweʔ.
1SG open Besa door
'I opened a door for Besa.'
(lit. 'I opened Besa a door.')

Similarly to the ditransitive alternation, benefactive serial verb constructions of the type above and benefactive constructions are paraphrases of each other. They have the same

---

7 A serial verb construction with *neĩ* 'give' is only employed for introducing a recipient-beneficiary NP. Thus, sentences like **go tedũ pi neĩ mo* (intended for 'I waited here for you') are not acceptable, because in such a sentence, the participant for which an action is done receives nothing. This generalization seems contradicted by example (20) below, but this is not the case. In (20), Besa received *space* to go through, created by the speaker when they opened a door.

8 As one of the reviewers suggested, it is possible to regard the verb *neĩ* 'give' as a preposition, rather than the (second) verb in a serial verb construction, if it has undergone grammaticalization.

truth-conditional meaning: *AGENT acts on PATIENT for the sake of RECIPIENT-BENEFICIARY*. Thus, (20) and (21) share the meaning that the speaker opened the door so that Besa could go through it.

Unlike the ditransitive alternation, the benefactive alternation in Lamaholot is productive. A variety of transitive verbs can take part in this alternation, as long as they fit with the basic meaning of the benefactive construction above. To illustrate this, consider the examples in (22) through (29), for instance.

(22) ***gute* 'get':**

a. *mo gute kursi neĩ go.*
2SG get chair give 1SG
'You get a chair for me.'

b. *mo gute go kursi.*
2SG get 1SG chair
'You get me a chair.'

(23) ***lərə̃* 'load':**

a. *ra lərə̃ semen ia oto neĩ go.*
3PL load cement LOC car give 1SG
'They loaded cement on the car for me.'

b. *ra lərə̃ go semen ia oto.*
3PL load 1SG cement LOC car
'They loaded cement on the car for me.'
(lit. They loaded me cement on the car.')

(24) ***kriə̃* 'work, build':**

a. *Nius kriə̃ laŋoʔ neĩ ra.*
Nius work housegive 3PL
'Nius built a house for them.'

b. *Nius kriə̃ ra laŋoʔ.*
Nius work 3PL house
'Nius built them a house.'

(25) ***namo* 'sweep':**

a. *Lin namo laŋoʔ neĩ Eli.*
Lin sweep house give Eli
'Lin swept a house for Eli.'

b. *Lin namo Eli laŋoʔ.*
Lin sweep Eli house
'Lin swept a house for Eli.'
(lit. 'Lin swept Eli a house.')

(26) ***məla* 'steal':**

a. *mo* *məla* *manu* *neĩ* *go.*
2SG steal chicken give 1SG
'Steal a chicken for me!'

b. *mo* *məla* *go* *manu.*
2SG steal 1SG chicken
'Steal a chicken for me!'
(lit. 'Steal me a chicken!')

(27) ***pulə* 'choke (a chicken)':**

a. *Srinu* *pulə* *manu* *neĩ* *Rani.*
Srinu choke chicken give Rani
'Srinu choked a chicken for Rani.'

b. *Srinu* *pulə* *Rani* *manu.*
Srinu choke Rani chicken
'Srinu choked a chicken for Rani.'
(lit. 'Srinu choked Rani a chicken.')

(28) ***sepa* 'kick':**

a. *go* *sepa* *bal* *neĩ* *Siku.*
1SG kick ball give Siku
'I kicked the ball for Siku.'

b. *go* *sepa* *Siku* *bal.*
1SG kick Siku ball
'I kicked the ball for Siku.'
(lit. 'I kicked Siku the ball.')

(29) ***basa* 'read':**

a. *Dani* *basa* *buku* *neĩ* *anaʔ* *nəʔẽ.*
Dani read book give child 3SG.NMLZ
'Dani read a book for his child.'

b. *Dani* *basa* *anaʔ* *nəʔẽ* *buku.*
Dani read child 3SG.NMLZ book
'Dani read a book for his child.'
(lit. 'Dani read his child a book.')

As demonstrated by the examples above, a variety of verbs can occur in the benefactive alternation in Lamaholot. For example, sentences like *I opened Mary a door* may not sound totally grammatical in the English benefactive construction (Shibatani 1996) but are perfectly fine in Lamaholot, as in (21). Other verbs that can appear in this alternation include *ø-ahuʔ* 'get (water)', *bahaʔ* 'wash', *bao* 'pour', *behĩ* 'pour (tuak)', *dʒãʔũ* 'snatch', *ø-ãʔã* 'make', *ø-əte* 'hold', *gnato* 'send', *hope* 'buy', *huka* 'steam', *itiʔ* 'lift', *lətəʔ* 'close', *rekam* 'record', *rodu* 'push', *səba* 'search', *segu* 'mix', *siap* 'prepare', *taʔo* 'put', *takuʔ* 'feed (a baby)', *tulũ* 'help', etc.

Note that some Lamaholot benefactive constructions look like English double object constructions in terms of surface structure, but semantically, they are completely different. See (30) and (31), for example.

(30) ***nə̃nə̃* 'ask':**

a. *Hans nə̃nə̃ doi neĩ go.*
Hans ask money give 1SG
'Hans asked for money on my behalf.'

b. *Hans nə̃nə̃ go doi.*
Hans ask 1SG money
'Hans asked for money on my behalf.'
(lit. 'Hans asked me money'; does not mean 'Hans asked me for money.')

(31) ***paʔu* 'feed (an animal)':**

a. *Sius paʔu wawe neĩ go.*
Sius feed pig give 1SG
'Sius fed a pig for me.'

b. *Sius paʔu go wawe.*
Sius feed 1SG pig
'Sius fed a pig for me.'
(lit 'Sius fed me a pig'; does not mean 'Sius made me eat pork.')

The two sentences in (30) mean that Hans asked for money (from someone) for the sake of the speaker. The benefactive construction *Hans nə̃nə̃ go doi* has a similar structure to the English double object construction *Hans asked me money*, but they have different meanings. Also, note that the verb *paʔu* 'feed' in (31) is involved in the benefactive alternation, while the verb *feed* in English appears in the dative alternation.

Similar to the ditransitive alternation, a sharp syntactic difference can be observed between the two competing constructions of the benefactive alternation (see Section 5). However, semantic differences between the two constructions are not clear. Unlike benefactive constructions in other languages (cf. Shibatani 1996; Van Valin 2007), semantic factors such as animacy, definiteness, and affectedness do not substantially affect the choice of one construction over the other in Lamaholot. Compare, for example, (32) and (33) below. Neither construction necessarily means that a recipient-beneficiary participant receives a theme element. It appears that this alternation has nothing to do with the degree of affectedness of a recipient-beneficiary participant.

(32) **Benefactive serial verb construction + event cancelation:**
go hope buku neĩ Nia, kũ na ai həlaʔ.
1SG buy book give Nia but 3SG get NEG
'I bought a book for Nia, but she didn't get (it).'

(33) **Benefactive construction + event cancelation:**
go hope Nia buku, kũ na ai həlaʔ.
1SG buy Nia book but 3SG get NEG
'I bought Nia a book, but she didn't get (it).'

Before closing this section, there is another aspect of the benefactive alternation worthy of note. Interestingly, the main verb of a benefactive construction can optionally be followed by the verb *neĩ* 'give'. Compare (29) with (34) below. Note that (34) allows a benefactive reading only, and must mean that Dani read a book for his child; it cannot mean that Dani gave a book to his child.

(34) *Dani basa* ***neĩ*** *anakʔ nəʔẽ buku.*
Dani read **give** child 3SG.NMLZ book
'Dani read a book for his child.'

As far as my current database shows, there is no benefactive construction that cannot take the verb *neĩ* 'give' after the main verb, and no syntactic or semantic differences between constructions with the verb *neĩ* and those without are attested. This may be considered as an incipient stage of grammaticalization of the verb of giving to an applicative marker, but further research is necessary to examine this hypothesis.[9]

5 Similarities and differences between the two alternations

This section discusses syntactic similarities and differences between the ditransitive and benefactive alternations. In Section 5.1, it is claimed that the two alternations share the syntactic function of realigning the primary object grammatical relation from one argument to another. Section 5.2 illustrates how this function is correlated with topic-related syntactic phenomena. In Section 5.3, I show that the two alternations behave differently with regard to the syntactic status of non-primary object NPs.

5.1 Shared syntactic function of ditransitive and benefactive alternations

As mentioned in Sections 3 and 4, it does not appear that the ditransitive and the benefactive alternations create any noticeable difference in (truth-conditional) meaning. Foremost, native speakers of Lamaholot judge both to have the same meaning. Additionally, no linguistic test appears to support the existence of semantic differences.

However, in my analysis, these alternations share an important syntactic function of realigning the **primary object grammatical relation** (PO: a grouping of patient arguments in mono-transitive constructions and recipient arguments in ditransitive constructions; Dryer 1986) from one argument to another, with the alternating constructions differing with regard to which NP bears this grammatical relation. The summary of my analysis is given in (35) and (36).

(35) **Ditransitive alternation:**
a. Prepositional-recipient: A *theme NP* is in PO.
b. Ditransitive: A *recipient NP* is in PO.

(36) **Benefactive alternation:**
a. Benefactive SVC: A *patient NP* is in PO.
b. Benefactive construction: A *recipient-beneficiary NP* is in PO.

---

[9] According to Matt Shibatani (pers. comm.), this type of benefactive construction is common in Sika, a language spoken to the west of the region where the Lewotobi dialect of Lamaholot is spoken. See (a).

(a) Petrus boter beli inan payung.
Petrus buy give mother umbrella
'Petrus bought her mother an umbrella.

The above-mentioned analysis is supported, among others, by coreferentiality with the pronominal enclitic *=roʔ* '3SG'. First, observe that patient NPs in transitive constructions can be coreferential with the pronominal enclitic *=roʔ* '3SG'.

(37) **Transitive construction:**
go k-oi Nia.
1SG 1SG-know Nia
'I saw Nia.'

(38) **Transitive construction + =roʔ:**
go k-oi =roʔ na. [*=roʔ*→ patient NP]
1SG 1SG-know =3SG 3SG
'I saw him/her.'

Second, observe below that what is coreferential with *=roʔ* is a theme NP in a prepositional-recipient construction ((39) vs. (40)) but a recipient NP in a ditransitive construction ((41) vs. (42)).

(39) **Prepositional-recipient construction:**
go neĩ gula ia Nia.
1SG give candy LOC Nia
'I gave candy to Nia.'

(40) **Prepositional-recipient construction + =roʔ:**
go neĩ =roʔ ia na. [*=roʔ*→ theme NP]
1SG give =3SG LOC 3SG
'I gave it to him/her.'
cf. **go neĩ roʔ gula ia na.*[10]

(41) **Ditransitive construction:**
go neĩ Nia gula.
1SG give Nia candy
'I gave Nia candy.'

(42) **Ditransitive construction + =roʔ:**
go neĩ =roʔ na gula. [*=roʔ*→ recipient NP]
1SG give =3SG 3SG candy
'I gave him/her candy.'
cf. **go neĩ roʔ Nia.*

Likewise, what is coreferential with *=roʔ* is a patient NP in a benefactive serial verb construction, but a recipient-beneficiary NP in a benefactive construction. See examples (43) through (46).

[10] The enclitic pronoun *=roʔ* cannot be used when its referent is expressed by a (non-topicalized) lexical noun.

(43) **Benefactive serial verb construction:**
go hope gula neĩ Ika.
1SG buy candy give Ika
'I bought candy for Ika.'

(44) **Benefactive serial verb construction + =*ro?*:**
go hope =ro? neĩ Ika. [=*ro?*→ patient NP]
1SG buy =3SG give Ika
'I bought it for Ika.'

(45) **Benefactive construction:**
go hope Ika gula.
1SG buy Ika candy
'I bought Ika candy.'

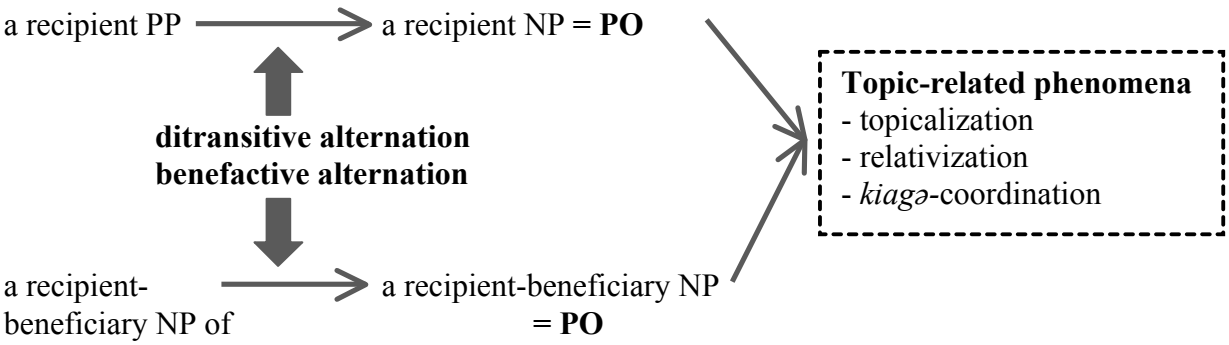
(46) **Benefactive construction + =*ro?*:**
go hope =ro? na gula. [=*ro?*→ recipient-beneficiary NP]
1SG buy =3SG 3SG candy
'I bought him/her candy.'
cf. **go hope ro? Ika.*

In summary, the following arguments behave alike with regard to =*ro?*: (i) a patient argument of transitive constructions, including prepositional-recipient and benefactive serial verb constructions; (ii) a recipient argument in ditransitive constructions; and (iii) a recipient-beneficiary argument in benefactive constructions. In Dryer's (1986:814) definition, a nominal counts as a primary object if it is a recipient-like argument in a ditransitive clause or a patient-like argument in a mono-transitive clause. On the basis of this observation, therefore, we can postulate the primary object relation in Lamaholot.

The data in (37) through (46) also show that different constructions of the ditransitive and benefactive alternations have different arguments for the primary object. This means that the syntactic function of the ditransitive and benefactive alternations is that of realigning the primary object grammatical relation from one argument to another.

5.2 Primary objecthood and topic-related phenomena

The fact that the ditransitive and benefactive alternations change the alignment of the primary object grammatical relation is quite significant for Lamaholot syntax. This is because arguments bearing the primary object relation can be topicalized and thus can serve as a controller or pivot in topic-related syntactic phenomena, such as relativization and *kia gə*-coordination (Nagaya 2013). See Figure 5.1.



a serialized verb

Figure 5.1: The syntactic function of the ditransitive and the benefactive alternations

To begin with, let us confirm that primary object NPs can be involved in topic-related phenomena, but not adjuncts (Nagaya 2013). As in (47) and (48), the primary object pronoun *mo* '2SG' can be topicalized and relativized. In addition, examples in (49) show that the topicalized primary object NP can serve as a controller for the *kia gə*-coordination construction.

(47) **Topicalization of a primary object NP:**

go	*bəŋo*	*mo.*	→	*mo*	*go*	*bəŋo* __.
1SG hit	2SG			2SG	1SG	hit
'I hit you.'				'You, I hit.'		

(48) **Relativization of a primary object NP:**

teʔẽ	*anaʔ*	*yang*	*go*	*bəŋo*	__.
DEM.PROX.NMLZ	person	REL	1SG	hit	

'This is the person I hit.'

(49) ***Kia gə*-coordination:**

a.

go	*bəŋo*	*mo,*	*kia*	*gə*	*plaʔe.*	[controller = topic SUBJ]
1SG	hit	2SG	PROS	CONJ	run	

'I hit you, and (I) ran.'

b.

mo	*go*	*bəŋo,*	*kia*	*gə*	*plaʔe.*	[controller = topicalized PO]
2SG	1SG	hit	PROS	CONJ	run	

'You, I hit, and (you) ran.'

In contrast, (50) and (51) illustrate that the adjunct NP *laŋoʔ* 'house' cannot be topicalized or relativized. The NP *laŋoʔ* 'house' is marked by the locative *ia* and is an adjunct NP rather than a primary object NP.

(50) **Topicalization of an adjunct:**

go	*tei*	*ia*	*laŋoʔ.*	→	**laŋoʔ*	*go*	*tei*	*ia* __.
1SG live		LOC	house		house	1SG	live	LOC
'I live in the house.'					Intended for 'As for the house, I live in (it).'			

(51) **Relativization of an adjunct:**

**teʔẽ*	*laŋoʔ*	*yang*	*go*	*tei*	*ia*	__.
DEM.PROX.NMLZ	house	REL	1SG	live	LOC	

Intended for 'This is the house in which I live.'

Then, observe in (52) and (53) that a theme NP of a prepositional-recipient construction and a patient NP of a benefactive serial verb construction can be directly topicalized, but a recipient PP of a prepositional-recipient construction and a recipient-beneficiary NP of a benefactive serial verb construction cannot.

(52) **Prepositional-recipient construction:**

a. *go neĩ doi ia Ika.*
1SG give money LOC Ika
'I gave money to Ika.'

b. *doi go neĩ __ ia Ika.* [Topicalization of a theme]
'As for the money, I gave (it) to Ika.'

c. **Ika go neĩ doi ia __.* [*Topicalization of a recipient]
Intended for 'As for Ika, I gave money to (her).'

d. *doi yang go neĩ __ ia Ika* [Relativization of a theme]
'the money I gave to Ika.'

e. **anaʔ yang go neĩ doi ia __* [*Relativization of a recipient]
Intended for 'the person to whom I gave money.'

(53) **Benefactive serial verb construction:**

a. *go biho lama neĩ Nia.*
1SG cook rice give Nia
'I cooked rice for Nia.'

b. *lama go biho __ neĩ Nia.* [Topicalization of a patient]
'As for the rice, I cooked (it) for Nia.'

c. **Nia go biho lama neĩ __.* [*Topicalization of a recipient-beneficiary]
Intended for 'As for Nia, I cooked rice for (her).'

d. *lama yang go biho __ neĩ Nia.* [Relativization of a patient]
'the rice I cooked for Nia.'

e. **anaʔ yang go biho lama neĩ __* [*Relativization of a recipient-beneficiary]
Intended for 'the person for whom I cooked rice.'

Now, observe in (54), (55), and (56) that a recipient NP in a ditransitive construction can be topicalized and be a pivot of relativization and a controller for the *kia gə*-coordination construction. Importantly, a recipient NP can display these syntactic behaviors only when it occurs in a ditransitive construction.

(54) **Ditransitive construction + topicalization of a recipient NP:**
mo go neĩ __ gula.
2SG 1SG give candy
'As for you, I gave (you) candy.'

(55) **Ditransitive construction + relativizationof a recipient NP:**
teʔẽ anaʔ yang go neĩ __ gula.
DEM.PROX.NMLZ child REL 1SG give candy
'This is the child I gave candy.'

(56) **Ditransitive construction + *kia gə*-coordination construction:**
mo go neĩ __ gula kia gə plaʔe.
2SG 1SG give candy PROS CONJ run
'As for you, I will give (you) candy and then (you should) run.'
cf. *go neĩ mo gula kia gə plaʔe.*
'I will give you candy and then (I will) run.'

This is also the case with a recipient-beneficiary NP in a benefactive construction. The recipient-beneficiary NP *mo* '2SG' is topicalized in (57), is relativized in (58), and is a

controller for the *kia gə*-coordination in (59). A recipient-beneficiary NP can display these syntactic behaviors only when it appears in a benefactive construction.

(57) **Benefactive construction + topicalization of a recipient-beneficiary NP:**
mo go hope __ gula.
2SG 1SG buy candy
'As for you, I bought (you) candy.'

(58) **Benefactive construction + relativizationof a recipient-beneficiary NP:**
teʔẽ anaʔ yang go hope __ gula.
DEM.PROX.NMLZ child REL 1SG buy candy
'This is the child for whom I bought candy.'

(59) **Benefactive construction + *kia gə*-coordination construction:**
mo go hope __ gula kia gə plaʔe.
2SG 1SG buy candy PROS CONJ run
'As for you, I will buy (you) candy and then (you should) run.'
cf. *go hope mo gula kia gə plaʔe.*
'I will buy you candy and then (I will) run.'

To conclude, two constructions with a double object construction-like structure, namely, ditransitive and benefactive constructions, have in common the function of "promoting" a non-patient/theme (or recipient-like) argument into primary object. In this sense, they are functionally equivalent to applicative constructions (but with no overt marking on verbs).

5.3 Symmetric applicative vs. asymmetric applicative

Despite the facts discussed above, the ditransitive and benefactive double object construction-like structures differ with regard to the grammatical nature of a patient/theme NP (i.e., a non-applied object NP or a non-primary object NP).[11] On the one hand, a theme NP in a ditransitive construction still counts as a grammatical argument and can be topicalized and therefore show topic-related features. (60), for instance, indicates that a theme NP can still be topicalized in a ditransitive construction. Furthermore, (61) illustrates that it can be directly relativized.

(60) **Ditransitive construction + topicalization of a theme NP:**
gula go neĩ mo __.
candy 1SG give 2SG
'As for the candy, I gave you (it).'

[11] 'Applied objects' refer to those objects that are added in applicative constructions, while 'non-applied objects' refer to those that are not. In the ditransitive construction, a recipient NP is an applied object NP bearing the primary object grammatical relation, while a theme NP is a non-applied object NP that is not in the primary object grammatical relation. Likewise, in the benefactive construction, a recipient-beneficiary NP is an applied object NP, whereas a patient NP is a non-applied object NP.

(61) **Ditransitive construction + relativization of a theme NP:**
teʔẽ *gula* *yang go* *neĩ* *mo* __.
DEM.PROX.NMLZ candy REL 1SG give 2SG
'This is the candy that I gave you.'

By contrast, a patient NP in a benefactive construction does not work as a grammatical argument. It cannot be topicalized or display topic-related properties. (62), for example, demonstrates that a patient NP cannot be topicalized in a benefactive construction. Likewise, (63) shows that a patient NP cannot be relativized.

(62) **Benefactive construction + topicalization of a patient NP:**
**gula* *go* *hope mo* __.
candy 1SG buy 2SG
Intended for 'As for the candy, I bought you (it).'

(63) **Benefactive construction + relativization of a patient NP:**
**teʔẽ* *gula* *yang go* *hope mo* __.
DEM.PROX.NMLZ candy REL 1SG buy 2SG
Intended for 'This is the candy I bought you.'

To summarize, both recipient and theme NPs work as grammatical arguments in ditransitive constructions, while only recipient-beneficiary NPs serve as such in benefactive constructions. In other words, the ditransitive construction is functionally equivalent to a **symmetric applicative**, and the benefactive construction, an **asymmetric applicative**. In the former, both applied and non-applied NPs work as grammatical arguments; in the latter, only applied NPs do so. Although both have a double object construction-like structure, ditransitive and benefactive constructions display different syntactic behaviors.

6 Conclusions

This paper has provided a description and analysis of the ditransitive and benefactive alternations in Lamaholot. The findings of this paper are summarized as in Table 6.1.

Table 6.1: Summary

	DITRANSITIVE ALTERNATION	BENEFACTIVE ALTERNATION
PRODUCTIVITY	Non-productive	Productive
TARGET	Recipient	Recipient-beneficiary
SYMMETRICITY	Symmetric applicative	Asymmetric applicative

The ditransitive alternation is a non-productive verb alternation where one and the same verb of giving, telling, teaching, or showing appears in prepositional-recipient and ditransitive constructions. By contrast, the benefactive alternation is a productive verb alternation in which a transitive verb profiling an action done for the sake of another participant is used in benefactive serial verb and benefactive constructions. Although the semantic differences between the two alternating constructions of each type of verb alternation seem unclear, there is an obvious syntactic difference between them. A theme NP bears the primary object grammatical relation in prepositional-recipient constructions,

while a recipient NP does so in ditransitive constructions; a patient NP is in primary object grammatical relation in benefactive serial verb constructions, whereas a recipient-beneficiary NP attains such status in benefactive constructions. In other words, these two alternations have the shared syntactic function of realigning the primary object grammatical relation from one argument to another.

This study also showed that, although ditransitive and demonstrative constructions are functional equivalents of applicative constructions, they differ according to the syntactic status of non-applied objects. A non-applied theme NP still counts as a grammatical argument in ditransitive constructions, but a non-applied patient NP does not play such a role in benefactive constructions.

Of course, this study does not solve all problems concerning ditransitive and benefactive verb alternations in Lamaholot. In particular, the question of why Lamaholot has developed two distinct kinds of double object construction phenomena still remains unanswered. Most probably, this question may be explained from a historical perspective (i.e., in the spirit of Kikusawa 2012). The contrast we see between the ditransitive and benefactive alternations in modern Lamaholot might be a remnant of some important semantic or syntactic distinctions that once existed (for example, locative focus vs. benefactive focus constructions in the Philippine languages). This question would be a fruitful topic for future studies.

References

Arka, I Wayan. 2007. Creole genesis and extreme analyticity in Flores languages. Presented at the 5th International East Nusantara Conference on Language and Culture (ENUS), Kupang, Indonesia, August 2007.

Bowden, John. 2008. A grammar of Lamaholot, Eastern Indonesia: The morphology and syntax of the Lewoingu dialect (review). *Oceanic Linguistics* 47.1:247–251.

Bresnan, Joan, Anna Cueni, Tatiana Nikitina & Harald Baayen. 2007. Predicting the Dative Alternation. In G. Bouma, I. Kraemer & J. Zwarts (eds.) *Cognitive Foundations of Interpretation*, 69–94. Amsterdam: Royal Netherlands Academy of Arts and Sciences.

Donohue, Mark. 2007. Word order in Austronesian from north to south and west to east. *Linguistic Typology* 11: 349–391.

Dryer, Matthew S. 1986. Primary objects, secondary objects, and antidative. *Language* 62.4:808–845.

Dryer, Matthew S. 2007. Clause types. In Timothy Shopen (ed.), *Language Typology and Syntactic Description*, Volume 1: *Clause Structure*, Second Edition, 224–275. Cambridge: Cambridge University Press.

Grimes, Charles E., Tom Therik, Barbara D. Grimes & Max Jacob. 1997. *A Guide to the People and Languages of Nusa Tenggara*. Center for Regional Studies, Paradigma, Series B, No 1. Kupang: Artha Wacana Press.

Haspelmath, Martin. 2004. Explaining the Ditransitive Person-Role Constraint: a usage-based account. *Constructions* 2/2004, 49pp.

Haspelmath, Martin. 2005. Argument marking in ditransitive alignment types. *Linguistic Discovery* 3.1:1–21.

Haspelmath, Martin. 2007. Ditransitive alignment splits and inverse alignment. *Functions of Language* 14.1:79–102.

Himmelmann, Nikolaus P. 2005. The Austronesian languages of Asia and Madagascar: Typological Characteristics. In Alexander Adelaar & Nikolaus P. Himmelmann (eds.), *The Austronesian Languages of Asia and Madagascar*, 110–181. London: Routledge.

Keraf, Gregorius. 1978. *Morfologi Dialek Lamalera Disertasi*. Ende: Arnoldus.

Kikusawa, Ritsuko. 2012. Indoneshiago-no *-i* applicative-no hattatsushi-ni kansuru shiron [An attempt at the history of the development of *-i* applicative in Indonesian]. A talk given at the ILCAA of Tokyo University of Foreign Studies, May 19, 2012.

Kittilä, Seppo. 2005. Recipient-prominence vs. beneficiary-prominence. *Linguistic Typology* 9.2:269–297.

Klamer, Marian. 2002. Typical features of Austronesian languages in Central/Eastern Indonesia. *Oceanic Linguistics* 41.2:363–384.

Lewis, M. Paul, Gary F. Simons & Charles D. Fennig (eds.), 2013. *Ethnologue: Languages of the World*, Seventeenth edition. Dallas, Texas: SIL International. Online version: http://www.ethnologue.com.

Malchukov, Andrej, Martin Haspelmath & Bernard Comrie. 2010. Ditransitive constructions: a typological overview. In Andrej Malchukov, Martin Haspelmath & Bernard Comrie (eds.), *Studies in Ditransitive Constructions: A Comparative Handbook*, 1–64. Berlin: De Gruyter Mouton.

Margetts, Anna & Peter K. Austin. 2007. Three-participant events in the languages of the world: towards a crosslinguistic typology. *Linguistics* 45.3:393–451.

Musgrave, Simon. 2008. Typology and geography in eastern Indonesia. In Timothy Jowan Curnow (ed.), *Selected Papers from the 2007 Conference of the Australian Linguistic Society*. http://www.als.asn.au.

Nagaya, Naonori. 2010. A phonological sketch of Lewotobi Lamaholot. *Asian and African Languages and Linguistics* 5:157–195.

Nagaya, Naonori. 2013. Voice and grammatical relations in Lamaholot of eastern Indonesia. In Alexander Adelaar (ed.), *More on Voice in Languages of Indonesia*, 67–101. *NUSA* 54.

Shibatani, Masayoshi. 1996. Applicatives and benefactives: A cognitive account. In Masayoshi Shibatani & Sandra A. Thompson (eds.), *Grammatical Constructions: Their Form and Meaning*, 157–194. Oxford: Oxford University Press.

Van Valin, Robert D., Jr. 2007. The Role and Reference Grammar analysis of three-place predicates. *Suvremena Lingvistika* 63:31–63.

13 *Tao[1] Voice Affixes: Derivation or Inflection or Both?*

STACY WAN TIN HUANG

1 Introduction

The aim of this paper is to investigate the properties of the verbal voice affixes *ma-, mi-, man-, -om-, -en, -an,* and *i-* in Tao, an Austronesian language spoken on the Orchid Island located near the southeastern coast of Taiwan, and to justify their functional category. Starosta (2002) has argued, on the basis of evidence from nominalisation, that the Austronesian ***focus*/*voice*** affixes are lexical derivations. However, his proposal seems to contradict the present treatment of the verbal voice affixes in Austronesian languages. Verbal affixes in Tao are found to both trigger subject-verb agreement[2] and participate in nominalization processes at the same time. In signaling the agreement between the verbal affix and the theta role of the grammatical subject, the Tao verbal affix is similar to English voice morphology. However, the possibility that all occurrences of this affix can be treated as derivational is cast into doubt by several of its properties, including its participation in both nominalization processes and verbal usage, the systematic gaps and idiosyncratic behavior found in the verbal usages, and the degree of productivity and predictability of these affixes, whether as verbal affixes or nominal affixes. Starosta, Pawley, and Reid (1982) have proposed that the original function of Austronesian verbal focus affixes is as a nominalizer, and the verbal usage is 'a secondary development' (p.148). Ross (2009), in reconstructing Pre-Proto Austronesian verbal morphology, has also suggested that the present day verbal affixes are the results of the reanalysis of nominalizations. Support for this proposal can be found in Tao. The part of speech of a lexical item in Tao has to be determined by the syntactic position the item occupies, as in (1). The word *kanen* in (1a) is a noun meaning 'food' as it occupies the slot after the case marker *so*, but it is a verb meaning 'eat' in (1b) as it takes the sentence-initial position.

(1) a.	*sino*	*o*	*ya*	*ni-k-om-an*	*so*	***kanen***	*mo?*
	who	NOM	AUX	PERF<AV>eat	SO	food	2.S.GEN

'Who ate your food?'

(Rau and Dong: 98)

[1] Both Tao and Yami refer to the language spoken on the Orchid Island in Taiwan.

[2] Unlike English subject-verb agreement that agrees in person, number, and plurality. The subject-verb agreement in Tao refers to the verbal voice affix that agrees with the theta role of the nominative case-marked noun phrase.

Asia-Pacific Linguistics, 2014.

b.	*kan-en*	*na*	*ni*	*Salang*	*o*	*wakay.*
	eat-PV	3.S.GEN	GEN	PN	NOM	sweet_potato

'Salang ate the sweet potato. (lit.) What Salang ate was the sweet potato'
(Rau and Dong: 87)[3]

At first glance, the affixation of these affixes seems to change the word class of the stem from either V to N or N to V. However, based on the aforementioned studies on reconstructing Austronesian voice affixes, and evidence found in Tao, it seems reasonable to suggest that the verbal affixes are the results of reanalysis, and that their newly-developed function is to show agreement with the grammatical subject. The present study suggests that these affixes should be listed as derivational affixes when involved in nominalization, and inflectional affixes when they participate in verbal uses to serve the grammatical function of subject-shifting. This can better account for the language facts, and also explains the cross-category distribution of these affixes. The systematic gaps and idiosyncratic behavior found in the verbal usages, their productivity, and their predictability are discussed and explained.

The organization of the paper is as follows. In the next section, previous literature on Austronesian voice affixes is discussed followed by a brief sketch on Tao voice affixes and their distribution within various voice constructions. Section three discusses various criteria that are commonly used by researchers for distinguishing inflectional and derivational morphology. Section four examines the distribution of the affixes *ma-, -om-, -en, -an, and i-* in Tao to determine their functional category. Section five concludes the paper.

2 Literature Review and A Brief Syntactic Sketch of Tao

2.1 Literature Review

Earlier studies on Austronesian languages have observed rich verbal morphology, and focused on its relation with the subject NP. The term *focus* is used to refer to the agreement relation between the verbal affixes and the theta role of the subject NP. Later on, researchers (Chang, 1997; Ross and Tsang, 2005, and others) found that the term *focus* is misleading and suggested that *voice* should be adopted to describe the agreement between the verbal affix and the theta role of subject NP. Voice seems to be a better way to describe this agreement relationship but the functional category of these verbal affixes remains controversial.

> *According to O'Grady, "[v]oice is the term used to describe the grammatical system concerned with the correspondences or 'mapping' between thematic roles and grammatical relations, especially the subject relation. The most widely used voice category is the active, in which the agent role (if present) must linked to the subject." (O'Grady 2004: 110)*

[3] Interlinear glosses in the examples are modified for the present study. NOM – nominative; GEN – genitive; LOC – locative; OBL – oblique; AUX– auxiliary; PERF – perfective; AV – agent voice; PV – patient voice; LV – Locative voice; IV – Instrumental voice; S – Singular; P – plural; PN – personal name; TNS – tense; PRO – Pronoun; CM – Case marker; 1 – 1st person; 2 – 2nd person; 3 – 3rd person.

Notice that the major difference between derivational and inflectional morphology is that derivational morphology results in creation of a new word with new meaning, while the inflectional morphology involves a modification of a word to indicate its grammatical subclass, for instance, marking the syntactic function of a nominal in a phrase or a clause. If these verbal affixes are voice affixes, and their presence is mainly for the grammatical purpose of signaling syntactic relations, then they should be inflectional affixes. If not, the relationship between the verbal affix and the theta role of the subject NP should not be treated as voice.

A lexical item that can either take clause-initial predicate position or noun position after the determiner/case marker without any change in form is commonly observed across Austronesian languages, including Tao. Hence, researchers have suggested the precategorial hypothesis for the lack of verb-noun lexical distinction among root words and have proposed that lexical category can be determined by the syntactic position in the clause (Foley 1998). When the lexical category is underspecified, the functional category of voice affix remains mysterious.

Starosta et al. (1982), in reconstructing Proto-Austronesian morphology and syntax, claimed that the present-day verbal forms originated from nominalization, while the voice affixes originated as nominal affixes. Ross (2009), in reappraising Proto Austronesian verbal morphology, also suggested that there is a stage in which the reanalysis of nominalization has taken place, and that the present-day verbs, including the verbal affixes, are the result of reanalysis. He referred to these voice affixes as 'second-generation verbal affixes' (Ross 2009: 303). Furthermore, Ross also suggested that the jumble of voice affixes – prefix, infix, and suffix – might be the results of disparate origins of the nominalizers (Ross 2009: 301). Reid (1992), in analyzing the development of the aspect system in some Philippine languages, also suggested that the derivational characteristics of focus/voice are a result of their historical origin. All of these studies seem to indicate that the nouns have reanalyzed as verbs and their present day verbal usage is a newly-developed function that resulted from reanalysis in Proto Austronesian.

Starosta (2002), who investigated the evolution of voice in Austronesian, claimed that Austronesian voice affixes are a form of lexical derivation, based on the assumption that Austronesian languages are ergative, and that the presence of these voice affixes is for the purpose of coding intransitive and transitive clauses. The agent voice (AV) affixes code intransitive clauses, and non-AV voice affixes including patient voice (PV), location voice (LV), and instrumental voice (IV) code transitive clauses. He suggested that the existence of more than one transitive voice affixes could be explained using a common term, *applicative*, that is widely used in the study of Bantu languages.

Himmelmann (2008) questioned the validity of the precategorial proposal that lexical categories such as nouns and verbs in Tagalog are underspecified until their position in syntactic slots is identified. He argued that Tagalog roots could be classified into different morpho-lexical classes based on the observation of different uses and meanings of Tagalog action roots, and suggested that all kinds of root in Tagalog could be used without further affixation. He suggested that the voice affixes in Tagalog are derivational, based on the observation that voice affixes change the meaning and category of the roots to which they apply after affixation.

De Guzman (1997), who regarded the difference between inflectional and derivational morphology as scalar, suggested that Tagalog voice affixes exhibit more inflectional features than derivational features, and hence these voice affixes should be considered as inflectional. Researchers (De Guzman 1997, Himmelmann 2008) working on the same language – Tagalog – have opposite views regarding the functional category of the voice affixes. This may be an indication that the voice affixes in Tagalog exhibit some

inflectional and derivational features at the same time, which makes the classification of their functional category difficult.

After three decades, the proposal of Starosta et al. (1982) on the evolution of the focus/voice affixes in Austronesian languages remains a plausible supposition, and it has been accepted by researchers of this subfield (Reid, 1992; Rau, 2002; Himmelmann, 2008; Kaufman, 2008; Ross, 2009; among others) that the verbal usages are result of reanalysis of nominalizations. The voice affixes in Tao are found to participate in both noun and verb formation. Rau (2002), who worked on nominalization in Tao, claimed that the three stages of historical reconstruction proposed by Starosta et al. (1982) are reflected in Tao data. She found that indicative verbs and nominalization are indistinct in Tao so long as the pronoun is in genitive case. She suggested that these two constructions could be distinguished via movement of the genitive enclitics – "[the] movement of the genitive pronouns [in Tao] corresponds to the derivation of indicative verbs from nominalization (Rau 2002: 186). In her observation, in the first stage the genitive pronoun takes the position after the nominalized verb, as in (2a); in the second stage it takes the position before the nominalized verb, as in (2b); and in the third stage it moves to the sentence-initial position, as in (2c). Once the genitive pronoun is moved to sentence-initial position, the clause should be indicative rather than nominal. Rau (2002) found some support for the evolution of focus as proposed by Starosta et al. (1982) while working on nominalization in Tao. The functional category of the voice affixes remains unexplored in her study. However, her study on Tao nominalization, and observations on the different stages of the re-analysis processes that are reflected in the present-day language, provides a solid foundation for the present study.

(2)

a. *aon-en* *mo* *o* *kanen* *do* *vanga,* *ka-ap* ***mo***
take_out-PV 2.S.GEN NOM food LOC pot ka-fetch 2.S.GEN
so *asoy,* *ta* *nyo.*
OBL soup because 2.P.GEN
'Take the taros or potatoes out of the pot, then fetch the soup for you to drink.'
(Rau 2002: 187)

b. *ya* *aro* *[o* ***na*** *ni-akan* *no* *kanakan].*
TNS many NOM 3.S.GEN PERF.PV-eat GEN child
'(What) the child ate was a lot.'
(Rau 2002: 187)

c. ***da*** *i-ka-gogolang* *no* *koyis* *nyo* *ya?*
3.P.GEN IV-reason-thin GEN pig 2.P.GEN this
'Why are your pigs so skinny?'
(Rau 2002: 188)

2.2 A Brief Syntactic Sketch of Tao

Tao is a verb-initial language, and like other Austronesian languages, it exhibits a unique voice system, in that the theta role of the subject NP is encoded on the verb, as in (3). Notice that the PV verbal affix *-en* and the perfective aspect affix *ni-* are in complementary distribution. The voice affixes are listed in Table 2.1. It is found that stative-like predicates are sometimes not affixed with voice morphology, as in (3e). The absence of voice morphology is possible in Tao in agent and patient voice constructions, as

shown in (3e-f). The claim that the PV construction in (3f) is zero-marked is controversial. The absence of the PV affix in (3f) is due to the fact that the perfective marker *ni-* is a portmanteau morph that serves the functions of marking both patient voice and the perfective aspect.

(3)

a.	*[k-om-an*	*so*	*wakay]*	*[si*	*Salang]*
	<AV>eat	SO	sweet.potato	NOM	PN

'Salang wants to eat a sweet potato. (lit.) The one who wants to eat a sweet potato is Salang.'

b.	*[kan-en*	*na*	*ni*	*Salang]*	*[o*	*wakay]*
	eat-PV	3.S.GEN	GEN	PN	NOM	sweet.potato

'Salang ate the sweet potato. (lit.) What Salang ate was the sweet potato.'

c.	*[ni-akan-an*	*na*	*[o*	*mogis*	*ori]*	*ni*	*Salang]*
	PERF-eat-LV	3.S.GEN	NOM	rice	that	GEN	PN

'Salang ate some rice from there. (lit.) What Salang ate a little bit from was rice.'

d.	*[i-akan*	*na*	*ni*	*Salang]*	*[o*	*among*	*ya]*
	IV-eat	3.S.GEN	GEN	PN	NOM	fish	this

'Salang took this fish and ate it. (lit.) What was given for Salang to eat was fish.'
(Rau and Dong 2006: 87)

e.	*ya*	*Ø-apiya*	*o*	*Wakay*
	TNS	AV-good	NOM	Yam

'The yam is good.' (Shih 1997: 34)

f.	*ya-na*	*ni-kan-Ø*	*ni*	*mapapo*	*o*	*soli*
	AUX-3.S.GEN	PV.PERF-eat	GEN	PN	NOM	taro

'Mapapu has eaten up the taros.' (Ho 1990: 74)

Table 2.1: Verbal Affixes in Tao (Rau and Dong 2005: 88)

	AV	**PV**	**LV**	**IV/BF**
Affix	-om-/m-/ma-/Ø	-en/-ni	-an	i-

Besides clause-initial predicates, bound pronouns, linkers, and particles, other elements in the sentence are always preceded by a function word: a case marker. The case markers in Tao are listed in Table 2.2. The subject is nominative case-marked; the agent is genitive case-marked when it is not the subject of the sentence; instrument is also genitive case-marked; location is locative case marked; elements that are not subject, instrument, location, doer of the action are *so* case marked[4].

[4] The case marker *so* is glossed as an oblique case marker under the ergative analysis found in some of the literature on Tao (Shih 1997; Rau 2002, 2006; Deng 2005), but it is not treated as an oblique case marker under the symmetrical voice system. In a symmetrical voice system, there is no distinction between basic and derived voice categories, and demotion of the core argument is not proposed. In an asymmetrical voice system, such as that of English, the basic voice is active, and the derived voice is passive. In a derived voice such as passive, one of the core arguments is demoted. Hence, the case marker *so* that marks one of the two noun phrases in AV construction is glossed as oblique in an asymmetrical analysis.

Table 2.2: Case distinctions in Tao (Rau and Dong 2006: 118)[5]

	Nominative	**Genitive**	**Locative**	**So**
Common nouns	o	no	do	so
Singular personal names and kinship terms	si	ni	ji	--
Plural personal names and kinship terms	sira	nira	jira	--

When a case marker is observed in front of an element, this element plays a crucial role in the sentence – either as an argument or an adjunct. In other words, it occupies an NP position in the sentence. Except in the predicate-initial position, it is impossible to find a lexical item/content word that stands alone in the sentence in this language, as exemplified in (4). The sentence in (4) is a topicalized construction; the topicalized NP is fronted to the sentence-initial position. Notice that words excluding the adverb *rana* demonstrative *ya*, particle *am*, bound pronoun *na*, and the main predicate *ni-pi-vat-vatek* are all preceded by a function word, namely a case marker. The structure of (4a) is briefly analyzed and illustrated in (4b). The particle and adverb are omitted for the sake of simplicity. Two important observations can be made from this example. First, it shows the phrase structure of this language is [$_{XP}$ [X] [Y]], where X is the head, i.e.. the case marker, and Y is the lexical item, namely the position for the content words.

(4)

a.

o	*i-pi-vat-vatek*	*rana*	*ya*	*am,*
NOM	IV-VF-RED-carve	already	this	TOP
ni-pi-vat-vatek	*na*	*no*	*mehakay*	*so*
PERF-IV-RED-carve	3.S.GEN	GEN	man	SO
ngaran	*no*	*mavakes*	*do*	*vakong*
name	GEN	woman	LOC	paper

'As for this pen, it was used by the man to write a woman's name on the paper.'
(Rau and Dong 2006: 118)

b. [I [TP [$_{NOM}$ o][$_{CW}$i-pi-vat-vatek][INFL [$_{CW}$ ni-pi-vat-vatek][$_{GEN}$ na][GP [$_{GEN}$ no][$_{CW}$ mehakay] [OP[$_{OBL}$ so][$_{CW}$ ngaran] [GP[$_{GEN}$ no][$_{CW}$ mavakes] [LP[$_{LOC}$ do][$_{CW}$vakong]]]]]]]

Second, the Y position can be occupied by any content word even it is affixed with voice verbal morphology. The phrase structure in (4b) illustrates that every noun phrase in Tao must be headed by a case marker and anything following the case marker is part of the NP, even it is affixed with an apparent voice verbal affix in this language. Notice that the topicalized NP is affixed with instrumental affix *i-* and is almost the same as the main predicate, except that the main predicate is affixed with the perfective aspect marker *ni-*. This example not only shows the necessity of the co-occurrence of a case marker and a content word; it also shows that words that are affixed with voice morphology can occupy other syntactic positions, in this case, an NP position. The example in (4) provides some support for the precatergorial hypothesis (Foley, 1998) and contradicts the morpho-lexical classes proposed by Himmelmann (2008).

[5] This table is adopted from Rau and Dong (2006) with minor modification of the last column. The case marker for common noun *so* is analyzed as an oblique case marker in Rau and Dong's study.

Furthermore, in the symmetrical voice system (Huang, 2011) proposed for Tao, there is no unmarked main predicate in the basic clause and the verbal morphology always codes the choice of the sentence pivot. In all voice constructions, verbs are coded with voice affixes *(-om-, ma-, -en, -an, i-,* etc.), except the perfective PV construction due to the complementary distribution of the perfective prefix *ni-* and the PV affix *-en*.

Based on these previous works done on the verbal affixes in Austronesian languages and Tao, this paper investigates the properties of the verbal voice affixes in Tao and seeks to justify the functional category of these voice affixes that trigger subject shifting. In the following section, properties and characteristics that have been applied to distinguish inflectional from derivational morphology cross-linguistically are listed and discussed and used as the basis for the further examination of the voice affixes *–ma-, -om-, -en, -an*, and *i-* in section four.

3 Properties of Inflectional and Derivational affixes

It is generally accepted that derivational morphology is less productive than inflectional morphology, due to the syntactic obligatoriness of inflectional morphology. Inflectional morphology is concerned with syntactically driven word-formation (Katamba 1993: 205), while derivational morphology creates new lexical items. Furthermore, inflectional morphology does not change the word category of the stem, whereas derivational morphology may or may not do so. Lastly, inflectional morphology does not cause meaning change; the form and meaning of inflected words are predictable. Derivational morphology, by contrast, tends to bring in new meaning, or cause the meaning to change. These properties are briefly summarized in Table 3.1.

Table 3.1: Properties of Inflectional and Derivational morphology

	Inflection	**Derivation**
Syntactic Obligatoriness	✔	✗
Productivity	✔	✗
Predictability	✔	✗
Create a new lexical item	✗	✔
Change the word-class	✗	✔
Change meaning	✗	✔

Syntactic obligatoriness is illustrated by the English examples in (5); a well-known example of productivity is the English inflectional past tense suffix *-ed* that is attached to verbs to indicate past events. Conversely, the English derivational nominal-forming suffix *-ant* can only be attached to a certain class of words, as shown in (6). A new lexical item is created when a derivational affix is attached to it, as shown in (7); at the same time the word-class has changed from noun to verb, and the newly-attached derivational affix has changed the meaning of the base word. The English examples for the property of predictability are illustrated in (8), where the function of the affix is predictable and widespread.

(5) **Syntactic Obligatoriness**

a. *this book* *these book**s*** **these book* **this books* → inflectional

b. *The farm**er** is in the barn.* → derivational

The cow is in the barn.

(Katamba 1993: 206)

(6) **Productivity**

a. English inflectional past tense suffix *-ed*

b. Derivational agentive nominal forming suffix *-ant*

apply → *applicant*

donate → **donant*

(Katamba 1993: 207)

(7) **New Lexical Item; Different Word-Class; New Meaning**

ducks → *duckling*	*-ling*: derivational; it changes meaning to 'small duck'
sheep → *sheepish*	*-ish*: derivational; it changes word-class and meaning to 'like a sheep'
humour → *humourless*	*-less*: derivational; it turns a noun into an adjective and adds the meaning 'lacking' (e.g. humour)'

(Katamba 1993: 48)

(8) **Predictability**

Affix	**Stem**	**Function**	**Example**
-s	N	plural	*book-s*
-s	V	3rd person, singular, present tense	*sleep-s*
-ed	V	past tense	*walk-ed*
-ing	V	progressive (incomplete action)	*walk-ing*

(Katamba 1993: 51)

The above properties seem to suggest that the distinction between inflectional and derivational morphology can be easily maintained. However, the distinction between these two notions does not seem to have a clear boundary, and the properties suggested above cannot always serve to distinguish inflectional from derivational morphemes. As for the property of syntactic obligatoriness, there are cases where, for the sentence to be grammatical, a form affixed with a particular derivational morpheme is required, as shown in (9). In order for the sentence in (9a) to be grammatical, the adverb '*awkwardly*' has to be used instead of adjective '*awkward*'. Similarly, the noun '*teacher*' has to be used instead of the verb '*teach'* in (9b). Besides, some morphemes seem to have dual roles. For example, as suggested by Matthews (1974: 53-4), the V-*en*/V-*ed* forms could either be classified as an inflectional morpheme when marking the past participle of the verb, or a derivational morpheme when it marks the change of a verb into a participial adjective, as illustrated in (10). A participial adjective like *crowded* could be modified by the adverb *very* as in (10a') while the past participle *heated* cannot be so modified, as in (10b'). Hence, the boundary between inflectional and derivational morphemes is not clear-cut.

(9) a. *I opened it* ***awkwardly****. / *I opened it* ***awkward****.*

b. *The* ***teacher*** *is in the office. / *The* ***teach*** *is in the office.*

(Katamba 1993: 207)

(10) a. *a crowded room*→ participial adjective

a'. a very Adj N →*a very crowded room; a very small/clean/cold room*.

b. *a well heated room*→ verbal past participle

b'. a very Adj N →**a very heated room*

(Katamba 1993: 207)

Moreover, cross-linguistic comparison has shown that the same notion might be encoded by inflectional morphology in one language, but by derivational morphology in another language. For example, the diminutive suffix *-ling* (e.g., *duckling*) is derivational in English, but is found to be inflectional in many African languages. If the presence or absence of a morpheme is solely determined by the syntactic position it occupies and is purely syntactically determined, then it is inflectional rather than derivational. The distinction between inflectional and derivational morphemes is relative, not absolute, and is also language-specific. With these properties in mind, in the following section the Tao data are examined with the relevant criteria discussed in this section to determine whether the voice affixes in Tao are inflectional or derivational.

4 Tao Voice Affix – Inflectional?

4.1 Syntactic obligatoriness, productivity, and predictability

Tao, as discussed earlier, uses various voice constructions to signal the theta role of the subject NP. It has also been shown that in AV and PV constructions, some verbs are zero-marked for voice, as in (3e-f), reproduced in (11). This seems to indicate that the voice affix is not obligatory to ensure the grammaticality of a sentence. This is not the case in Tao. It is found that only a minority set of verbs is not affixed with voice affix in this language. Furthermore, the absence of the voice affix seems to occur with a specific set of verbs, such as stative-like predicates: e.g., *arako* 'big' in (12). Similarly, English has a set of verbs that has irregular past tense inflectional morphology. The absence of past tense suffix *-ed* in the past tense verb *hit* in the sentence '*he hit me yesterday'* does not indicate that the past tense suffix *-ed* in English is not obligatory. Hence, the absence of voice affixation in a minority set of verbs in Tao as in (11a) and (12) does not constitute a convincing evidence for non-obligatoriness of voice affixes in this language. For the absence of a patient voice affix in (11b), as mentioned in an earlier section, it could be easily explained that the perfective prefix *ni-* is a portmanteau morph that has dual functions, one to mark patient voice and the other to indicate perfective aspect. It can be stated that the voice affixes always attach to verbs in Tao. In other words, from this perspective, the affixation of voice affixes is productive.

(11) a.	*ya*	*Ø-apiya*	*o*	*wakay*		
	AUX	AV-good	NOM	yam		
	'The yam is good.'				(Shih 1997: 34)	
b.	*ya-na*	*ni-kan*	*ni*	*mapapo*	*o*	*soli*
	AUX-3.S.GEN	PV.PERF-eat		GEN PN	NOM	taro
	'Mapapu has eaten up the taros.'			(Ho 1990: 74)		

(12)	*ya*	*Ø-arako*	*o*	*kayo*
	AUX	AV-big	NOM	tree
	'The tree is big.'			(Shih 1997: 41)

Unlike other voice constructions (i.e., PV, LV, IV) it seems that the agent voice in Tao has more than one form, as illustrated in (13). The set of verbs that these agent voice affixes can be attached to remains unclear and requires further study. The diversity of agent voice affixes in this language might raise questions regarding their productivity and predictability. Nonetheless, these affixes share a common property, that is, their forms all

contain the letter *m*, despite the diversity of their forms and distributions. The affix *om* surfaces either as a prefix or infix as in (13a) and (13b), respectively; *man-*, *ma-*, and *mi-* surface as prefixes as in (13c-e). It can be claimed that the underlying form of agent voice affix in Tao is /m/. This claim has support in the language data. Shih (1997) claimed that there are phonological changes involved during the affixation process. For instance, for verbs that are prefixed with *pi-*'do, make or possess', affixation involves the deletion of the bilabial consonant [p], while the *m-* agent voice affix is prefixed to the stem *pi-talili* (*m-pi-talili* → *mi-talili*), as illustrated in (14). This example provides supporting evidence for the claim that the underlying form of agent voice affix in Tao is /m/. Once this is established, the productivity of the voice affixation process is without any doubt in Tao. Verbs are always coded with voice information via affixation of voice affixes – */m/, -en, -an, i-*, …etc. The origin of the agent voice affixes may come from different origins, or maybe historical residues. All of these are beyond the scope of the present study and require further investigation.

(13)

a. *ya* ***om****-lavi* *o* *kanakan*
AUX AV-cry NOM child
'The child is crying.' (Shih 1996: 27)

b. *ya* *k-**om**-alat* *do* *tokon* *si* *Mapay*
AUX AV-climb LOC mountain NOM PN
'Mapay is climbing in the mountain.' (Shih 1996: 40)

c. *ya* ***man****-bakbak* *si* *yama* *so* *kanakan.*
AUX AV-hit NOM father SO child
'My father is hitting/is about to hit the child.' (Chang 1997: 35)

d. ***ma****-rakat* *o* *ino.*
AV-kill NOM dog
'The dog is dead.' (Chang 1997: 13)

e. *ya* ***mi****-vali-valiat* *so* *vatu*
AUX AV-RED-turn SO stone
'He turns around the stone.' (Shih 1996: 59)

(14) *ya* ***mi****-talili (**m-pi**-talili)* *si* *Mapay*
AUX AV-jacket NOM PN
'Mapay was wearing a jacket.' (Shih 1996: 74)

Starosta (2002) has claimed that, if focus/voice affixes were inflectional morphemes, every verb should have a form for every voice construction in the paradigm. He suggested that it is rare to find a verb that has forms for all the voice variants in Tagalog (p. 569); in other words, voice affixes are not as productive as would be expected if voicing is derivational. In Tao as in Tagalog, it is rare to find a verb that has all the voice variants. Stative-like predicates or one-place predicates only have one voice variant, AV, as such predicates only involve one participant, and the sole participant may be a theme instead of an agent, as in (15). Predicates that involve more than one participant usually have at least two voice variants, depending on the thematic structure of the predicate. Hence, where

verbs lack the full set of voice variants in this language, this is mainly due to semantic, not morphological factors.

(15) *ma-apsi* *o* *gazas.*
AV-break NOM glass
'The glass broke.' (Chang 1997: 14)

4.2 Change lexical meaning, change word class, and create new lexical item

Judging by certain criteria, the Tao language data does not support the claim that voice affixes in the language are inflectional morphemes: specifically, the affixes do not exhibit the properties of syntactic obligatoriness, productivity, or predictability which are typical of inflectional morphology. However, examination of other properties of these voice affixes might help us to justify their morphological category. Relevant properties include whether the affix changes lexical meaning, changes word classes, or creates new lexical items. If Tao voice affixes are derivational affixes, then the affixation of these voice affixes will change lexical meaning, change lexical class, and create new lexical items after affixation.

It is not an easy task to determine the word classes of Tao lexical items alone. The lexical category of Tao words is not distinguishable unless they are put into syntactic slots, as suggested by the precategorial hypothesis discussed earlier (Foley, 1998). The direction of noun and verb formation in this language would still be difficult to determine even with the evolution of focus/voice in Austronesian languages proposed by Starosta, Pawley, and Reid (1982) and supports by researchers working on this subfield (Reid, 1992; Rau, 2002; Himmelmann, 2008; Kaufman, 2008;Ross, 2009; among others).

On the basis of these previous studies done on Austronesian languages and voice affixes, and Rau's work (2002) on Tao nominalization showing that Tao reflects the three stages of reconstruction proposed by Starosta et al. (1982), it is reasonable to claim that present-day Tao verbs are the result of reanalyzed nominalization. Tao verbs originated from nouns historically, thus explaining why Tao voice affixes are also found in nominalizations. This cross-category distribution seems to indicate that these affixes can change the lexical category of words in Tao. However, in view of Starosta et al.'s (1982) suggestion as to the evolution of Austronesian focus/voice, this apparent derivational property should not be taken into consideration in justifying the functional category of Tao voice affixes.

Grammaticalization is the term that is widely used to refer to this type of linguistic change (Hopper and Traugott, 2003). Hopper and Traugott have suggested some characteristics of the linguistic form after grammaticalization takes place. First, the coexistence of earlier and later forms. Second, the meaning of the earlier form may constrain the meaning or structural characteristics of the resulting derived form (Hopper and Traugott, 2003: 17). Both characteristics are observed in Tao. Nouns and verbs with identical forms are found and the voice affixes that originated from nominalization constrain the choice of sentence pivot in this language. Hence, Tao nouns and verbs should be treated differently as well as the voice affixes and their counterparts in nominalizations. That Tao voice affixes do not create new lexical items and do not change meaning are here proposed based on the premise that the verbs including voice affixes are the results of reanalysis. This premise can be supported with new coined words in Tao in (16). The root *vatek* meaning 'carving' is used to coin new word for new concepts, such as the action of writing or the one who writes. Notice that the *mi-* prefixes to two different stems: *vat-vatek* (CVC reduplication) and *vate-vatek* (CVCV reduplication) to coin new verb and noun in (16a) and (16b), respectively. These two words indicate that the native speakers do

distinguish nouns from verbs even though these two words are affixed with the same affix *mi-*, and the function of the prefix *mi-* in these two examples is not the same. The *mi-* in (16a) is clearly an agent voice affix that signals that the theta role of the subject NP is agent (referring to the one who writes or attends school) while the *mi-* in (16b) is a nominalization affix. Notice that *mi-* is the result of phonological change at the point in the derivation when *m-* is prefixed to the affix *pi-* meaning 'do, make or possess'. The verbal form of 'write' is derived while the root undergoes CVC reduplication, the prefix *mi-* is attached later on to signal the theta role of the sentence pivot. Affixation of voice affix *mi-* does not change lexical meaning, nor change word class nor create a new lexical item in (16a).

(16) a. *mivatvatek* (CVC) V 'write, attend school'
b. *mivatevatek* (CVCV) N 'student'

The inflectional and derivational properties of Tao voice affixes are summarized in Table 4.1. Notice that the column for changing the word-class for Tao voice affixes could be either left unspecified based on the precategorial hypothesis or as not changing the word-class based on the process of new coined word discussed in the last paragraph. The category of words is mainly determined by their syntactic position in Tao, except coined words. Furthermore, it is found that Tao voice affixes behave more like inflectional morphemes. They exhibit more properties that are associated with inflectional morphemes. Hence, it seems more plausible to consider Tao voice affixes as inflectional affixes despite the presence of negative evidence.

Table 4.1: Properties of Tao Voice Affixes

	Tao Voice Affixes	Inflection	Derivation
Syntactic Obligatoriness	✔	✔	✗
Productivity	✔	✔	✗
Predictability	[✔]	✔	✗
Create a new lexical item	✗	✗	✔
Change the word-class	-- (✗)	✗	✔
Change meaning	✗	✗	✔

Furthermore, the ungrammaticality of forms with two different voice affixes and their relative order with other affixes suggest that they should be classified as inflectional morphemes. A predicate that is affixed with two voice affixes is not possible, as shown in (17). The co-occurrence of agent voice affix *ma-* and patient voice affix *-en* renders the sentence ungrammatical. Notice that multiple derivational processes are allowed if the meanings of the affixes are compatible, as the affixation of *-ize* after the affixation of *-al* to form the verb *nation-al-ize* in English.

(17) **ma-apsi-en* *u* *gazas.*
AV-break-PV NOM glass (Chang 1997: 14)

While multiple affixation of voice affixes is not allow in Tao, Tao voice affixes always attach to the edges of the word with few exceptions. These exceptions include the agent voice infix *-om-*and the agent voice affixes prefix /m/ that is attached to the stem before the perfective affix *ni-*, as in (18a). The relative order of voice affixes can be observed from examples in (18b-d). The agent and beneficiary voice affixes *i-* and *ma-* in (18b-c) are further away from the stem while the causative prefix *pa-* is affixed closer to the stem.

From the example in (18d), it can be observed that the reduplication process has taken place earlier than voice affixation.

(18)

a. *ya* ***ni-mi****-ka-za-zakat* *sira* *o* *koyis*
AUX PERF-AV-KA-RED-kill 3.P.NOM NOM pig
'The pigs died one after the other.' (Shih 1996: 119)

b. *ko* ***i-pa****-sne-snek* *si* *kaka*
1.S.GEN BV-CAUS-RED-shamed NOM brother
I made my brother fell ashamed. (lit. I ashamed my brother)
(Shih 1997: 87)

c. *ya* ***ma-pa****-mying* *si* *Mapay*
AUX AV-CAUS-laugh NOM PN
Mapay is good at making others laugh.
(Shih 1997: 81)

d. *ya* *ni-**mi**-ka-za-zakat* *sira* *o* *koyis*
AUX PERF-AV-KA-RED-kill 3.P.NOM NOM pig
'The pigs died one after the other.' (Shih 1996: 119)

Lastly, pair examples in (19) that differ only in the voice affixation of the predicate are frequently found in Tao and these examples show the function of voice affixes in Tao is solely to trigger subject shifting. The irregular form, gaps, unpredictability, semi-productive forms, etc. discussed earlier in this section might be due to the historical residue or different degree of grammaticalization of individual words.

(19)

a. *ya* *k-**om**-alat* *do* *tokon* *si* *Mapay*
AUX AV-climb LOC mountain NOM PN
'Mapay is climbing in the mountain.' (Shih 1996: 40)

b. *na* *kalat-**en*** *o* *tokon* *ni* *Mapay*
3.S.GEN climb-PV NOM mountain GEN PN
'Mapay is climbing the mountain.' (Shih 1996: 40)

5 Conclusion

The function of signaling the agreement between the theta role of the subject and the Tao voice affix can easily lead people to consider the voice affix in Tao to be an inflectional morpheme, by analogy with English voice morphology. Researchers working on other Austronesian languages have claimed that voice affixation is derivational. Our attempt to resolve the contradiction lead us to a more detailed examination of Tao voice affixes to justify their inflectional nature. The majority of Tao voice affixes are found to be syntactically obligatory, their form and meaning are predictable, they are just as productive as English inflectional affixes are, and they do not contribute additional meaning to the word they are affixed to. However, some negative evidence such as the diversity of AV verbal voice affix and absence of PV voice affix in perfective sentence is found in the language data. It seems that the choice of AV verbal affix is constrained by verb classes that have weakened our proposal. But the diversity of AV voice affix and complementary distribution of perfective prefix *ni-* and PV voice suffix *-en* might be the residues from

historical changes. Notice that these verbal voice affixes were proposed as "second-generation verbal affixes" (Ross 2009:301) in historical linguistic research. In examining current morphological properties of Tao verbal affixes, I have found that the distinction between inflection and derivation in Tao is not absolute, but relative. Hence, a continuum with inflectional and derivational morphemes at two extremes should be proposed for Tao morphology. As Tao voice affixes have exhibited more inflectional properties as language data have shown, it seems more plausible to classify them as inflectional morphemes while they participate in signaling the theta role of the nominative case-marked NP. However, these affixes should be treated differently when they are found in noun phrases that are headed by case markers. Further investigation is required to clarify the diversity of the AV voice affixes found within the language data.

References

Chang, Henry Yungli. 1997. Voice, case and agreement in Seediq and Kavalan. Ph. D. dissertation, National TsingHua University, Hsinchu.

––– 2004. AF verbs: transitive, intransitive, or both? *Language and Linguistics MonographSeries Number W-4*: 95-119. Taipei: Institute of Linguistics, Academia Sinica.

Chang, Claire Hsun-huei. 1997. On thematic structure of Yami verbs. NSC report 85-2418-H-004-001.

––– 2000. *Yami reference grammar*. Taipei: Yuen Liu.

Deng, Dun-Hong. 2005 *Argument-Function Linking in Yami: An Optimality-Theoretic Account.* MA Thesis. National Cheng-Chi University.

Foley, William, A. (1998) Symmetrical voice systems and precategorialtiy in Philippine languages. University of Sydney.

Himmelmann, Nikolaus P. 2008. Lexical categories and voice in Tagalog. In Peter K. Austin and Simon Musgrave, eds. *Voice and Grammatical Relations in Austronesian Languages*. Standford: CSLI.

Ho, Arlene Yue-ling. 1990. *Yami Structure: A Descriptive Study of the Yami Language.* M.A. Thesis, National TsingHua University.

Hopper, Paul J. and Elizabeth C. Traugott (2003). Grammaticalization. Cambridge: Cambridge University Press.

Katamba, Francis. 1993. *Morphology*. New York: Palgrave.

Kaufman, Daniel. 2009. Austronesian nominalism and its consequences: A Tagalog case study. *Theoretical Linguistics* 35.1:1-49.

O'Grady, Willian. 2004. The syntax files: an introductory survey of basic syntactic concepts and phenomena. Manuscript. Honolulu: Department of Linguistics, University of Hawai'i at Mānoa.

Rau, D. Victoria. 2002. Nominalization in Yami. *Language and Linguistics* 3.2: 165-195

––– 2004. Iconicity and temperal-aspectual-modality in Yami. Paper presented to the Ninth Insternational Symposium on Chinese Languages and Linguistics, Graduate Institute of Linguistics, National Taiwan University, November 2004.

Rau, D. Victoria. 2002. Nominalization in Yami. *Language and Linguistics* 3.2: 165-195. Taipei: Academia Sinica.

Rau, D. Victoria. & Tung, Ma-nuy. 2006. *Tao Dictionary: Reference Grammar, Texts, and Dictionary*. Language and Linguistics Monograph. Taipei: Academia Sinica.

Reid, Lawrence A. 1992. On the development of the aspect system in some Philippine languages. *Oceanic Linguistics* 31.1: 65-69

Ross, Malcolm D. 2009, Proto Austronesian verbal morphology: A reappraisal, in Alexander Adelaar and Andrew Pawley (ed.), *Austronesian Historical Linguistics and Culture History: A festschrift for Robert Blust*, Pacific Linguistics, Canberra Australia, pp. 295-326.

Ross, Malcolm &Teng, Stacy Fang-ching. 2005. Formosan Languages and Linguistic Typology. *Language and Linguistics* 6.4:739-781.

Shih, Louise Yu-mien, 1997. Yami morphology. MA thesis. Providence University, Taichung, Taiwan.

Starosta, Stanley, Andrew K. Pawley, and Lawrence A. Reid. 1982. The evolution of focus in Austronesian. *Papers from the Third International Conference on Austronesian Linguistics*, Vol. 2: *Tracking the Travellers*, ed. by AmranHalim, Lois Carrington and S. A. Wurm, 145-170. Pacific Linguistics C-75. Canberra: Australian National University.

Starosta, Stanley. 2002. Austronesian 'Focus' as derivation: Evidence from nominalization. *Language and Linguistics* 3.2:427-479

Traugott, Elizabeth Closs. 2010. Grammaticalization, in Silvia Luraghi and VitBubenik, eds., *Continuum Companion to Historical Linguistics*, 269-283. London: Continuum Press.

14 *Benefactives in the Sembiran Dialect of Balinese*

I NYOMAN SEDENG

1 Introduction

The Sembiran dialect of Balinese (SDB), a western Austronesian language, is spoken by 4883 language users inhabiting the mountainous village of Sembiran, Bali.[1] This dialect is actively used by the population of the village in every aspect of their daily activities. Morphologically, SDB is an agglutinative and head-marking language; verb marking indicates the argument structure of the clause. There are four diathesis alternations found in the dialect: agentive, objective, passive, and medial, although only the first three are discussed here. Accordingly, this paper aims at giving a description of the characteristics of benefactives in SDB. The discussion is framed in certain aspects within the RRG framework (Foley and Van Valin 1984; Van Valin and LaPolla 1997). There are several strategies in SDB for expressing benefactive meaning. In particular, the dialect has two productive verbal suffixes, *-ang* and *–in*[2], which have an applicative function, introducing a benefactive argument in certain contexts. The main claim about benefactive in SDB is that they are constructionally expressed by means of applicative morphology, in combination with argument flagging involving the (first) object NP and oblique PP, and that there is no single dedicated construction for benefactive meaning. That is, the same applicative morphology on a verb can enter different beneficiary-related constructions, giving rise to variations in the nature of benefactive meanings.

The paper is structured as follows. Section 2 provides a brief overview of basic sentence structure and grammatical relations in SDB, and section 3 discusses morphosyntactic

[1] SBD should be distinguished from the Plains Balinese dialect, which, lexically, has been influenced by many other languages, namely, Javanese, Sanskrit, English, Arabic, and Indonesian. Morphologically, both dialects have slight differences in prefix and suffix systems, but syntactically, both dialects have the same syntactic marking typology.

[2] These two suffixes, *-ang* and *-in,* play an important role in verb formations. As mentioned, their applicative function (e.g., *Iyya* ***ngesop*** *nasi* 'She eats rice', *Iyya* ***ngesopin*** *panakanné nasi* 'She is feeding her child') is mainly what concerns us here. However, both can have another valence-increasing function, as a causative morpheme (e.g., *Andipi-né* ***mati*** 'The snake is dead', *Okké* ***nge-mati-ang*** *andippiné nto* 'I killed the snake.'*).* Furthermore, both suffixes may also function as a verbaliser that may be attached to the noun (e.g., *bubuh* 'porridge', *bubuh-ang* 'cook some porridge for someone'), adjective (e.g., *tuh* 'dry', *tuh-ang* 'cause something to dry'; *belah* 'broken', *belah-ang* 'cause something to break'), or bound root (e.g., *orah* 'say'*, orah-in* 'tell', *orah-ang* 'mention').

Asia-Pacific Linguistics, 2014.

alignment. In section 4, I turn to the discussion of benefactive constructions, including the applicative morphology of SDB. Discussion of the semantically determined variation in the coding of beneficiaries, in section 5, and polysemy of benefactive markers, in section 6, then follows. The conclusion is given in section 7.

2 Grammatical Relations in SDB

Central to the discussion of grammatical relations, such as subject and object, are the relationships between the verb and the arguments it assigns. As such, in analysing clausal structures, the first concern is to identify the predicate, and the semantic roles of its arguments.

Either an agent or patient can be the sole argument of an intransitive verb, as illustrated in (1) and (2) below. In (1), the zero-marked verb *hhug* 'damage' assigns a patient, *umah-é* 'the house'. By contrast, in (2), the nasal verb[3] *mlahib* 'run' assigns an agent, *Maman Bangsing* 'Uncle Bangsing'.

(1) *umah-é* *hhug*
house-DEF OV.damage
'The house was damaged.'

(2) *Maman Bangsing* *m-lahib*
uncle [name] AV-run
'Uncle Bangsing ran away.'

However, in both cases, the sole argument linearly precedes the verb. We conclude that the subject must precede the predicate (although this may be reversed for pragmatic purposes).

Intransitive nasal verbs can also be followed by oblique arguments. In (3), *ntakon* is a di-intransitive verb that assigns an agent as its subject, but also assigns an oblique argument:

(3) *iyya* *langsung* *n-takon* *kén* *sommah-a-né*
3SG directly AV-question PREP wife-3-POSS
'He asked his wife something directly.'

In (4), *ntakon* takes the applicative suffix *–in,* and the oblique argument rises to become a core object argument.

---

[3] In Balinese varieties, including SBD, the nasal prefix undergoes morphophonological assimilation in which the nasal prefix is homorganic with the initial sound of the base verb and it has four possible realizations. The allomorphs of the prefix are as follows:

/ny-/ before consonants /s, j/, e.g., *sampat* 'broom' → *nyampat* 'sweep', *jepit*→*nyepit* 'hold something by clamping';

/m-/ before bilabial consonants /b, p/, e.g., *pecik* → *mecik* 'press', *batek*→ *matek* 'pull';

/n-/ before dental consonants /d, t/, e.g., *tanem* → *nanem* 'bury', *dacin* → nacin 'put sth. on scales';

/ng-/, a velar nasal, before velar consonants /k, g/ and all vowels, e.g., *getok*→*ngetok* 'knock', *kelés* → *ngelés* 'to loosen', *amah* →*ngamah* 'to eat', etc.

All initial consonants in the base forms of the verb are dropped after the nasal is attached.

(4) *iyya* *langsung* *n-(t)akon-in* *sommah-a-né*
3SG directly AV-question-APPL wife-3-POSS
'He asked his wife something directly.'

After being attached by suffix {-in} the initial consonant in the base verb is dropped. This phenomena take place when the initial sound of the base verbs are dental /d/ and /t/, such in, *Okké ba ntemu jak dadong* 'I have met the grandma' is di-intransitive clause and the transitive counterpart *Beli adda nemunin tappih adin okkéné* 'Brother, have you accidently met my younger sister's clothes?', and *Nanang ba ulih ndahhar ba? 'Father, have you eaten your meal?'*intransitive form becomes *Iyya nahhar nasin nanangé* 'He is consuming your food.'

The same contrast holds for monotransitive verbs. In (5), the zero-marked *gugut* 'bite' selects a patient as its subject, and in (6), the nasal *ngugut* 'bite' selects an agent as its subject. SDB therefore exhibits a two-way alternation between zero-marked verb forms, where the patient is the grammatical subject, and nasal verbs, where the agent is the grammatical subject. I call the former 'Objective voice' ('OV'), and the latter, 'Agentive voice' ('AV')

(5) *Nyoman Sadra* *gugut* *legu* (Objective voice)
[name] OV.bite mosquitos
'Mosquitos bit Nyoman Sadra.'

(6) *Legu mannya* *ng-(g)ugut* *Nyoman Sadra* (Agentive voice)
Mosquitos PART AV-bite [name]
'It seems that the mosquitos bit Nyoman Sadra.'

In (7), however, the base form *gugut* takes a suffix, *-ca*[4], marking Passive voice ('PASS.V'). Here, the patient is still the subject, but the agent is demoted to oblique status.

(7) *Nyoman Sadra* *gugut-ca* *ken* *legu-ne* (Passive voice)
[name] bite-PASS PREP mosquitos-DEF
'Nyoman Sadra was bitten by the mosquitos.'

Structures (5) – (7) therefore show that SDB exhibits a three-way distinction between OV, in which the verb is marked by a Ø prefix; AV, marked by a nasal prefix; and PASSV, marked by passive suffix.

SDB is a language with symmetrical double-object constructions (i.e. both objects of a di-transitive structure can alternate to become subject in OV structures). This phenomenon can be seen from the di-transitive predicate *behhang* in (8)-(10). In (8), where the verb has AV marking, the agent must be the grammatical subject. However, in (9) and (10), where both verbs have OV marking, either non-agent argument (i.e., either of *mémmé* 'mother' or *kaccang* 'long bean') can occupy the grammatical subject position:

[4] The phonetic realization of the passive suffix or third person singular clitic *-a* varies based on the last consonant closing the preceding syllable, as the last consonant is geminated, e.g., *mak-ka* 'take-PASS', *gugut-ca* 'bite-PASS', *amah-ha* 'eat-PASS', *ulihang-nga* 'return-PASS', and *allap-nya* 'pick up-PASS'.

(8) *Okké* *m-(b)ehhang* *mémé* *kaccang*
1SG AV-give mother long bean
'I gave mother some long beans.'

(9) *Kaccang* *behhang* *okké* *mémé*
long bean Ø.give 1SG mother
'It was some long beans that I gave to mother.'

(10) *Mémé* *behhang* *okké* *kaccang*
Mother Ø.give 1SG long bean
'It was mother to whom I gave some long beans.'

Typologically speaking, this distinguishes the SDB sentences (8)-(10) from their English counterparts with the verb *give,* also a three-argument verb. Example (11) shows several possible alternations where arguments of *give* may occupy subject position; as (11d) shows, however, the theme cannot do so.

(11) a. My uncle gave me some money.
b. I was given some money by my uncle.
c. Some money was given to me by my uncle.
d. *Some money was given me by my uncle

There are only two pure di-transitive verbs in SDB, namely *behhang* 'give' and *bayah* 'pay'. Other three-argument verbs are derived via applicative process. Some of the verbs are *mak* 'take' <x, y>, which becomes *makkang* 'take something for someone' <x, y, z>; *ejuk* 'catch' <x, y>, which becomes '*ejukkang* 'catch something for someone' <x, y, z); *beli* 'buy' <x, y>, which becomes *beliang* 'buy something for someone' <x, y, z>, and many others.

3 Active Grouping

Typologically speaking, languages can be classified into three types with respect to their syntactic alignment: accusative, ergative, and active-stative. Criteria defining each type are based on three primitive relations proposed by Dixon (1994).

S–sole argument of intransitive clause

A–agent of transitive clause

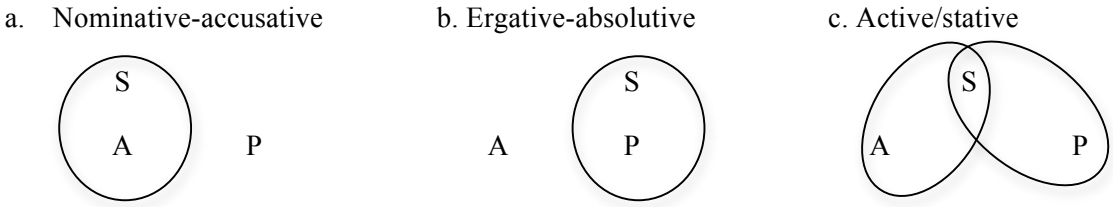
O/P–patient of transitive clause

In an accusative language, S and A have the same coding, which is distinct from O, i.e., S=A ≠ O. On the other hand, in an ergative language, S is coded in the same way as O/P, and is distinct from A, i.e., S= O ≠ A. Many languages have characteristics of both accusative and ergative alignment, in certain grammatical contexts coding S and A together, and in others coding S and O/P together.

Verhaar (1996:218) uses the term 'active-stative system' to describe alignment 'in between' accusative and ergative systems. In active-stative languages, the sole argument S is sometimes marked in the same way as A, as in the nominative system, and sometimes as O, as in the ergative system, depending on whether the sole argument is agent-like or

patient-like; i.e., $S_A = A \neq S_O = O$. The three different alignment systems are represented schematically below, as in a language where arguments are marked for case.

Diagram 3.1 Accusative, Ergative and Active-stative alignment



SDB exhibits active-stative alignment, in between accusative and ergative alignment, because the S of an intransitive predicate that assigns an agent argument is usually marked by the nasal prefix, which also encodes the A of a transitive predicate. By contrast, a patient S argument is marked by zero, which also encodes O/P in a transitive predicate.

(12) and (13) show patterns of accusative marking for agentive S. In (12), the verb *nahar*, marked by the nasal prefix realized as /n-/, assigns one argument: *nannang*, which has the semantic role agent. The transitive predicate *ngubuhin* in (13) is also marked by the nasal prefix, here realized as /ng-/, and the predicate assigns two arguments: *kaka*, an agent, and *panak brarakan*, a patient. Therefore, the S in (12) is encoded in the same way as the A in (13), i.e., Sa = A ≠ O; both verbs are marked by the nasal prefix.

(12) *Nanang iyya ba ulih n-(d)ahar*
father 3SG PERF from AV-eat
'Father has eaten his meal.'

(13) *Kaka ng-ubuh-in panak brarakan*
1SG AV-take.care-APPL children many
'I take care of many children.'

Items (14) and (15) illustrate that patient-like S shows ergative marking. That is, the S of an intransitive predicate is encoded in the same way as the O/P of transitive clauses, or $S_O = O \neq A$. Both verbs that fill the predicates of sentence (14) and (15) are zero-marked, and consequently both subject argument NPs have the thematic role of patient.

(14) *Gigi-na-né ba* <u>*hilang*</u> *asibak* OV:S_U=SUBJ
tooth-3-POSS PERF OV.loose partially
'His teeth have partially fallen out.'

(15) *Nang Landep cotot andipi* OV:U=SUBJ
[name] OV.bite snake
'Mr. Landep was bitten by a snake.'

4 Benefactives in SDB

Attention turns now to benefactives in SDB. Benefactive is a concept that describes a situation or an event in which someone receives benefit from someone else's action. Accordingly, the term 'benefactive' refers to a construction which introduces the person or thing that benefits from the action expressed by the verb. This person or thing is the beneficiary, a participant that is advantageously affected by an event without being its obligatory participant, either agent or primary target (i.e. patient) (Kittilä 2010:15). Furthermore, Kittilä (2010:15) notes that "since normally only animate participants are capable of making use of the benefit bestowed upon them, beneficiaries are typically animate."

In Role and Reference Grammar terms, the applicative construction belongs to the class of accomplishment verbs, with the logical structure [do' (x, ǿ)] CAUSE [BECOME **pred**' (y, z)]. *Pred* in the second part of the formula may be equal to class of verb meanings (e.g., to have, to know, to experience, to receive, to be at a given location). The following example makes this clear:

(16) *Senir m-(b)eli-ang kaka behhas*
[name] AV-buy-APPL 1SG rice
'Senir bought me some rice.'

The logical structure of (16) is [do' (Senir, ǿ)] CAUSE [BECOME **have**' (kaka, behhas)]. As for their argument structure, accomplishment verbs can be two-place or three-place verbs in SDB, as in (17) and (18) below.

(17) *Nyoman m-(b)ubuh-ang dadong*
[name] AV-porridge–APPL grandmother
Nyoman cooked some porridge for grandmother.'

(18) *Mbok Rati n-nyonyo-in panak-a-né.*
sister [name] AV-breast-APPL child-3-POSS
'Sister Rati breastfed her child.'

In (17), the verb *mubuhhang* 'cook porridge for', being derived from the one-place verb *mubuh* 'cook porridge', assigns two arguments, with semantic roles affecter and beneficiaries. In (18), the verb *nnyoyoin* 'breastfeed' is also a two-place verb that assigns two arguments, feeder and recipient. This verb similarly derived from the one-place verb *nnyonyo,* as in *Anake senik nto sedekan nnyonyo* 'The baby is breastfeeding.'

When someone conducts an action, the action may give benefit to someone else, or to the doer himself. This can be seen from the following two examples.

(19) *Okké m-(b)ehhan parang-é ené sig Rangsasa-né*
1SG AV-get pumpkin-DEF DEM PREP giant-DEF
'I got this pumpkin from the giant.'

(20) *I Rangsasa m-(b)ehhang okké parang-é ené*
ART giant AV-give 1SG pumpkin-DEF DEM
'The giant gave me this pumpkin.'

The logical meaning of the two sentences is the same, namely the process of transferring

the argument theme *parang* 'pumpkin' from the giant's possession to the recipient, *okké* 'I'. The same semantic relation involving the benefactive is expressed by two distinct, but semantically related, reverse lexical items, *behhang* 'give' and *behhan* 'get'. In the first sentence (19), it is the subject *okké* 'I' who is the benefactive of the event, and the oblique *sig Rangsasané* 'from the giant' specifies the source. On the other hand, in (20), the subject *I Rangsasa* 'the giant' is the source, and the first object *okké* 'I' is the beneficiaries of the act of giving.

4.1 Formal Mechanisms of Beneficiary Coding in SDB

Kittilä (2010:7) identifies four major mechanisms for encoding beneficiariess across languages, which include case marking, adpositions, serial verb constructions, and applicativization. SDB employs all these strategies except case marking, because SDB is a head-marking language. However, the first object grammatical function, distinguished via linear order, similarly encodes a beneficiary.

4.1.1 First Object

As mentioned by Kittilä, cross-linguistically, case-marking is usually used to encode the beneficiary. However, since SDB is not a case-marking language, the first object, as distinguished via linear order, encodes this function. There are a limited number of 'pure give' verbs which select for a first object, namely *behhang* 'give' and *bayah* 'pay'. However, there are more derived through applicative processes, examples of which can be seen below.

(21)	*Sidikara-né*	*ditu*	*m-(b)ehhang*	*okké*	*poh-hé*	*né*
	neighbour-DEF	there	AV-give	1SG	mango-DEF	DEM

'The neighbour there gave me these mangoes.'

(22)	*Merta*	*m-(b)ayah*	*okké*	*Rp. 300.000*	*me-gai*	*petang*	*dina*
	[name]	AV-pay	1SG	Rp. 300.000	AV-work	four	day

'Merta paid me IRD 300.000 for four days' work.'

(23)	*Kaka*	*ny-(s)erah-in*	*mémé*	*pipis*	*dasa*	*ringgit*
	1SG	AV-hand-APPL	mother	money	ten	ringgit

'I handed ten ringgit to my mother.'

The predicates *mehhang* 'give' (21) and *mayah* 'pay' (22) both assign three core arguments, i.e., Pred' <x, y, z>, and the argument *okké* 'I' is the beneficiary in both constructions. Both sentences have no alternative or derived form to express the same message of the sentences, unlike e.g., the equivalent English alternation between *She gave me the book* or *She gave the book to me.* The predicate *nyerahin* 'hand' <x, y, z> in (23) derives from the bound root *serah* and the verbal suffix *-in*; however, the root *serah* can also take the verbal suffix *-ang*, forming *nyerahang* 'hand', which assigns two core arguments and one oblique, i.e. <x, y><z>. The logical structure of sentence (21) is [do' (Sidikarané ditu, ø)] CAUSE [BECOME **have**' (*okké*, *pohhé né*)].

4.1.2 Prepositions

There are two prepositions in this dialect, namely *kén* 'to' and *sig* 'at', which mark the oblique grammatical function, and can introduce a benefactive:

(24) *Nang tuwa* *juwa* *ng-adep* *nyuh plukutus* *puhun*
uncle PART AV-sell coconut eighteen tree

kén *beli* *Dora*
PREP brother [name]
'It was uncle who sold eighteen coconut trees to Dora.'

(24) *Nang tuwa* *juwa* *ng-adep* *nyuh plukutus* *puhun*
Uncle PART AV-sell coconut eighteen tree

kén *beli* *Dora*
to brother Dora
'It was uncle who sold eighteen coconut trees to Dora.'

(25) *Mbok Sriman* *ng-ejot* *bé* *sig* *nanang tua*
sister [name] AV-send fish PREP father old
'Sister Sriman sent some food to uncle.'

Both verbs that fill the predicates of sentences (24) and (25) assign two core arguments and one oblique argument, **Pred'**<x, y><z>, and have corresponding applicative forms, in which the oblique argument is promoted to core argument status. The applicative suffix *-in* is attached, forming the derived structures, as can be seen in (26) and (27) below.

(26) *Nang tuwa* *juwa* *ng-adep-in* *<u>beli Dora</u>* *nyuh* *plukutus puhun*
Uncle PART AV-sell-APPL brother [name] coconut eighteen tree
'It was uncle who sold eighteen coconut trees to Dora.'

(27) *Mbok Sriman* *ng-ejot-in* *<u>nanang tua</u>* *bé*
sister [name] AV-send-APPL uncle fish
'Sister Sriman sent uncle some food.'

All the noun phrases underlined in the constructions above refer to the beneficiaries of the events expressed by their respective clauses. *Beli Dora* and *nanang tua*, which function as obliques, are promoted to first object position in the derived forms.

4.1.3 Serial Verb Constructions (SVCs)

Cross-linguistically, serial verb constructions are widely found in isolating languages, unlike SDB, which is agglutinative, as mentioned. Nonetheless, a limited number of such constructions are found in SDB. The term 'verb serialization' was developed by Foley & Van Valin (1984:198), based on language data from Anyi (Van Leynseele 1975), Kasem (Hewer 1976, cited in Foley & Van Valin 1984:199), Alambak (Bruce 1988:39), Abelam (Laycock 1965:55), and Imonda (Seiler 1986:14, cited in Durie 1997:307). There are two productive verbs in SDB which form serial verb constructions, and these reverse each other in meaning. They are *behhan* <x, y><z> 'get' and *behhang* <x, y, z> 'give'. Both are three-place verbs: the former assigns two core arguments and one oblique, while the latter assigns three core arguments. In a serial verb construction, both verbs occupy the first

verbal position, and can form simplex characteristic SVCs (i.e., the serialization occurs at the nuclear juncture level)[5] as in (29), or complex characteristic of SVCs (i.e., the serialization occurs at the core juncture level), as in (28).

(28) *Okké* ***m-(b)-ehhang*** *Man Muji* ***ny-(s)ilih*** *pipis*
1SG AV-give [name] AV-borrow money
'I let Man Muji borrow some money.'

(29) *Okké* ***m-(b)ehhang*** ***ng-idih*** *Komang* *pipis*
1SG AV-give AV-beg [name] money
I gave Komang some money as a gift.'

In both (28) and (29), the first verb in the SVC, *behhang*, expresses the process of releasing ownership of something to the recipient. The second verb specifies the method intended by the giver in releasing his ownership, either by a kind of non-permanent movement, as in *silih* 'borrow', or by the owner releasing ownership as a gift, such that the recipient has legal ownership of the item, as in *idih* 'beg'. Both sentences (28) and (29) have corresponding non-serialised forms derived by the attachment of the applicative suffix *-ang*, as can be seen in examples (30) and (31). One important point to be seen in these two examples is that in SDB, the meaning of 'lend' and 'borrow' is based on the verb base *silih:* the form *nyilih* means 'borrow', while *nyilihang* means 'lend'. The same process also applies in the case of 'beg' and 'give': is based on the base verb *idih* 'beg' and *ngkidihang* 'give'.

(30) *Okké* ***ny-(s)ilih-ang*** *pipis* *kén* *Man Muji*
1SG AV-borrow-APPL money PREP [name]
'I lent Man Muji some money.'

(31) *Okké* *ng-**kidih**-ang* *pipis* *kén* *Komang*
1SG AV-beg-APPL money PREP [name]
'I gave Komang some money as a gift.'

Viewed from the derived forms of both sentences, the second verbs *nyilih* and *ngidih* function as the 'major' verb in their respective SVCs, since they function as the main verb in the derived form, whereas *behhang* functions as the 'minor' verb, which then becomes the applicative suffix *-ang* in the derived forms.

The other verb that may emerge in the first position of the serial verb construction is *mehhan,* derived from the base form *behhan,* 'get'. There is a completive aspectual meaning in this context, but it may also be used to express future expectation. In the situation when a speaker doesn't have money that he urgently needs, he may go to someone who he thinks has some, and say '*Madak mehhan nyilih pipis sig I Jigehé',*

[5] RRG proposes two types of complex predicate; i) the 'nexus' type, which covers subordination, coordination, and co-subordination, and ii) the 'juncture' type, which covers nuclear, core, and clausal junctures. A nuclear-level juncture means that the predicate is constructed from two verbs, but without being interrupted by object of the first verb, and that the SVC behaves like a simplex predicate (e.g., ***m-(b)ehhang ng-idih***). A core-level juncture means that the predicate is a serial verb construction, but the verbs are interrupted by the object (e.g., ***m-(b)-ehhang Man Muji ny-(s)ilih***), and this structure behaves like complex predicate.

'Hopefully Jigeh may lend me some money.'

In the context of the cultural use of the verb, when you bring something home, e.g., money, (in a greater amount than what your relative thinks that you might have), your relative at home may ask you the following question:

(32) *Japa jah ngko m-(b)ehhan pipis a yuta?*
where PART 2SG AV-get money a lot of
'Where did you get so much money?'

A typical answer to such a question may employ an SVC, where *behhan* is the major verb, and the minor verb specifies the manner, e.g., *mehhan ngidih* 'beg', *mehhan nyilih* 'borrow', *mehhan nuduk* 'find by chance', or *mehhan mutranin* 'get as a loan'. The structure of this SVC is generally at the nucleus juncture level.

(33) *Uwa Budi m-(b)ehhan m-(b)eli bé-né nto di pasih*
uncle [name] AV-get AV-buy fish-DEF DEM PREP beach
'Uncle Budi got the fish by buying it on the beach.'

(34) *Okké m-(b)ehhan n-(d)uduk dompét-é né di marga-né*
1SG AV-get AV-pick.up purse-DEF DEM PREP street-DEF
'I found the purse by chance on the street.'

(35) *Okké m-(b)ehhan ng-idih labu-né né sig*
1SG AV-get AV-beg pumpkin-DEF DEM PREP

dadong raksasa-né
grand giant-DEF
'I got the pumpkin from the grand giant for free.'

The verb *behhang* 'give' may also function as the minor verb that occupies the second verbal position of the SVC. This serial verb construction is constructed by two- and three-place predicates, and is always in the structure of bi-clausal, i.e., Pred' <x, y> Pred' <x, y, z>. To clarify, below are some examples found in the data corpus.

(36) *Miasa m-(b)eli baju behhang-nga mémé*
[name] AV-buy shirt OV.give=3 mother
'Miasa bought a shirt for me.'

(37) *Okké ng-alih basé behhang I dadong*
1SG AV-pick betel.leaves OV.give ART grandmother
'I picked up betel leaves for grandmother.'

(38) *Okké ng-(g)oréng jaja behhang Wayan Ganing*
1SG AV-fry cakes OV.give [name]
'I fried some cakes for Wayan Ganing.'

The first verbs in the construction above assign two core arguments: (36) *meli <Miasa, baju>*, (37) *ngalih < Okké*, *basé* >and (38) *ngoréng < Okké, jaja>*. These verbs are major verbs, whereas the second verb, *behhang,* is a minor verb. Since in this dialect there is no preposition corresponding to the English preposition *for*, the serial verb construction with *behhang* has a similar function to the preposition. The verb *behhang* assigns three core

arguments: in item (36), *mehhang <Miasa, mémé, baju>,* (37) *mehhang <okké, dadong, basé >,* and (38) *mehhang <okké, Wayan Ganing, jaja>*. The characteristics of each verb can be seen in its derived applicative form below, in which it is the first verb that maintains its status as the main predicate, while the second verb is replaced by the applicative suffix *-ang*.

(39) *Miasa m-(b)eli-ang mémé baju*
[name] AV-buy-APPL mother blouse
'Miasa bought mother a blouse.'

(40) *Okké ng-alih-ang I dadong basé*
1SG AV-pick-APPL ART grandmother betel.leaves
'I picked grandmother some betel leaves.'

(41) *Okké ng-(g)oréng-ang Wayan Ganing jaja*
1SG AV-fry-APPL [name] cakes
'I fried some cakes for Wayan Ganing.'

4.1.4 Applicativization

Beneficiaries are often introduced via applicative constructions. SDB applicative morphology is introduced below. There are two productive applicative suffixes found in SDB, *-ang* and *–in*. There are three types of prominent applicative constructions proposed by Katamba (1993:270-273): namely, those that promote beneficiaries, those that promote locatives, and those that promote possessor*s;* however, additional types exist, as shown below.

Applicative *-ang*

emak 'to take'	*mak-ang*	'take something for someone' (benefactive)
bedbed 'to bandage'	*bedbed-ang*	'bandage something on' (instrument)
adep 'to sell'	*adep-ang*	'to sell something for someone' (benefactive)

Applicative *-in*

emak 'to take '	*emak-in*	'to take something from someone' (source)
pula 'to plant'	*pula-in*	'to plant something on' (locative)
adep 'sell'	*adep-in* 'to sell something to someone' (goal)	

There are two applicative suffixes in SDB, namely *-ang* and *-in*, and applicative verbs can be derived from various underlying structures. In (42), *ngemak* 'take' assigns two core arguments, the second of which is an NP is in a possessive structure. The derived form (43) is an applicative construction, derived via possessor rising: the possessive *nannangané* rises to the position of first object, so the verb *ngemakin* assigns three arguments <taker, source, taken>. The applicative suffix applied in this context is *–in*, and the beneficiary *nannangané* 'his father' has the semantic role of source. The father is not the primary target of the events he benefits from; rather, he gets some benefit from the result of the action. The target of the action is *babayan* 'the load', and after it is fetched from *nannangané,* he is relieved of his burden

(42) *Senir ng-emak babayan nannang-a-né*
[name] AV-take load father-3-POSS
'Senir took/fetched his father's load.'

(43) *Senir* *ng-emak-in* *nannang-a-né* *babayan*
[name] AV-take-APPL father-3-POSS load
'Senir took/helped with his father's load.'

The verb root *emak* can also take the verbal suffix *–ang*, as in (44). Here, the first object function is assigned to the theme *babayan,* so it is called the putative beneficiaries.

(44) *Senir* *ng-emak-ang* *nannang-a-né* *babayan*
[name] AV-take-APPL father-3-POSS load
'Senir took/helped with his father's load.'

Applicative verbs can also substitute for serial verb constructions with *behhang* as the second verb, at either core or bi-clausal levels, as can be seen in (45). In the corresponding applicative construction (46), the beneficiary is in the first object position.

(45) *Nanang* *ng-alap* *poh* *behhang-nga* *Nyoman Gari*
father AV-pick.up mango OV.give=3 [name]
'Father picked up some mangoes for Nyoman Gari.'

(46) *Nanang* *ng-alap-ang* *Nyoman Gari* *poh*
father AV-pick.up-APPL [name] mango
'Father picked up some mangoes for Nyoman Gari.'

Applicativisation may also involve the alternation of an oblique element to become the first object. Two prepositions in this dialect, *kén* 'to' and *sig* 'at', function as oblique markers.

(47) *Kaka* *ng-adep* *coklat* *sig* *Made Runti-né*
1SG AV-sell cocoa PREP [name]-DEF
'I sold chocolate to Made Runti.'

(48) *Kaka* *ng-adep-in* *Made Runti* *coklat*
1SG AV-sell-APPL [name] cocoa
'I sold chocolate to Made Runti.'

The oblique function in sentence (47) is marked by a combination of the preposition *sig* and definite clitic *-né* attached to the noun head, such that the name *Made Runti* has a locative meaning. (48) is derived through the process of applicativisation: the preposition *sig* and the clitic *-né* that form the PP oblique are deleted in the derived form, and only the head noun is promoted to the first object position.

5 Semantically determined variation in the coding of beneficiaries

As Kittilä (2010:14) notes, beneficiaries may benefit from an event in different ways, and the benefactive role may be encoded differently in each case. Kittilä (2010:14) follows Van Valin and La Polla (1997:384) in distinguishing three subtypes of benefactive, as it has been defined here: namely, plain-beneficiaries, deputative–beneficiaries, and recipient-benefactives.

In this first subtype, the benefaction consists of amusing or entertaining the

beneficiaries, as in the following examples:

(49) *Sekka* *angklung-é* *ng-gambel-ang* *rejang-é*
player.group angklung-DEF AV-play-APPL rejang.dancers-DEF
'The angklung player group played for the rejang dancers.'

(50) *Mbok Ganing* *ng-gending-ang* *panak-a-né*
Sister [name] AV-sing-APPL baby-3-POSS
'Sister Ganing sang for her baby.'

The second subtype, deputative-beneficiaries, specifies that someone is substituting for the beneficiaries as the agent of the denoted event, as can be seen in the example below.

(51) *Beli Nengah* *ng-alap-ang* *okké* *basé*
brother [name] AV-pick-APPL 1SG betel.leaf
'Brother Nengah picked up some betel leaves for me.'

In the benefactive-recipient subtype, the beneficiaries are also the recipients of the action or event conducted by the agent. The bound root *enjuh* 'hand' in the following example can take the suffixes *-ang* and *–in,* both encoding a beneficiary argument, but with different grammatical structures. In (52), the predicate *ngenjuhin* assigns three core arguments <*Pan Belog, dagang bébéké, ringgit a kéténg* >, while in (53), the predicate *ngenjuhang* assigns two core arguments and one oblique: <*Pan Belog, ringgit a kéténg*>< *kén dagang bébéké* >.

(52) *Pan Belog* *ng-enjuh-in* *dagang* *bébék-ké* *ringgit* *a* *kéténg*
[name] AV-hand-APPL seller duck-DEF ringgit one piece
'Pan Belog handed the duck seller one ringgit.'

(53) *Pan Belog* *ng-enjuh-ang* *ringgit* *a kéténg* *kén* *dagang* *bébék-ké*
[name] AV-hand-APPL ringgit one piece PREP seller duck-DEF
'Pan Belog handed one ringgit to the duck seller.'

After these suffixes are attached to the verb, the derived form can be classed as a verb of giving, but with a different basic characteristic to that of the 'pure' verb of giving, *behhang* 'give', which has only one argument structure, <giver, recipient, gift>. The argument *dagang bébékké* in both constructions has the semantics of a benefactive-recipient: in (52) it has the function of first object, while in (53), it in the oblique position.

6 Polysemy of benefactive markers

Kittilä (2010:18) shows that benefactive arguments share common features with other arguments, such as recipients, experiencers, possessors, causees, indirect causes, and goals. This subsection will discuss relevant data from SDB, which support this observation.

6.1 Recipient- Beneficiaries

The examples (52) and (53) above, renumbered below as (54) and (55), respectively, are the type of benefactive in which the beneficiaries can be understood as being polysemous

with recipients, since it can be clearly understood that the beneficiaries directly receive the ringgit from Pan Belog, the agent who brings about the handing over.

(54) *Pan Belog ng-enjuh-in dagang bébék-ké ringgit a kéténg*
[name] AV-hand-APPL seller duck-DEF ringgit one piece
'Pan Belog handed the duck seller one ringgit.'

(55) *Pan Belog ng-enjuh-ang ringgit a kéténg kén dagang bébék-ké*
[name] AV-hand-APPL ringgit one piece PREP seller duck-DEF
'Pan Belog handed one piece of ringgit to the duck seller.'

Other verbs that can be grouped into this class are *behhang* 'give', *bayah* 'pay', *ngejot* 'send food to someone'.

6.2 Experiencer Beneficiaries

The other bound root form to which the suffixes *-ang* and *-in* can attach is *édéng* 'show'. The derived form *ngédéngin* assigns three core arguments, i.e. *Pred' <x, y, z>,* whereas *ngédéngang* requires two core arguments and an oblique argument, i.e. *Pred' <x, y>< z>*.

(56) *Man Manis ng-édéng-in mémé-na-né kelambi anyar*
[name] AV-show-APPL mother-3-POSS kabaya new
'Man Manis showed her mother a new kabaya.'

(57) *Man Manis ng-édéng-ang kelambi anyar kén mémé-na-né*
[name] AV-show-APPL kabaya new PREP mother-3-POSS
'Man Manis showed a new kabaya to her mother.'

The verbs *ngédéngin* and *ngédéngang* that fill the predicates of the sentence constructions above can be classed as verbs of attention. In both sentence constructions, the argument *méménanné* expresses a benefactive, of the semantic type experiencer-beneficiaries.

Similarly, the verbs *nyaduin* <x, y, z>, as in (58), and *nyaduang* <x, y><z>, as in (59), both derived from the base form *sadu* 'inform', can be classed as beneficiaries-experiencer verbs.[6] The two derived forms assign different argument structures, the first with three core arguments, the second with two core arguments and one oblique argument. There are some other lexical verbs belonging to this group, like *orahang* 'say', *nuturang* 'tell', and *ndingehin* 'inform indirectly'.

(58) *Iyya ny-(s)adu-in mémé-na-né tingkah panak–a-né*
3SG AV-report-APPL mother-3-POSS behaviour child-3-POSS
'She informed her mother about the behaviour of her child.'

(59) *Iyya ny-(s)adu-ang tingkah panak-a-né kén mémé-na-né*
3SG AV-inform-APPL behaviour child-3-POSS PREP mother-3-POSS
'She informed her mother about the behaviour of her child.'

---

[6] *Nyaduang* 'to report', *ngorahang 'to* say', *nuturang* 'to tell', and *ndingehin* 'to inform indirectly' are grouped into verbs of saying. The object or oblique argument denotes someone who receives information from the person speaking.

The beneficiary in each of these two sentences is the NP *méménané,* which has different argument positions across the two constructions: the former is in first object position, while the latter is in the PP oblique position.

6.3 Deputative Beneficiaries

The verb root *adep* may have the suffixes -*ang* and –*in* attached to it. When the suffix -*ang* is attached, the verb assigns three core arguments, i.e. *adepang* <Agent, Ben, Theme>, and the construction specifies beneficiaries of the *deputative–beneficiaries* type. In the case where the verb is suffixed with -*in*, the number of core arguments assigned is the same, *adepin* <Agent, Ben, Theme>, but the construction specifies beneficiaries of the *benefactive-recipient* type. In this case, the beneficiaries is the buyer (see Van Valin and La Polla 1997:383) and (Kittila 2010:14). Below are the two examples of the suffixed verbs.

(60) *Wayan Ragi ng-adep-ang kaka jehhuk-ké nto*
[name] AV-sell-APPL 1SG oranges-DEF DEM
'Wayan Ragi sold the oranges for me.'

(61) *Nyoman Kanti ng-umbah-ang kaka kelambi*
[name] AV-wash-APPL 1SG kabaya
'Nyoman Kanti washed the kabaya for me.'

Another verb root that has the same characteristics as *adep* 'sell' in this respect is e*mak* 'take', which can also be suffixed with –*ang*, to form *ngemakang.*

(62) *Mbok Balik ng-emak gajih mémén-né*
sister [name] AV-take salary mother-POSS
'Sister Balik took mother's salary.'

(63) *Mbok Balik ng-emak-ang mémé gajih*
sister [name] AV-take-APPL mother salary
'Sister Balik took mother her salary.'

The beneficiaries *kaka* in (60) and (61) and *mémé* in (63) receipt some benefit from the action carried out by the agents Wayan Rani, Nyoman Kanti, and Mbok Balik , and each of the beneficiaries is also the owner of the second NP object, in each construction. So in one sense, the first objects are beneficiaries, and in another sense, they are real possessor since those structures derive from object rising as shown in (62) and (63). The underlying structures of (60) is *Wayan Ragi ngaddep jehhuk kakané nto* 'Wayan Ragi sold my oranges' *and (61) Nyoman Kanti ngumbah kelambin kakané* ' Nyoman Kanti washed my kabaya.' In example (64), the first object benefit something from the action carried out by the creator and he can be classified into intended possessor, because the creation created by the creator is intended to the first object, in this case the beneficiary

(64) *Putu ng-gaé-nang adi-na-né layangan*
[name] AV-make-APPL sibling-3-POSS kite
'Putu made a kite for his younger sibling.'

6.4 Purposive Beneficiaries

Both of the bound roots *ampak* 'to open' and *tuduh* 'to point' can have the verbal suffix *-in* attached to them in constructions (65) and (66), to form three-place verbs in SDB. The verbal suffix *-ang* can also be applied to both categories to form a two-place predicate, rather than a three-place one.

(65) *Dadong* *ng-ampak-in* *nanang* *jalanan*
grandmother AV-open-APPL father door
'Grandmother opened the door for father.'

(66) *Anak-ké nto* *n-(t)uduh-in* *okké* *jalan-né* *maku*
person-DEF DEM AV-point-APPL 1SG way-DEF to.there
'That man pointed out the correct direction to me.'

The NP arguments *nanang* in (64) and *okké* (65) occupy the first object function, and both benefit from the action conducted by the agents. The benefit which the first object derives is not concrete, but is a kind of favour; hence, the construction can be said to have a purposive meaning. In (65), the agent opens the door in order for the beneficiaries to enter the room, and in (66), the agent points in the correct direction so that the beneficiaries will not get lost.

6.5 Goal Beneficiaries

The applicative predicates in the following examples derive from the mono-transitive verbs *abba* 'bring', and e*mak* 'take', respectively. The second object in both examples, the theme*,* moves through space towards the goal, in this case the beneficiaries of the actions conducted by the agent. All the applicative predicates here can be paraphrased into SVCs in which the main verbs take the first position in the SVC (in other words, they are major verbs) and the second verb, *behhang,* is a minor verb.

(67) *Dadong* *ng-abba-ang* *okké* *kamben*
grandmother AV-bring-APPL 1SG sarong
'My grandmother brought me a sarong.'

(68) *Kaka* *nge-mak-ang* *nanang* *kaka-né* *mamahan*
1SG AV-take-APPL father 1SG-POSS food
'I took my father some food.'

7 Conclusion

After discussing all the occurrences of benefactives in SDB, some conclusions can be drawn, as follows:

1. There are two productive verbal suffixes found in SDB, *-ang* and *-in*, that have an applicative function, in addition to causative and verbalizing functions.
2. Semantically determined variation in the coding of beneficiaries can be related to the polysemy of benefactive markers. The suffixes *–ang* and *–in*, in combination with the semantics of the root, give rise to particular benefactive readings, such as recipient-beneficiariess, deputative-benefactives, goalbeneficiaries, experience-

beneficiaries, and purposive -beneficiaries, etc. Some roots, when *–ang* and *-in* are attached, both express benefactive meaning, and they have reversed ordering of the arguments, e.g., *okke ngidih pipis sig nanang <Pred (Ben-rec, Theme, Source)>* 'I begged father for some money') and *nanang ngkidihang pipis ken okke <Pred (Agent, theme,Ben-goal)>* 'Father gave me money as a gift'.

3. Examining the benefactive types in SBD shows that the dialect exhibits each of the different strategies for expressing benefactive meaning indentified in earlier studies, including the use of adpositions, serial verb constructions, and applicativisation.

References

Alsina, Alex, Joan Bresnan, & Peter Sells. 1997. Complex Predicates: Structure and Theory. In Alex Alsina, Joan Bresnan, and Peter Sells (Eds.), *Complex Predicates*. Stanford: CSLI.

Baker, Mark C. 1993. Complex Predicates and Agreement in Polysynthetic Languages. In Alex Alsina, Joan Bresnan, and Peter Sells (Eds.), *Complex Predicates*. Stanford: CSLI.

Butt, Miriam. 1993. Complex Predicates in Urdu. In Alex Alsina, Joan Bresnan, and Peter Sells (Eds.), *Complex Predicates*. Stanford: CSLI.

Censabella, Marisa. 2010. Beneficiaries and recipients in Toba. In Fernando Zúñiga and Seppo Kittilä (Eds*.), Benefactives and Malefactives: Typological perspectives and case studies.* Philadelphia: John Benjamin's publishing Company.

Comrie, Bernard. 1981. *Language Universals and Linguistic Typology*. Oxford: Blackwell.

Creissels, Denis. 2010. Benefactive applicative periphrases: A typological approach. In Fernando Zúñiga and Seppo Kittilä (Eds.), *Benefactives and Malefactives: Typological perspectives and case studies.* Philadelphia: John Benjamin's publishing Company.

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

Durie, Mark. 1997. Grammatical Structures in Verb Serialization. In Alex Alsina, Joan Bresnan, and Peter Sells (Eds.), *Complex Predicates*. Stanford: CSLI.

Foley, William A. & Robert D. Van Valin, Jr. 1984. *Functional Syntax and Universal Grammar*. Cambridge: Cambridge University Press.

Foley, William A. & Robert D. Van Valin, Jr. 1992. A typology of information packaging in the clause. In Timothy Shopen (Ed.), *Language Typology and Syntactic Description*, *Vol I: Clause Structure*, 362-336. Cambridge: Cambridge University Press.

Foley, William A. 1993. Polysynthesis and Complex Verb Formation: The Case of Applicative in Yimas. In Alex Alsina, Joan Bresnan, and Peter Sells (Eds.), *Complex Predicates*. Stanford: CSLI.

Hale, Ken and Jay Keyser. 1997. On the Complex Nature of Simple Predicators. In Alex Alsina, Joan Bresnan, and Peter Sells (Eds.), *Complex Predicates*. Stanford: CSLI.

Kaoru, Kiyosawa & Donna B. Gerdts. 2010. Benefactive and malefactive uses of Salish applicative. In Fernando Zúñiga and Seppo Kittilä (Eds.), *Benefactives and Malefactives: Typological perspectives and case studies.* Philadelphia: John

Benjamin's publishing Company.

Katamba, Francis. 1993. *Morphology*. London: Macmillan Press Ltd.

Kittila, Seppo & Fernando Zuniga. 2010. Benefaction and malefaction from a cross-linguistic perspective. In Fernando Zúñiga and Seppo Kittilä (Eds.), *Benefactives and Malefactives: Typological perspectives and case studies.* Philadelphia: John Benjamin's publishing Company.

Quirk, Randolph, et al. 1972. *A Grammar of Contemporary English*. London: Longman.

Radetzky, Paula & Tomoko Smith. 2010. An areal and cross-linguistic study of benefactive and malefactive constructions. In Fernando Zúñiga and Seppo Kittilä (Eds.), *Benefactives and Malefactives: Typological perspectives and case studies.* Philadelphia: John Benjamin's Publishing Company.

Radford, Andrew. 1997. *Syntactic theory and the structure of English: a minimalist approach*. Cambridge: Cambridge University Press.

Rosen, Carol. 1993. Auxiliation and Serialization: On Discerning the Difference. In Alex Alsina, Joan Bresnan, and Peter Sells (Eds.), *Complex Predicates*. Stanford: CSLI.

Schmidtke-Bode, Karsten. 2010. The role of benefactives and related notions in the typology of purpose clauses. In Fernando Zúñiga and Seppo Kittilä (Eds.), *Benefactives and Malefactives: Typological perspectives and case studies.* Philadelphia: John Benjamin's Publishing Company.

Shibatani, Masyoshi & Sandra A. Thompson. 1996. *Grammatical Constructions: Their Form and Meaning.* Oxford: Clarendon Press.

Sedeng, Nyoman. 2010. *Morfosintasis Bahasa Bali Dialek Sembiran*. Denpasar. Udayana Press.

Song, Jae Jung. 2010. Korean benefactive particles and their meanings. In Fernando Zúñiga and Seppo Kittle (Eds.), *Benefactives and Malefactives: Typological perspectives and case studies.* Philadelphia: John Benjamins Publishing Company.

Van Leynseele, H. 1975. Restrictions on serial verbs in Anyi. *Journal of West African Languages*, *10*(2), 189-218.

Van Valin, Jr. Robert D. 1990. *A Synopsis of Role and Reference Grammar*. Philadelphia: John Benjamins.

Van Valin, Jr. Robert D., & Randy J. LaPolla. 1997. *Syntax: Structure, Meaning, and Function*. Cambridge: Cambridge University Press.

Yamashita, Tomoko Smith. 2010. Cross-linguistic categorization of benefactives by event structure: A preliminary framework for benefactive typology. In Fernando Zúñiga and Seppo Kittilä (Eds.), *Benefactives and Malefactives: Typological perspectives and case studies.* Philadelphia: John Benjamins Publishing Company.

Zuniga, Fernando. 2010. Benefactive and malefactive applicativization in Mapundungun. In Fernando Zúñiga and Seppo Kittilä (Eds.), *Benefactives and Malefactives: Typological perspectives and case studies.* Philadelphia: John Benjamins Publishing Company.

15 *Passive actors are not adjuncts – Consequences for the distinction between symmetrical and asymmetrical voice alternations*

SONJA RIESBERG

1 Introduction

One of the most common views on passive constructions found in the literature is to analyse the adpositionally marked agent argument as an adjunct. Under such an approach, the passive agent is prevented from being linked to a syntactic function. The details of this process differ from theory to theory, but in most cases the passive morphology is assumed to be responsible for this "suppression" of the agent argument. This argument suppression is also one of the most salient differences when comparing the active-passive alternation with so-called SYMMETRICAL VOICE alternations (Himmelmann 2005; Foley 2008) where a change in voice takes place without any suppression of arguments, and where the non-subject remains a direct core argument in all voice constructions.

There are, however, also other views that either acknowledge some kind of special status of this passive adjunct or dispense with the adjunct analysis altogether. Grimshaw, e.g., introduces the term ARGUMENT-ADJUNCT, claiming that the passive agent shows properties of both arguments (due to its "relationship to an argument structure") and adjuncts (due to it "not satisfy(ing) argument structure positions") (1990: 107). Likewise, in Role and Reference Grammar (RRG) the status of the passive agent is not clear: as it has a slot in the logical structure of the verb, it should be analysed as a core argument of the verb; however, being syntactically realised in the periphery, it shows adjunct properties (cf. Van Valin 2005). Dowty (2003), on the other hand, comes to the conclusion that there are no reliable criteria for distinguishing arguments from adjuncts in the first place and that passive agents should be analysed as arguments.

In this paper, I want to argue that argumenthood and adjuncthood might better be conceived of as gradient concepts, with the passive agent being closer to the argument side than to the adjunct side. The first piece of evidence for this claim will come from German data, based on work by Zifonun, Hoffmann, Strecker, et al. (1997), who developed a series of tests to distinguish arguments from adjuncts. Applying these tests results in a very fine-grained distinction, showing strong evidence that the passive agent in German is clearly more argument-like than adjunct-like. The second piece of evidence will come from two Austronesian languages – Totoli and Balinese – that display symmetrical voice alternations. Just like in asymmetrical active-passive alternations, Balinese arguments

Asia-Pacific Linguistics, 2014.

differ in their syntactic behaviour (Arka 2009, 2013), giving rise to the claim that they gradually vary in their argument status.

These observations have two consequences: First, it is no longer necessary to analyse the passive agent as some kind of hybrid category (e.g. either as an argument-adjunct, or an argument realised in the periphery). Instead, the difference between the non-subject in an active construction and the non-subject in a passive construction consists in the former being a direct core argument and the latter being an oblique core argument. Second, and resulting from the first point, the difference between asymmetrical and symmetrical voice systems is then not as fundamental as it is widely assumed. Both alternation types can be analysed without argument suppression. The difference between the two systems would then only be that in symmetrical voices non-subject agents are direct core arguments while in asymmetrical voices they are oblique core arguments.

2 Direct Cores, oblique Cores, adjuncts, and the argument-adjunct distinction

As in most fields of linguistics, there are some terminological inconsistencies when it comes to the concept of argumenthood. For the following discussion, the two conceptual pairs CORE VS. NON-CORE and ARGUMENT VS. ADJUNCT are of special interest here. While in some theoretical frameworks (e.g. RRG) these pairs denote more or less the same thing, some authors (most notably those working within LEXICAL FUNCTIONAL GRAMMAR (LFG)) make a distinction between the concept of core/non-core and argumenthood, even though their meaning partly overlaps.

Arguments are traditionally assumed to be those elements in a clause that are licensed by the verb, i.e. that are part of the verb's argument structure. Adjuncts, on the other hand, are not required by the verb but can rather be added to or omitted from the clause freely. This difference – the obligatoriness of arguments vs. the optionality of adjuncts – has often been taken as one of the defining criteria of these two concepts. However, it is well known that the fact that an element can be omitted from a clause does not mean that this phrase necessarily has to be an adjunct. Even though there are of course certain elements that must not be omitted from a clause, this is certainly not true for all arguments. Omissibility is thus a necessary but not a sufficient criterion for adjuncthood: If a phrase cannot be omitted, it is an argument, but the reverse does not hold, as not all phrases that are omissible can automatically be analysed as adjuncts.

A second criterion for determining whether a phrase has argument status or whether it is an adjunct is government. The concept of government describes a dependency relation between two elements within a sentence, in which the governing element determines the form of the governed element with respect to a given category. If the form of a given phrase is governed by the verb, it is an argument; if not, it is an adjunct. In languages that exhibit case marking, for example, the verb determines (i.e. governs) the case of its object arguments, as illustrated in the German example below, where *unterrichten* 'to teach' governs the accusative, while *helfen* 'to help' governs the dative.

(1)	*Der*	*Lehrer*	*unterrichtet*	***die***	***Kinder***.
	DET.3s.NOM.M	teacher	teach:3s.PRS	DET:3p.**ACC**	child:p.**ACC**

'The teacher teaches the children.'

(2) *Der* DET.3s.NOM.M *Lehrer* teacher *hilft* help:3s.PRS ***den*** DET.3p.**DAT** ***Kindern***. child:p.**DAT**

'The teacher helps the children.'

The distinction between arguments and adjuncts is a substantial one, made in traditional grammar as well as in probably all modern syntactic theories. Even though in praxis the assignment of a noun or a phrase to either of these two categories might not always be unproblematic, most authors agree on making this distinction, and there is broad consensus – at least theoretically – about which properties are attributed to arguments and which ones are attributes to adjuncts. This kind of consensus does not exist when it comes to concepts such as (NON-) CORE and OBLIQUE, and these terms have, in fact, been used quite differently within different traditions or different frameworks. Within LFG, for example, a first broad distinction is made between arguments and non-arguments. The latter include adjuncts, the former are further divided into core and non-core. Core arguments subsume subjects and what has been traditionally termed direct and indirect objects (OBJ and OBJ_θ in LFG terminology), and non-core arguments subsume a grammatical relation called OBLIQUE (OBL_θ), which is an argument marked by an adposition. In LFG, an oblique argument is thus understood to be an adpositionally marked, non-core argument. This kind of use differs slightly from the way the term is used e.g. in RRG. Like LFG, RRG makes a distinction between arguments and non-arguments, with adjuncts falling under the second category. The term CORE (-ARGUMENT) is however used for all "arguments in semantic representation of the predicate" (Van Valin/LaPolla 1997: 27), i.e. all semantic arguments of the verb, including those that are marked by an adposition. Unlike in LFG then, oblique arguments in RRG are considered to be core. Table 2.1 and Table 2.2 how the different concepts are categorised in the two frameworks.

Table 2.1: Core-Non-core distinction RRG

ARGUMENTS		NON-ARGUMENTS
Core		**Non-Core**
direct core arguments	oblique core arguments	adjuncts

Table 2.2: Core-Non-core distinction LFG

ARGUMENTS		NON-ARGUMENTS
Core	**Non-Core**	
direct arguments	oblique arguments	adjuncts

In this paper, I will adopt the terminology as used within RRG, making a distinction between DIRECT CORE ARGUMENTS, i.e. arguments of the verb that are not marked by an adposition, OBLIQUE CORE ARGUMENTS, i.e. arguments of the verb that are adpositionally marked, and ADJUNCTS, i.e. phrases that are not arguments of the verb.

Now even though there is a rather broad consensus about the usefulness of the argument-adjunct distinction, this distinction is by no means a trivial one, and quite some literature is devoted to the question of how to find reliable criteria to distinguish arguments from adjuncts. Omissibility was already said to be only of little use, as it constitutes only a necessary but not a sufficient criterion for adjuncthood. Likewise, government is not a reliable criterion to determine whether a phrase is an argument or not. This is especially the case with prepositional phrases.

Prepositional phrases differ in whether their head is governed by the verb or not. Those prepositions that are not governed by the verb have been termed PREDICATIVE PREPOSITIONS while those that are governed have been called NON-PREDICATIVE (cf. Bresnan 1982). As the following two examples illustrate, the form of the preposition in a non-predicative PP is fixed, just like the case of a governed object phrase, while it is freely

interchangeable in predicative PPs. The former can therefore be claimed to be an argument of the verb; the latter should be regarded to be an adjunct.

(3) *Hannah lent a pair of gloves* ***to****/*in/*on/*under Katie.*

(4) *Hannah is reading a book* ***on/under/next to*** *the sofa.*

However, there are cases where this correspondence between government of the preposition and argument status of the PP does not hold. Consider the following example containing the English verb *put*, which takes an agent argument, a theme argument, and a locative argument. The non-omissibility of the locative phrase shows that it clearly bears argument status. Nevertheless its preposition can be chosen freely:

(5) *Hannah puts her toothbrush* ***on/under/next to*** *the shelf.*

Government of the preposition is thus not a reliable test to determine the argument status of a prepositional phrase in English. In a language like German, however, government *into* the PP is possible: even though the preposition itself can be chosen freely the case of the contained noun phrase is governed by the verb.

(6)	*Hannah*	*stellt*	*ihre*	*Zahnbürste*	***auf/unter/neben***
	PN	put:3s.PRS	3s.POSS.ACC.F	toothbrush	on/under/next to
	den/	****der/***	****dem***	*Schrank.*	
	DET.3s.ACC.N	DET.3s.NOM.N	DET.3s.DAT.N	cupboard	

'Hannah puts her toothbrush on/under/next to the cupboard.'

But on closer inspection, even government into a PP becomes problematic. As Zifonun, Hoffmann, and Strecker et al. (1997) point out, a positional verb like *hängen* 'to hang' can either take a locative argument that bears dative case or a directional argument that bears accusative case. As the following two examples show, neither the preposition nor the case of the noun inside the PP is fixed:

(7)	*Der*	*Lappen*	*hängt*	***auf/hinter/an***
	DET.3s.NOM.M	cloth	hang:3s.PRS	on/behind/at
	der	*Heizung.*		
	DET.3s.DAT.F	radiator:DAT		

'The cloth hangs on/behind/at the radiator.'

(8)	*Der*	*Lappen*	*hängt*	***auf/hinter/an***
	DET.3s.NOM.M	cloth	hang:3s.PRS	on/behind/against
	die	*Heizung.*		
	DET.3s.ACC.F	radiator:ACC		

'The cloth hangs on/behind/against the radiator.'

Thus neither government of the preposition nor government of the noun inside the PP can be used as clear evidence for the argument status or the adjunct status of PPs. Many authors have discussed the difficulties that arise when trying to distinguish arguments from

adjuncts, some of them having come to quite some pessimistic conclusions. Thus, for example Dowty (2003: 2):

> *(...) no diagnostic criteria have emerged that will reliably distinguish adjuncts from complements in all cases – too many examples seem to fall into the crack between the two categories, no matter how theorists wrestle with them.*

Other authors have yet tried to find workable procedures to nevertheless come to a satisfying result. Many of them point out that a larger number of (language specific) diagnostics should be applied, and that one should dispense with the idea that argumenthood and adjuncthood are clear cut categories. Jacobs (1994, chapter 4), for example, introduces seven syntagmatic relations that might or might not hold between a verb and a given phrase. Most of these relations have already been applied in former approaches, but Jacobs shows that they can in fact occur independently from each other. A given phrase thus does not have to show all these characteristics in order to be classified as an argument. Rather, some phrases may stand in many of these relations to the predicate, other may do so only in a few or in none. Argumenthood and adjuncthood therefore become gradient concepts (i.e., a phrase can be more or less argument properties thus being more or less argument-like) (Jacobs 1994: 57). This view is strongly supported in this paper, and the discussion of data from an asymmetrical voice language (German) as well as data from the symmetrical voice languages Balinese and Totoli will give further evidence for this claim.

3 The status of the passive-actor - a multi-step approach to the argument-adjunct distinction[1]

One approach that builds strongly on Jacobs' ideas introduced above has been developed by Zifonun, Hoffmann, Strecker et al. 1997 (henceforth ZHS), who make use of three different kinds of tests in order to determine the argument status of phrases in German. The tests are applied consecutively: Depending on the result of the first test, a phrase will either be classified as an argument, or further examination is needed and the phrase will be passed on the second test. Depending on the result achieved in this second test, the phrase will either be directly classified as an adjunct, or it will again be passed on to the next test. The last test then decides whether a phrase that did not yield clear results in the first two tests should be analysed as an argument or as an adjunct. One interesting feature of ZHS's approach which distinguishes it from other, similar approaches is that phrases can not only fail or pass a test but they can also be valued as passing the test only under certain conditions. This kind of multi-step approach has thus two advantages over the application of single criteria such as obligatoriness or government: First, it allows phrases not only to be classified as clear adjuncts or clear arguments, but also as being an intermediate argument candidate. Second, phrases showing adjunct properties in one test might well behave like arguments in another test. This leads to a finer-grained classification than the one achieved by the traditional approaches, and it follows naturally that argumenthood and adjuncthood become gradient concepts: Depending on the results within the different tests, phrases turn out to be prototypical arguments, a prototypical adjuncts, or something in between these two categories.

---

[1] This section closely follows section 4.3.2 in Riesberg (2014).

The first test, called reduction test ('Reduktionstest' in ZHS: 1043ff.), separates obligatory arguments from facultative arguments and adjuncts. As a starting point, it takes a complete clause that forms a proposition. This clause is then reduced by the respective phrase that is supposed to be tested for its argument status. If the reduced clause does not form a proposition anymore, or if the propositional content has changed, the omitted phrase is assumed to be an argument. This is exemplified below, where the reduced clause in (9b) does not have the same propositional content as (9a). In fact, it is not even grammatical in the first place (both examples are taken from ZHS: 1044). The prepositional phrase *auf mich* 'on me' is therefore classified as an (obligatory) argument. However, as already discussed in the previous section, it does not automatically follow that those phrases which can be left out without changes in meaning or grammaticality are automatically adjuncts. Therefore, phrases like *ein Bier* 'a beer' in (10), passive actor phrases like *von dem Bauern* 'by the farmer' in (11), and locative phrases like *im Garten* 'in the garden' in (12) need further testing, even though they can be left out easily and without any loss of grammaticality.

(9) a. *Hans* *verlässt* *sich* ***auf*** ***mich***.
PN.NOM rely:3S.PRS REFL on 1S.ACC
'Hans relies on me.'

b. **Hans* *verlässt* *sich.*
PN.NOM rely:3S.PRS REFL

(10) a. *Hans* *trinkt* ***ein*** ***Bier***.
PN.NOM drink:S.PRS DET.S.N.ACC beer:S.N.ACC
'Hans is drinking a beer.'

b. *Hans* *trinkt.*
PN.NOM drink:3S.PRS
'Hans is drinking.'

(11) a. *Die* *Kuh* *wird* ***von*** ***dem***
DET.S.F.NOM cow:S.F.NOM AUX.3S.PRS by DET.S.M.DAT
Bauern *gemolken.*
farmer:S.M.DAT milk:PART
'The cow is being milked by the farmer.'

b. *Die* *Kuh* *wird* *gemolken.*
DET.S.F.NOM cow:3S.F.NOM AUX.3S.PRS milk:PART
'The cow is being milked.'

(12) a. *Sarah* *grübelt* ***im*** ***Garten***.
PN.NOM muse:3S.PRS in:DET.S.M.DAT garden:S.M.DAT
'Sarah is musing in the garden.'

a. *Sarah* *grübelt.*
PN.NOM muse:3s.PRS

'Sarah is musing.'

Second, following the reduction test, the so-called inference test ('Folgerungstest' ZHS: 1046ff.) is applied. It tests the status of those phrases that turned out to be freely omissible in the reduction test. Drawing on the involvement of the participants in the event (ZHS: 1046), it investigates the sematic relationship between the phrase and the verb. For this test, the reduced clause from the reduction test is taken as input and the formerly omitted phrase is replaced by its indefinite form. If it is not possible to infer this indefinite use of the reduced phrase from the reduced clause, then the omitted phrase is analysed as an adjunct. If such an inference is possible, the omitted phrase might be an argument. In the first one of the following two example sets it seems to be unproblematic to infer the meaning of the b.-sentence from the a.-sentence, in the way it is paraphrased under c. The two phrases that have been determined as possible argument candidates by the reduction test in (10) and (11) – *ein Bier* 'a beer' and *von dem Bauern* 'by the farmer' – can therefore not be directly classified as adjuncts. Again, further testing is needed. As example (15) shows, their behaviour contrasts with the behaviour of the locative phrase: an inference from sentence a. to sentence b. is not possible in the same way as it has been possible for the two aforementioned example sets. It is certainly true that the act of musing has to take place somewhere, just as it will take place at some time and probably also for a certain reason. However, all events in this world take place in space and time without this information being explicitly included in the verb's semantics denoting this event. The locative phrase *im Garten* 'in the garden' is therefore classified as an adjunct.

(13) a. *Hans* *trinkt.*
PN.NOM drink:3s.PRS

'Hans is drinking.'

b. *Hans* *trinkt* *etwas.*
PN.NOM drink:3s.PRS something:ACC

'Hans is drinking something.'

c. *There is something so that Hans is drinking it.*

(14) a. *Die* *Kuh* *wird* *gemolken*
DET.S.F.NOM cow:S.F.NOM AUX.3s.PRS milk:PART

'The cow is being milked.'

b. *Die* *Kuh* *wird* *von*
DET.S.F.NOM cow:S.F.NOM AUX.3s.PRS by

jemandem *gemolken.*
somebody:DAT milk:PART

'The cow is being milked by somebody.'

c. *There is somebody so that the cow is being milked by her/him.*

(15) a. *Sarah* *grübelt.*
PN.NOM muse:3s.PRS

'Sarah is musing.'

b. *Sarah* *grübelt* *irgendwo.*
PN.NOM muse:3s.PRS somewhere

'Sarah is musing somewhere.'

c. ??*There is a place where Sarah is musing.*

Besides those cases where the indefinite reading clearly follows from the reduced clause and those cases where this is clearly not the case, there are often instances in which a straightforward classification is not possible. It might be context-dependent whether an additional phrase is inferable or not. For example, German verbs denoting an action can take an additional dative argument in contexts where the actor and the beneficiary do not denote the same referent. Thus in (16), the dative phrase *Eva* can be omitted and its indefinite reading in c. can be inferred from b. in the sense of the paraphrase in d. In this context the omissible dative phrase is classified as an intermediate argument candidate whose status has to be further examined.

(16) a. *Ich* *habe* ***Eva*** *ein* *Fahrrad*
1s.NOM AUX.3s.PRS PN.DAT DET.s.N.ACC bike:s.N.ACC

gekauft.
buy:PART

'I bought Eva a bike.'

b. *Ich* *habe* *ein* *Fahrrad* *gekauft.*
1s.NOM AUX.3s.PRS DET.s.N.ACC bike:s.N.ACC buy:PART

'I bought a bike.'

c. *Ich* *habe* ***jemandem*** *ein* *Fahrrad*
1s.NOM AUX.3s.PRS somebody:DAT DET.s.N.ACC bike:s.N.ACC

gekauft.
buy:PART

'I bought somebody a bike.'

d. *There is somebody so that I bought I bike for her/him.*

The last test in ZHS's approach is the so-called connection test ('Anschlusstest' in ZHS: 1051ff.). It aims to clarify the status of those phrases that did not achieve clear results in the preceding tests. Again, the reduced clause is taken as the starting point, which is then extended by the phrase *und das X* 'and this X' (X standing for the previously omitted phrase). If the phrase that is substituted for X is an adjunct, such a conjunct construction is easily possible, as exemplified in (17a). Note that the PP *im Garten* 'in the garden' has already been determined to be an adjunct with the help of the inference test above (examples (12) and (15)). The behaviour of an NP that has been clearly classified as an argument by the reduction test is illustrated in (17b).

(17) a.

Sarah	*grübelt*	*und*	*das*	*im*	*Garten.*
PN.NOM	muse:3s.PRS	and	that	in:DET.S.M.DAT	garden:S.M.DAT

'Sarah is musing and (she is doing) that in the garden.'

b.

**Hans*	*verlässt*	*sich*	*und*	*das*	*auf*	*mich.*
PN.NOM	rely:3s.PRS	REFL	and	that	on	1s.ACC

So far, the connection test thus confirms the results achieved by the first two tests. The more interesting question is how the non-straightforward cases will behave. As illustrated below, in these cases, the grammaticality seems to vary. Some phrases clearly yield an ungrammatical result in the connection test, such as the NP *ein Bier* 'a beer' in example (18a). In other cases, the grammaticality of a phrase in this construction is slightly more difficult to judge. The dative phrase that passed the inference test only in certain contexts (cf. example (16)) is only marginally acceptable in the connection construction, and the same seems to be true for the passive actor. ZHS judge both phrase types as marked, as illustrated in (18b) and (18c) respectively.

(18) a.

**Hans*	*trinkt*	*und*	*das*	*ein*	*Bier.*
PN.NOM	drink:S.PRS	and	that	DET.S.N.ACC	beer:S.N.ACC

b.

?Ich	*habe*	*ein*	*Fahrrad*	*gekauft*
1s.NOM	AUX.3s.PRS	DET.S.N.ACC	bike:S.N.ACC	buy:PART
und	*das*	*Eva*		
und	that	PN.DAT		

'I bought a bike, and (I did) that for Eva.'

c.

?Die	*Kuh*	*wird*	*gemolken*	*und*
DET.S.F.NOM	cow:S.F.NOM	AUX.3s.PRS	milk:PART	and
das	*von*	*dem*	*Bauern*	
that	by	DET.S.M.DAT	farmer:S.M.DAT	

'The cow is being milked, and this (is done) by the farmer.'

However, the dative phrase in (18b) and the passive actor in (18c) do not exhibit the same degree of argumenthood. This is due to only the latter easily passing the inference test, while the former can only be classified as an intermediate argument candidate by the inference test. This shows why argumenthood in ZHS's approach has to be a gradual concept: A phrase that is an argument candidate due to its behaviour in the inference test and which is only marginally acceptable in the connection test is more argument-like than a phrase that is equally marked in the connection construction, but is only an intermediate argument candidate with respect to the inference test.

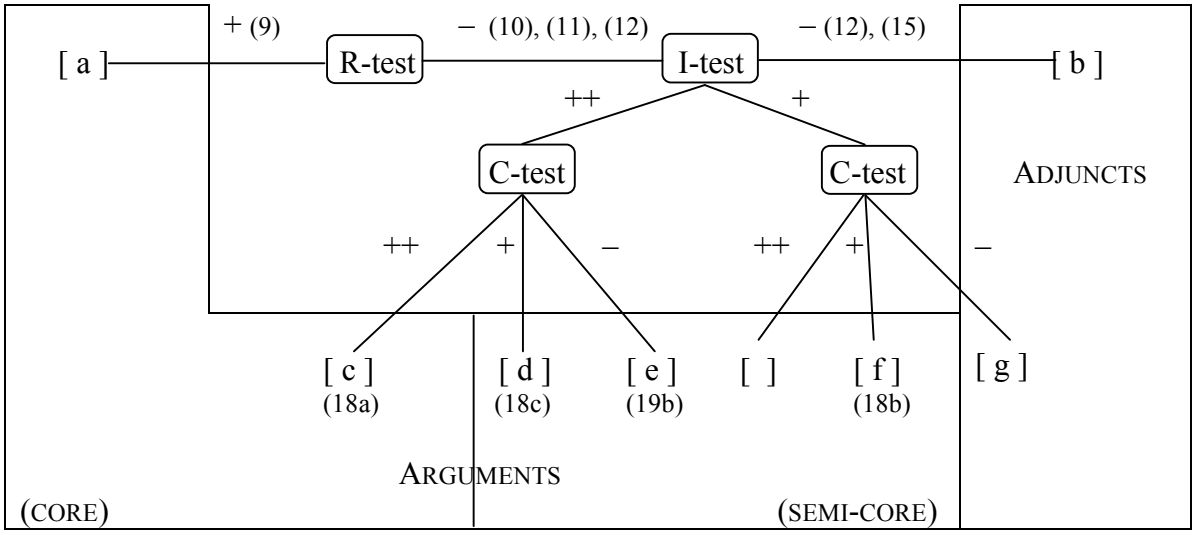


Figure 3.1: Distinction of arguments and adjuncts in German

The results of the testing procedure are summarised in Figure 3.1 (adopted from ZHS: 1059), illustrating how phrases can be arranged on a scale that ranges from prototypically argument-like to prototypically adjunct-like. The letters [a] to [g] represent the different phrase types, arranged according to their behaviour in the three above-mentioned tests. Depending how well the single candidates behave in the tests – i.e. depending on their grammaticality in the different constructions – they are marked as positive (++), intermediate positive (+), or as negative (–). Most of these possibilities have been discussed above. The most uncontroversial cases, such as the respective phrase in (9), pass the reduction test. They are subsumed under phrase class [a] in Figure 3.1. Phrase class [b] is also easy to determine: This class includes those cases that fail in the inference test, as illustrated in (12) and (15). The two phrase classes [a] and [b] form the two endpoints on the argument-adjunct continuum, with [a]-phrases being prototypical arguments and [b]-phrases being prototypical adjuncts. The phrases that end up in the classes [c], [d], [e], [f], and [g] constitute the larger group of intermediate candidates. The first three of these classes contain those phrases that positively pass the inference test. Phrase class [c] consists of those phrases that furthermore positively passed the connection test (cf.(18a)). Members of class [d] on the other hand are judged as marginally acceptable in the connection construction. This class also includes the prepositional actor phrase of the passive construction, as shown in (18c). Phrases that do not pass the connection test, while at the same time passing the inference test, are subsumed under [e]. This phrase type has not been discussed, so far. ZHS cite the actor phrase of the German *durch*-passive as an example. The German passive can, under certain conditions, be formed in two different ways, the actor phrase being either introduced by the preposition *von*, or by the preposition *durch*. According to ZHS, these two PPs behave differently with respect to the connection test. The examples in(19) show that, unlike in (18c), the *durch*-PP can be used easily in the connection construction. The two passive actor phrases thus belong into two different phrase classes, namely [d] and [e] respectively.

(19) a. *Die* *Stadt* *wurde* *durch* *ein*
DET.S.F.NOM town:NOM AUX.PST through DET.S.N.ACC
Erdbeben *zerstört.*
earthquake:ACC destroy:PART

'The city was destroyed by an earthquake.'

b.	*Die*	*Stadt*	*wurde*	*zerstört*	*und*
	DET.S.F.NOM	town:NOM	AUX.PST	destroy:PART	and
	das	*durch*	*ein*	*Erdbeben.*	
	that	through	DET.S.N.ACC	earthquake:ACC	

'The city was destroyed and this (happened) through an earthquake.'

The phrases within the last two classes, [f] and [g], only achieve an intermediate positive result in the inference test, but they differ with respect to their behaviour in the connection test. While members of phrase class [f] perform only in parts positively (see e.g. the dative phrase in example (18b)), members of phrase class [g] yield negative results.[2]

ZHS (p. 1059) classify phrase classes [a], [c], [d], and [f] as arguments and analyse the members of [g] and [b] as adjuncts. They furthermore distinguish arguments in the core (including the classes [a] and [c]) from those in the periphery ([d], [e], [f]). As will be shown below, their classification is very similar to results achieved by Arka (2009 and 2013), who investigates the core status and the argument status of phrases in Austronesian languages. To avoid confusion with the RRG use of the terms CORE and PERIPHERY as described above, I will adopt Arka's (2009) terminology and call the two kinds of argument types CORE and SEMI-CORE (cf. Figure 3.1).

Of course, ZHS's approach is not unproblematic. The fine-grained distinctions within the semi-core often depend on very subtle changes in grammaticality which are sometimes difficult to judge. For example, the difference between the *von*-phrase and the *durch*-phrase in (11a) and (19a) seems to be rather small, and the difference in grammaticality between examples (18c) and (19b) might not be directly obvious to many native speakers. It may be conceivable to group these two phrases together in phrase class [e], with both phrases performing equally negatively in the connection test. But even if one does not agree on every detail of their classification, two important conclusions can be drawn from the results of ZHS's multi-step testing approach: First, the syntactic status of (nominal) phrases reaches from prototypical argumenthood to prototypical adjuncthood. Between these two end points lies a continuum on which a number of intermediate candidates can be arranged according to their behaviour in the different test constructions. Depending on their position on the continuum, these phrases are more or less argument-like, or more or less adjunct-like than others. Second, following from ZHS's testing procedure, the actor phrase of the German passive is located within the semi-core, and in fact this is true for both the *von*-phrase as well as the *durch*-phrase. In any case, both phrases are closer to the argument side of the continuum than to the adjunct side. This is due to them both passing the inference test without any difficulties, which in turn reflects the fact that the referent denoted by the actor phrase in a passive construction is a participant in the event denoted by the verb.

4 Gradience of argumenthood in western Austronesian languages

The peculiarities of their voice system have long been noted as one of the most distinct features of western Austronesian languages. Being morphologically marked in all voices (which often include more than one undergoer voice in addition to one actor voice) and allowing for voice alternations without argument demotion, these languages have been

[2] Note that the last possible combination (a restricted result in the inference test and a positive result in the connection test) does not seem to be attested as indicated by the empty square brackets in Figure 3.1.

termed SYMMETRICAL VOICE LANGUAGES (cf. Himmelmann 2005, Foley 2008). Riesberg (2011: 11) gives the following heuristic description of symmetrical voice:

> A language displays a symmetrical voice system, if
> a. it has more than one basic transitive construction,
> b. the corresponding arguments behave equally in all different voices, and
> c. the verb is morphologically equally marked in all different voices.

With respect to these properties, symmetrical voice languages differ crucially from languages that display the well-known active-passive alternation (henceforth ASYMMETRICAL (VOICE) LANGUAGES): First, while traditionally the passive agent in asymmetrical languages is assumed to be an adjunct and the passive construction are thus analysed as intransitive, the agent of an undergoer voice clause in a symmetrical voice language remains a direct core argument, the construction clearly being transitive. Now, as the discussion in the previous section has revealed, there is strong evidence that the passive agent in asymmetrical voice languages like German is not an adjunct but rather an oblique core argument, so that the transitivity of passive constructions should be reconsidered. It seems to be clear, however, that the asymmetric passive construction and the symmetric undergoer voices differ in their degree of transitivity. Second, and following from the first point, as the agent arguments in the undergoer voices, just like the undergoer argument in the actor voice, remain direct core arguments, in symmetrical languages the non-subject arguments show the same syntactic properties in all voices. In asymmetrical languages on the other hand, the undergoer in the active clause and the passive agent behave differently, the former being a direct core argument, the latter being oblique. Third, asymmetrical and symmetrical languages differ with respect to their morphological marking on the verb, with only the passive being marked in the asymmetrical languages, while in symmetrical languages the verb carries voice morphology in all voices.[3]

Totoli, a Western Malayo-Polynesian language spoken in the northern part of Central Sulawesi, Indonesia, shows all of the above-mentioned properties and will serve here as an example of a prototypical symmetrical voice language. Totoli exhibits one actor voice and two lexically determined undergoer voices. Table 4.1 shows the complete Totoli voice paradigm. As in many Austronesian languages, voice morphology in Totoli is inseparable from mood morphology, i.e. there is an obligatory distinction between actor voice and undergoer voice on the one hand, and non-realis mood and realis mood on the other hand. Table 4.2 illustrates the applicative forms, which are closely intervened with the voice morphology and partly overlap. Note especially, that forms with *-i*/*ni-* *-an* are ambiguous between a simple undergoer voice and an undergoer voice with applicative function. The exact relationship between voice and applicative functions is not important for this paper. Table 4.2 is merely included in order to help the reader better understand the glossing in the Totoli examples to follow (e.g. *-i* being either glossed as UV2 or as APPL2). The

[3] Note that in many Austronesian languages that display symmetrical voice, there is usually one slot in the verbal paradigm that remains morphologically unmarked. However, language inherent evidence as well as cross-linguistic comparison give reason to assume that the non-marked forms are a historical coincident rather than representing the "unmarked" voice (in the sense that the active represents the "unmarked" voice in the active-passive alternation). Thus, language-internally non-marked forms always stand in paradigmatic opposition to marked ones. Cross-linguistically, the non-marked slots do not occur in the same position in the paradigm (e.g. in colloquial Indonesian, the actor voice verb remains unmarked, in Tagalog realis patient voice verbs do not carry voice morphology and in Totoli there is no voice marking in one of the two undergoer voices. For a more detailed discussion in this issue see Riesberg (2011, section 2.2.5).

interested reader is referred to Himmelmann/Riesberg (2013), where the interplay between voice and applicatives in Totoli is discussed in great detail.

Table 4.1: The Totoli voice paradigm

	NON-REALIS	REALIS
AV	*mo*(*g*)-/*moN*-	*no*(*g*)-/*noN*-
UV1	Ø	*ni*-
UV2	*-i*	*ni- -an*

Table 4.2: The Totoli applicative paradigm

	APPLICATIVE 1			
	NON-RLS		RLS	
AV	*mo*(*g*)-/*moN*- *-an*		*no*(*g*)-/*noN*- *-an*	
UV	*po*(*g*)-/*poN*- *-an* *-an*	INS/BEN THEME	*nipo*(*g*)-/*poN*- *-an* *ni-*	INS/BEN THEME
	APPLICATIVE 2			
	NON-RLS		RLS	
AV	*mo*(*g*)-/*moN*- *-i*		*no*(*g*)-/*noN*- *-i*	
UV	*-i*	GOAL	*ni--an*	GOAL

The following two example pairs illustrate the morphological marking of the two undergoer voices in the b.-examples (in the realis mood in (20) and in the non-realis in (21)) and their respective actor voice forms in the a.-examples. As it can be seen, both the actor voice construction and the two undergoer voice constructions are morphologically marked:

(20) a.

TOT	*I*	*Rinto*	*manaip*	*taipang.*
	i	Rinto	**moN-**taip	taipang
	HON	PN	AV-peel	mango

'Rinto is peeling a mango.'

b.

Taipang	*nitaip*	*i*	*Rinto.*
taipang	**ni-**taip	i	Rinto
manggo	RLS-peel:UV1	HON	mango

'Rinto has been peeling a mango.'

(21) a.

TOT	*I*	*Winarno*	*mongusut*	*kunji*	*motorna.*
	i	Winarno	**moN-**kusut	kunji	motor=na
	HON	PN	AV-look.for	key	scooter=3s.GEN

'Winarno is looking for the keys for his scooter.'

b.

Kunji	*itu*	*kusuti*	*i*	*Winarno.*
kunji	itu	kusut-**i**	i	Winarno
key	DET	look.for-UV2	HON	PN

'Winarno is looking for the keys.'

Further illustrated by the four example sentences above is the fact that the non-subject phrase, i.e. the phrase that directly follows the verbs, is marked by neither adposition nor (oblique) case in all voices. Furthermore, in addition to exhibiting the same coding properties, they also show the same behavioural properties, illustrated below with an example of relativisation. As is known for many Austronesian languages (e.g. Tagalog (Schachter 1977), standard Indonesian (Sneddon 1996: 287ff.), Pendau (Quick: 508ff.), and many more), Totoli allows only the subject argument to be relativized. The attempt to relativize the object argument results in non-grammaticality, and this holds true for all three voice constructions. In the following examples, the inserted space holder (__) indicates the position of the respective relativized grammatical function. In the three grammatical examples (22a), (23a) and (24a), it is the subject of the relative clause (*mangngana* (*itu*) 'the children' and *doi* 'the money'); in the ungrammatical examples (22b), (23b) and (24b), it is the object (*deuk itu* 'the dog', *guru itu* 'the teacher' and *Yusuf*).

(22)	a.	*I*	*aku*	*nokoita*	*mangngana*	
TOT		i	aku	noko-ita	-ngo-mangana	
		HON	1s	POT.AV.RLS-see	-COLL-child	
		anu	____	*nambbagi*	*deuk*	*itu.*
		anu		noN-babag-i	deuk	itu
		REL		AV.RLS-beat-APPL2	dog	DIST

'I saw the children who were beating the dog.'

	b.	**I*	*aku*	*mokoita*	*deuk*	*itu*
		i	aku	moko-ita	deuk	itu
		HON	1s	POT.AV-see	dog	DIST
		anu	*mangngana*	*nambbagi*	____.	
		anu	-ngo-mangana	noN-babag-i		
		REL	-COLL-child	AV.RLS-beat-APPL2		

for: 'I see the dog that the children were beating.'

(23)	a.	*I*	*aku*	*nokoita*	*mangana*	*itu*
TOT		i	aku	noko-ita	-ngo-mangana	itu
		HON	1s	POT.AV.RLS-see	-COLL-child	DIST
		anu	____	*nipiolan*	*guruna.*	
		anu		ni-pio'-an[4]	guru=na	
		REL		RLS-pinch-APPL2	teacher=3s.GEN	

'I saw the child who was pinched by its teacher.'

	b.	**I*	*aku*	*nokoita*	*guru*	*itu*
		i	aku	noko-ita	guru	itu
		HON	1s	POT.AV.RLS-see	teacher	DIST
		anu	*anak*	*nipiolan*	____.	
		anu	anak	ni-pio'-an		
		REL	child	RLS-pich-APPL2		

[4] The phoneme /l/ in Totoli has three allophones. Word-finally it is realized as lengthening of the preceding vowel, here represented as <'>.

for: 'I see the teacher by whom the child was pinched.'

(24) a. TOT

I	*aku*	*nokoita*	*doi*	
i	aku	noko-ita	doi	
HON	1s	POT.AV.RLS-see	money	
anu	____	*lau*	*kusuti*	*Yusuf.*
anu		lau	kusut-i	Yusuf
REL		presently	look.for-UV2	PN

'I saw the money that Yusuf was searching.'

b.

**I*	*aku*	*nokoita*	*Yusuf*
i	aku	noko-ita	Yusuf
HON	1s	POT.AV.RLS-see	PN
anu	*doi*	*kusuti*	____.
anu	doi	kusut-i	
REL	money	look.for-UV2	

for: 'I saw Yusuf by whom the money was searched.'

The limits of this paper do not allow for a comprehensive overview of the syntactic behaviour of Totoli arguments. For a detailed discussion including the presentation of a wide range of Totoli data see Riesberg 2014. Table 4.3 below summarises the results of this discussion.

Table 4.3: Syntactic behaviour of arguments in Totoli

	AV		**UV1**		**UV2**	
	Subj	non-Subj	Subj	non-Subj	Subj	non-Subj
Relativisation	✓	x	✓	x	✓	x
Raising	✓	x	✓	x	✓	x
Control	✓	x (✓)	✓	x (✓)	✓	x (✓)
Fixed position	x	✓	x	✓	x	✓
Reflexive Binding	✓	x	x	✓	x	✓

The table reflects the symmetry of the Totoli voice system: Whenever an argument possesses a certain property (symbolized here as ✓), it does so in all three voice constructions (i.e. in the actor voice as well as in the two undergoer voices). Likewise, if an argument does not possess a given property (symbolized as x), it will not do so in any of the voices. Thus, for example, only the embedded subject, but not the embedded non-subject argument, can be raised to be realised as one of the matrix clause arguments. With respect to control constructions, speakers vary as to whether or not they allow control of non-subject arguments. Most speakers allow only the subject argument of a subordinated clause to be controlled, rejecting sentences with non-subject control. A few speakers, however, did accept (and produce) these kinds of constructions (hence ✓in parentheses). The crucial point to note is that both groups of speaker adhere to the overall symmetry of the system, i.e. those speakers that allow non-subject control do so not only in one but in all voices. The symmetry of the Totoli voice system is also reflected in the word order possibilities: In all three voices the non-subject has to immediately follow the verb, thus

building a unit with it. The subject, on the other hand might either precede or follow this unit.

A few words are needed in order to explain the behaviour of arguments with respect to reflexive binding. At first sight, the symmetry seems to be abolished: While in the actor voice only the subject can bind a reflexive non-subject argument, the reverse holds for the two undergoer voice construction, where only the non-subject argument can function as a binder for a reflexive subject. Binding of a reflexive subject by the non-subject argument in the actor voice or of a reflexive non-subject by the subject in the undergoer voices is not possible, as shown in the examples below: The examples in (25) illustrate two actor voice clauses where, in the a.-example, the subject *isia* 'she' binds the reflexive phrase *batanganna*.[5] The attempt to put the reflexive in subject position (thus making it the actor of the clause) and having it bound by the undergoer object *Efi* results in ungrammaticality, as shown in (25b).

(25) a.

TOT	*Isia*	*geimo*	*mamarakai*	*batanganna.*
	isia	geimo	moN-paraka-i	batangan=na
	3s	NEG	AV-take.care-APPL2	body=3s.GEN

'She doesn't take care of herself.'

b.

**Bantanganna*	*mogitai*	*Efi*	*dei*	*sindongan.*
batangan=na	mog-ita-i	Efi	dei	sindongan
body=3s.GEN	AV.see-APPL2	PN	LOC	mirror

for: 'Efi watched herself in the mirror.'

The reverse holds for the following undergoer voice example: In (26a) it is grammatical for the reflexive phrase to be the subject of the clause, being bound by the actor object (in this case the pronominal clitic =*na* 'she'). (26b) shows that the reflexive cannot be bound by the undergoer subject *Efi*.

(26) a.

TOT	*Batanganna*	*geimo*	*parakaina.*
	batangan=na	geimo	paraka-i=na
	body=3s.GEN	NEG	take.care-UV2=3s.GEN

'She doesn't take care of herself.'

b.

**Efi*	*niitaan*	*batangan*	*dei*	*sindongan.*
Efi	ni-ita-an	batangan	dei	sindongan
PN	RLS.see-UV2	body	LOC	mirror

for: 'Efi watched herself in the mirror.'

The first conclusion to be drawn from data like (25) and (26) is that binding in Totoli, as in many other Austronesian languages, adheres to semantic rather than to syntactic restrictions, the generalisation being that the higher semantic argument can function as a binder to the lower one but not vice versa. That is, an actor argument can always bind a reflexive undergoer argument, regardless of either being the subject or the object. The second important fact about binding in Totoli is that the binding phrase has to be a direct core argument. It is this restriction that makes the binding data relevant for discussion

[5] *Batanganna* literally translates as 'his/her body'. The genitive clitic =*na* is, however, not obligatory, as can be seen in example (26b).

about symmetry: If only core arguments are possible binders and if the actor in the undergoer voice *can* bind the undergoer subject, this means that the undergoer voice actor is indeed a core argument (cf. Arka 2003, Arka/Manning 2008).

This restriction becomes especially interesting when looking at languages like Indonesian and Balinese, which show the same binding behaviour as Totoli but, in addition to an actor voice and an undergoer voice, also display a passive construction. The actor in the passive, being marked by a preposition, is not a direct core argument, and thus the prediction would be that the passive actor, unlike the undergoer voice actor, will not be able to bind a reflexive undergoer subject. And indeed this prediction is borne out, as the Balinese data below show. While in (27a) the direct actor argument *ida* 'she' can bind the reflexive argument *ragenne* in subject position, the prepositionally marked passive actor *teken ia* 'by him' in (27b) cannot function as a binder for the reflexive *awakne* 'himself'.

(27)	a.	*Ragenne*	*tan*	*kanggoang*	*ida.*	(Arka 2003: 178)
BAL		self.3	NEG	UV:accept	3	

'She didn't accept herself.'

b.	*Anake*$_i$	*cenik*	*ento*	*edengina*	*awakne*$_{i/*j}$
	anak-e	cenik	ento	edengin-a	awakne
	child-DEF	small	DET	show-PASS	self.3
	di	*kacane*	*teken*	*ia*$_j$.	(Arka 2008: 81)
	di	kaca-ne	teken	ia	
	LOC	mirror-DEF	by	3	

'The child$_i$ was shown himself$_{i/*j}$ in the mirror by him$_j$.'

As indicated by the indexation, sentence (27b) is only grammatical with a reading where *anake* 'the child' and the third person reflexive *awakne* are co-referent. Attempted co-reference between the third person pronoun *ia* and the reflexive *awakne* yields ungrammaticality, proving that the oblique actors in the passive, unlike the direct core actors of the undergoer voice, cannot function as a binder to a semantically lower reflexive pronoun. Thus, despite its superficial asymmetry, data on reflexive binding provides strong evidence for the core status of the non-subject arguments in symmetrical undergoer voices (cf. Wechsler/Arka 1998), and thereby proving the overall symmetry of the actor voice – undergoer voice alternation.

However, not all languages that display symmetrical voice alternations exhibit the same degree of symmetry. Totoli seems to exhibit a rather high degree of symmetry, but in many (or most) symmetrical languages the boundaries between direct argumenthood and oblique argumenthood (and adjuncthood) are just as fluid as has been demonstrated for the asymmetric language German. Balinese is a good candidate to illustrate the gradience of argumenthood in a language with symmetrical voice. In addition to one actor voice and one undergoer voice Balinese possesses two passive constructions. Just like in Totoli, the undergoer argument of the actor voice and the actor argument of the undergoer voice show symmetrical behaviour with respect to many syntactic constructions. So, for example, neither can be relativized, raised, or controlled, whereas the subject arguments in both voices do allow for relativization, raising, and control. With respect to other syntactic constructions, however, the symmetric behaviour is given up. This is the case, for example, for quantifier float constructions.

Balinese has two different kinds of quantifiers: simple ones, consisting of one word only, and complex ones, consisting of two words. These two kinds of quantifier behave

differently in two respects. First, complex quantifiers can only modify noun phrases that refer to animate (mostly human) entities (Wechsler/Arka 1998: 404). Simple quantifiers adhere to definiteness restrictions in that they are restricted to modifying definite NPs (cf. Arka 2003: 45ff.).[6] Second, simple and complex quantifiers show different modification properties in the different voices. If a complex quantifier occurs in non-adjacent position to either of the two arguments of a transitive clause, this clause will be ambiguous: The quantifier can be interpreted as modifying either the subject or the non-subject (direct core) argument. Importantly, this behaviour can be observed for both actor voice clauses (28a) and undergoer voice clauses (28b) (both examples from Wechsler/Arka 1998: 404).

(28) a. *Ia* *nakonin* *tiang* *ibi* *ajak* *makejang*
BAL 3 AV-ask 1 yesterday accompanying all

'They all asked me yesterday./(S)he asked us all yesterday.'

b. *Ia* *dengokin* *tiang* *ibi* *ajak* *makejang*
3 UV:visit 1 yesterday accompanying all

'(S)he was visited by us all yesterday./They all were visited by me yesterday.'

Simple quantifiers do not preserve this symmetry, as the ambiguous reading is only available in the actor voice (29a). In the undergoer voice this kind of construction is no longer ambiguous, and the quantifier can only be interpreted as modifying the subject (29b) (examples from Arka 2003: 44ff.).

(29) a. *Cherik-cherike* *meli* *jajane* *ibi* *onya.*
BAL cherik-cherik-e N-beli jaja-ne ibi onya
child-child-DEF AV-buy cake-DEF yesterday all

'All the children bought the cake [yesterday].
The children bought all of the cake [yesterday].'

b. *Nasine* *jemak* *cai* *ibi* *makejang*
nasi-ne jemak cai ibi makejang
rice-DEF UV.take 2 yesterday all

All the rice was taken by you yesterday.
?*The rice was taken by you all yesterday.'

The same behaviour just illustrated for simple quantifiers also holds for secondary predicates. Just like simple quantifiers, secondary predicates behave asymmetrically in actor and undergoer voice, as only in the actor voice can they 'float away' from the non-subject argument. In the undergoer voice, this option is blocked. And also with respect to word order, actor voice and undergoer voice are not symmetric. Non-subjects of the former can be fronted into clause initial position, thereby forcing a reading with contrastive focus. For non-subjects in the undergoer voice, this is not possible. With respect to these properties, non-subject arguments in Balinese undergoer voice constructions behave in parallel to the actor in passive clauses, giving reason to question the direct core argument

[6] Examples that are supposed to illustrate the ambiguity of constructions with floated quantifiers thus have to be chosen with care: Sentences with complex quantifiers should include two animate arguments while sentences with simple quantifiers have to involve two definite NPs – otherwise, one possibly available reading might be blocked due to restrictions on either animacy or definiteness, respectively.

status of the undergoer voice actor. However, as shown above, there are other properties that give contradicting evidence. Complex quantifiers and binding data prove that the non-subject argument in the undergoer voice behaves – just as in the actor voice – like a direct core argument, clearly having to be distinguished from the oblique passive actor.

The discussion of the Balinese data thus shows that, just like in asymmetrical languages, also in symmetrical voice languages argumenthood seems to be a gradient concept: The Balinese undergoer voice actor behaves less direct core-like than the actor voice undergoer, but more direct core-like than the passive actor. This is supported by work by Arka (2009), who studies the behaviour of non-subject arguments in a number of Austronesian languages (including Balinese). Arka concludes that in these languages some arguments show an intermediate status between direct core and oblique.[7] He develops a procedure to determine what he calls the CORE INDEX, in order to ensure better comparability of the core status of arguments language-internally, as well as cross-linguistically (Arka 2009: 7).[8] To measure the core index, he investigates the behaviour of arguments in certain syntactic constructions that distinguish core arguments from obliques on a language-specific basis, together with some more general characterisations of argument status (such as subcategorisation and obligatoriness).The argument types tested are the actor argument of a bivalent predicate (ACT of AV), the patient of a bivalent predicate (P), the recipient, benefactive, or goal of a trivalent predicate (G), and the theme of a trivalent predicate (T). Also included are the actor of the undergoer voice construction (ACT of UV), the passive actor (ACT of PASS), and an adjunct (non-ARG). Besides simple and complex quantifier float, reflexive binding, the question of whether the phrase is obligatory, and whether it is marked by an adposition or not, Arka (2009) investigates seven more core properties[9], which thus adds up to twelve core properties altogether. The core index of a given phrase is then calculated by dividing the number of core properties that have been tested positively for a given phrase by the overall number of core properties. So, for example, the agent of the actor voice construction (ACT of AV) has all twelve core properties and thus a core index of (12:12=) 1.00. For those arguments that are of interest to the discussion here (i.e. P and ACT of UV), Arka calculates the following numbers: While the non-subject in the actor voice has a core index of 0.83, the value of the undergoer voice non-subject only amounts to 0.75, thus clearly being less core like than its actor voice counterpart. The same gradience of core argumenthood described for the asymmetrical language German can thus also be observed for the symmetrical voice language Balinese. Furthermore, in a more recent study, Arka (2013) does not only investigate the distinction between core arguments and obliques, but extends his method of calculating indices on the argument-adjunct distinction[10], calling it the ARGUMENT INDEX as opposed to the core index. The result of this work strongly supports the former claim that passive agents are arguments, showing that in Balinese the passive actor displays an

---

[7] Note that Arka uses the terms CORE and OBLIQUE as defined within LFG (see section 2), i.e. in his use of the terms, oblique arguments are not core. For those issues of his work that are being discussed in this paper, this deviation in terminology is not problematic.

[8] Arka (2009) also briefly discusses the fact that the core index is a helpful tool to make claims about the transitivity of constructions and about the degree of symmetry exhibited by a given voice system.

[9] These seven further properties are: topicalization of possessor phrases, topicalization with resumptive pronoun, modification by depictive predicates, the possibility of having a zero imperative actor, control of complex arguments, and position (i.e. fixed position vs. non-fixed position). For more detail, see Arka (2009).

[10] Note that the values calculated within these two works (Arka 2009 and Arka 2013) cannot be compared directly. This is due to the fact that more properties are considered when calculating the argument index.

argument index of 0.29, as opposed to an (non-argument) adjunct whose argument index is 0.00.

5 Conclusion

This paper has shown that neither in the asymmetrical voice language German, nor in the symmetrical voice language Balinese, is a clear-cut distinction between argumenthood and adjuncthood or between direct core and oblique core possible. Instead, it has been demonstrated that argumenthood (and adjuncthood) is better conceived of as a gradient concept. Against the common view that the actor argument in a passive construction is an adjunct, ZHS's study has shown that at least in German, the passive actor is closer to the argument side, even if – of course – it is less core-like than the object argument of an active construction. Likewise, it has been demonstrated that the undergoer voice actor in Balinese is less core-like than the actor voice undergoer, though still more core like than the passive actor. Following from these two observations, it seems reasonable to claim that the difference between symmetrical voice languages and asymmetrical voice languages is less fundamental than widely assumed. In both voice types – symmetrical and asymmetrical – the non-subjects remain core arguments, the only difference being that in the former they remain direct core in all voices while in the latter they become oblique core in the passive, thereby producing the asymmetry of the system. Furthermore, as the discussion above has show, even symmetrical languages do not always behave one hundred per cent symmetrically. While Totoli has been shown to display a very high degree of symmetry, the Balinese data were shown to behave less symmetrically. This, in turn, prompts the conclusion that also the (a)symmetry of voice constructions should better be conceived of as being gradient (a fact that in fact follows from the gradience of argumenthood). The three languages discussed in the paper can be ordered on a continuum as illustrated in Figure 5.1, with the actor voice-undergoer voice alternation in Totoli being the most symmetrical one, the German active-passive alternation being the most asymmetrical one, and the Balinese actor voice-undergoer voice alternation[11] finding its place in between, probably closer to the Totoli alternations than to the German ones.

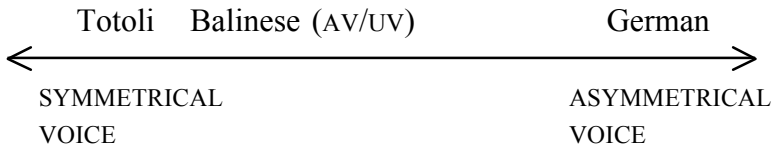


Figure 5.1: Totoli, Balinese and German ordered on a continuum of symmetry

References

[11] Note that Balinese also possesses a passive construction which is not considered in Figure 5.1.

Arka, I W., 2013, Locatives and the Argument-Adjunct distinction in Balinese. Paper presented at the Zentrum für Sprachenvielfalt und Mehrsprachigkeit, University of Cologne, April 2013.

Arka, I W., 2008, Voice and the Syntax of =a/-a Verbs in Balinese. In P. K. Austin & S. Musgrave, eds., Voice and Grammatical Relations in Austronesian Languages(pp. 70-89). Stanford: CSLI Publications.

Arka, I W., 2009, The Core-Oblique Distinction and Core Index in Some Austronesian Languages of Indonesia. Keynote paper presented at International ALT VI (Association of Linguistic Typology conference, Padang Indonesia, July 2005 under a different title 'On the distinction between Core and Obliques in the Austronesian languages of Indonesia'). http://rspas.anu.edu.au/linguistics/projects/iwa/Arka_papers.html (30th November, 2010).

Arka, I W., 2003, Balinese Morphosyntax: A Lexical Functional Approach. Canberra: Pacific Linguistics.

Bresnan, J., 1982, Control and Complementation. In J. Bresnan, ed., The Mental Representation of Grammatical Relations (pp. 282-390). Cambridge, Massachusetts/ London, England: The MIT Press.

Quick, P., 2007, A Grammar of the Pendau Language of Central Sulawesi, Indonesia. Canberra: Pacific Linguistics.

Dowty, D., 2003, The Dual Analysis of Adjuncts/Complements in Categorial Grammar. In E. Lang, C. Fabricius-Hansen & C. Maienborn, eds., Modifying Adjuncts (pp. 33-66). Berlin: Mouton de Gruyter.

Foley, W. A., 2008, The Place of Philippine Languages in a Typology of Voice Systems. In P. K. Austin & S. Musgrave, eds., Voice and Grammatical Relations in Austronesian Languages (pp. 22-44). Stanford: CSLI Publications.

Grimshaw, J., 1990, Argument Structure. Cambridge, Massachusetts/ London: The MIT Press.

Himmelmann, N. P., 2005, The Austronesian Languages of Asia and Madagascar: Typological Characteristics. In A. Adelaar & N. P. Himmelmann, eds., The Austronesian Languages of Asia and Madagascar (pp. 110-181). London/New York: Routledge.

Himmelmann, N. P., and S. Riesberg, 2013, Symmetrical Voice and Applicative Alternations: Evidence from Totoli. Oceanic Linguistics.

Jacobs, J., 1994, Kontra Valenz. Trier: Wissenschaftlicher Verlag Trier.

Riesberg, S., 2014, Symmetrical Voice and Linking in Western Austronesian Languages. Boston/Berlin: Pacific Linguistics – De Gruyter.

Schachter, P., 1977, Reference-Related and Role-Related Properties of Subjects. In P. Cole & J. M. Sadock, eds., Grammatical Relations (pp. 279-306). New York/San Francisco/ London: Academic Press.

Sneddon, J. N., 1996, Indonesian. A Comprehensive Grammar. London/New York: Routledge.

Van Valin, R. D., 2005, Exploring the Syntax-Semantic Interface. Cambridge: Cambridge University Press.

Van Valin, R. D., and R. J. LaPolla, 1997, Syntax. Structure, Meaning and Function. Cambridge: Cambridge University Press.

Wechsler, S., and I W. Arka, 1998, Syntactic Ergativity in Balinese: An Argument Structure Based Theory. Natural Language and Linguistic Theory 16:387-441.

Zifonun, G., Hoffmann, L. & Strecker, B. et al., 1997, Grammatik der Deutschen Sprache. Band 2. Berlin/New York: Walter de Gruyter.

16 *A comparative look at the major voice oppositions in Sama-Bajaw languages and Indonesian/Malay*

MARK MILLER

1 Introduction

The Sama-Bajaw (S-B) languages form a distinct subgroup of the Western Malayo-Polynesian branch of Austronesian languages (Ruhlen 1987:167). Most S-B languages are spoken in the Sulu (southern Philippines), with varieties also in northern Borneo and Sulawesi. Data for this paper are gathered from five S-B languages: Central Sama, Southern Sama, Sama Bangingi (all members of the Inner Sulu Sama subgroup), Sama Pangutaran (the sole member of the Western Sulu Sama subgroup), and West Coast (WC) Bajau (a member of the Borneo Coast Bajaw subgroup).[1] Of these five S-B languages, only WC Bajau does not have its geographical homeland in the Southern Philippines.

Despite having long historical contact with Philippine languages such as Tausug (Pallesen 1984), the major voice oppositions present in S-B languages show striking similarities to that of Indonesian/Malay. The S-B languages I have investigated share with Indonesian the following voice features:

(1) an actor voice (AV) construction marked by a nasal prefix;
(2) a transitive non-AV construction that occurs with the unaffixed verb root; and
(3) an alternate non-AV construction with morphological marking on the verb (usually *-in-* for S-B languages) and case marking on the agent.

Thus, in their basic voice oppositions the S-B languages pattern very similarly to Indonesian-type languages, even while they retain certain Philippine-type voice features such as additional "undergoer voices" mapped onto particular semantic roles.

2 Features of the voice system of Indonesian/ Malay

I will now summarize the primary features of the voice system in Indonesian/Malay. Indonesian is an SVO language, where the grammatical subject canonically appears before the verb. Indonesian transitive verbs show a tripartite system: they can appear with the

[1] The subgroups of the S-B languages are cited in the 16th Edition of the Ethnologue (2009).

Asia-Pacific Linguistics, 2014.

nasal prefix *meN-*, or with the *di-* prefix, or they may occur without affixation. When the verb is prefixed with *meN-* the actor is the grammatical subject and appears before the verb. This is the active voice (AV) construction:

(1) a. *Dia akan men-jual rumah itu.*
1sg FUT AV-sell house that
'He will sell the house.'

When the verb is prefixed with *di-* the undergoer is the grammatical subject and appears before the verb. The agent argument is expressed as a PP, is understood to be oblique, and may be optionally deleted. This is the affixed non-AV construction, generally analyzed as a passive:

b. *Rumah itu akan di-jual oleh dia.*
house that FUT PASS-sell by 3sg
'The house will be sold by him.'

b′. *Rumah itu akan di-jual.*
'The house will be sold.'

The affixed non-AV construction is limited to third person agents. Non-third person agents can be demoted using the syntactically similar*ter-* prefix on the verb, where the agent is also expressed as a PP. (However, the *ter-* prefix normally entails a non-volitional or abilitative meaning not exhibited by *di-*.)

When the verb is unaffixed, the agent is expressed by the preverbal pronominals *saya/kamu/dia* or by the proclitics *ku-/kau-*. This is the zero UV construction (what Arka and Manning call 'Objective Voice'):[2]

c. *Rumah itu akan dia Ø-jual*
house that FUT 3sg UV-sell
'He will sell the house.'

In the zero UV construction, the agent argument can be any person, but can only appear as a pronominal, never as a common noun. Some analysts (Chung 1976; Kana 1986) have treated the agent argument in the zero UV construction as a non-core argument, which would make this a passive construction. However, Arka and Manning (1998) have shown, mainly with evidence from binding, that the zero UV actor is still a term/ core argument. They also note that the zero UV actor cannot be omitted. These facts argue against a passive analysis for the zero UV construction.

It is important to note that the *di-* prefix in Indonesian does not necessarily mark a passive construction. In the construction [*di-* + verb + *-nya*], where the third person enclitic *-nya* is used on the verb to mark the agent, Arka and Manning (1998) show that the enclitic agent can bind the reflexive grammatical subject. This is evidence that the agent expresses a core argument rather than an oblique, even though the verb is marked with *di-*.

---

[2] In colloquial Indonesian, there is not only a zero UV but also a zero AV construction, since the AV verb is normally unaffixed (occurs without *meN-*).

3 Features of the voice systems of Sama-Bajaw (S-B) languages

Each of the S-B languages surveyed has, like Indonesian/Malay, an actor voice (AV) construction marked by a nasal prefix, an affixed non-AV construction, and an unaffixed ("zero") UV construction. In what follows, I will briefly describe these constructions in the different S-B languages according to: (1) morphology; (2) word order; (3) constraints on choice of UV agent; and (4) potential demotion of the agent in the affixed non-AV construction.

3.1 Morphology

In all S-B languages surveyed, the AV construction is marked by a nasal prefix (*N-* or *aN-*).[3] The affixed non-AV construction is marked by *-in-* or *-i-* and always co-occurs with a preposition for the agent. For each S-B language, the morphology of the AV construction, the affixed non-AV construction, and the preposition are shown in Table 3.1:

Table 3.1: Morphosyntactic expression of the major voice oppositions in S-B language

Sama-Bajaw language	**Morphology of AV construction**	**Morphology of affixed non-AV construction**	**Preposition associated with affixed non-AV construction**
Sama Bangingi'	*aN-*	*-in-/ni-*[4]	*e'* or *ni* (dialectical difference)
Central Sama	*aN-*	*-in-/ni-*	*e'*
Southern Sama	*N-*	*-in-/ni-*	*e'*
Pangutaran Sama	*N-*	*-i-/i-*	*uk*
West Coast Bajau	*N-*	*-in-/ni-*	*(o)le'*

In all S-B languages surveyed, when the affixed non-AV verb occurs, the preposition is obligatory. The following is an example of the affixed non-AV construction in WC Bajau:

(2)	*Beluang*	*e*	*pan*	*-in-buka*	*no*	*ole'*	*anak=ni*	*sioko.*
	door	DEM	TOP	-PASS-open	FOC	PREP	child=3s.I	oldest

[3] In some environments (generally before the voiceless obstruents and /b/), the nasal consonant assimilates to the point of articulation of the stem initial consonant and that consonant is deleted. In other environments (generally before continuants and voiced obstruents other than /b/) the *N-* is realized by allophonic variants and no coalescence occurs.

[4] The infix *-in-* is inserted immediately after the initial consonant of the verb stem; its allomorph, *ni-*, occurs before words beginning with a vowel, /h/ or /l/ (Gault 1999). The affixation rules appear to be similar for the other S-B languages surveyed.

'The door was opened by the oldest child.' (kerabaw 032)

Pallesen (1985:98) says that this "agentive construction," to his knowledge, is not found in other Philippine languages. Indonesian, on the other hand, is similar to the S-B pattern with regard to the affixed non-AV verb, since *di-* on the verb requires the use of *oleh* before the agent.

a. Word order

The word order(s) in the AV, zero UV and affixed non-AV constructions are shown in Table 3.2 for each of the S-B languages (where data is available):

Table 3.2: Preferred word orders for the major voice constructions in S-B languages

Sama-Bajaw language	**AV preferred word order**	**Zero UV preferred word order**	**Affixed non-AV preferred word order**
Sama Bangingi'	V A U		
Central Sama	V A U (but V U A when A is not a pronoun)	**V A** U	V U A or V A U (not sure which is preferred)
Southern Sama	V A U (but V U A when A is not a pronoun)	**V A** U	V U A (V A U also possible)
Pangutaran Sama	V A U	**V A** U	V A U or V U A (not sure which is preferred)
West Coast Bajau	A V U	**V A** U or U**V A**	U V A (less often V U A)

Pallesen (1985:95) notes that, with respect to word order, those S-B languages found in the Sulu (Philippines) are more like Philippine languages in that they prefer verb initial order. With these languages, the order S V U occurs only with preposed items such as negatives and aspect markers (which draw the syntactic subject to the preverbal position) or when the subject is topicalised for emphasis. In other S-B languages (such as West Coast Bajau) "the preverb ordering of subject appears to be unmarked, and occurs in narrative texts with much greater frequency" (Pallesen 1985:95). In this respect only the non-Sulu S-B languages pattern like Indonesian, which has S V U as its unmarked word order.

However, all S-B languages pattern similarly with regard to the zero UV construction, in that the agent follows the verb, and in at least some S-B languages (like Sama Bangingi' and West Coast Bajau) no intervening material is allowed. In Bangingi', the agent must be pronominal and is cliticized to the verb, so that "the verb stem plus pronoun are pronounced as one phonological word" (Gault 1999:11). In Indonesian, the OV construction shows the reverse order of agent followed immediately by the verb. But the

linkage between agent and verb is similarly rigid, in that the agent (whether free pronoun or proclitic) must occur immediately prior to the verb and no intervening material is allowed.

3.3 Constraints on choice of zero UV agent

As noted above, in Sama Bangingi' the agent in the zero UV construction must be a pronoun, and a similar constraint is observed in some other S-B languages. The constraints on non-AV actor (if any) that occur in the surveyed S-B languages are shown in Table 3.3:

Table 3.3: Constraints on the expression of non-AV actor in S-B languages

Sama-Bajaw language	**Zero UV agent**	**Affixed non-AV agent**
Sama Bangingi'	Pronominal clitic; more commonly 1st and 2nd person rather than 3rd person	Commonly 3rd person pronoun or full NP
Central Sama	Pronoun, sometimes (always?) cliticized; commonly 1st person (2nd and 3rd person only in grammatically restricted contexts)	Full NP or pronoun
Southern Sama	Pronoun, sometimes (always?) cliticized	Full NP or pronoun
Pangutaran Sama	Pronoun or full NP	Full NP or pronoun
West Coast Bajau	Pronoun or full NP	Full NP or pronoun

Table 3.3 shows that for the Inner Sulu Sama languages (Sama Bangingi', Central Sama, and Southern Sama) the zero UV agent is restricted to pronouns, more commonly 1st or 2nd person than 3rd person. In these languages, a non-AV agent expressed as a common or proper noun <u>must</u> take the affixed non-AV construction. Similar restrictions apply in Indonesian, where agents expressed as first and second person pronouns must take the zero UV construction, and agents expressed as full NPs must take the *di-* affixed construction (Sneddon 1996:250).

Note that in all surveyed S-B languages (including Pangutaran Sama and West Coast Bajau) <u>the zero UV agent is required</u>, as shown by the following pair of sentences in Pangutaran Sama:

(3) a. *Ø-b'lla* *d'nda* *kiyakan* *kami.*
UF-cook girl food our
'The girl cooked our food.' (Walton 1986:117)

b. **Ø-b'lla* *kiyakan* *kami.*
UF-cook food our

'Cooked our food.'

Similarly, the zero UV agent is required in Indonesian. Recall that Arka and Manning (1998) consider this as evidence for the zero UV agent being a non-subject term/core argument, though their claim is corroborated with evidence from the binding properties of the zero UV agent. To be more certain of the grammatical status of the zero UV agent in the S-B languages, its binding properties would also need to be tested in these languages.

Note finally that the zero UV agent tends to be highly topical, especially for the S-B languages spoken in the Sulu. As noted by Joanne Gault (p.c.): "An obligatory pronoun agent which is commonly first or second person, and which is phonologically attached to the verb means that the agent is highly topical." High discourse topicality may be additional evidence for the zero UV agent being a term/core argument. In Southern Sama the zero UV construction "appears to focus attention on the action. For future events, the construction implies that the action is imminent", or the end point is already specified (Brainard and Trick, unpublished workshop paper). Gault (1999:11) notes that for Sama Bangingi', the use of the zero UV construction "is commonly limited to the immediate context of the speaker and hearer, and normally refers to an action just completed or about to be accomplished by one or the other. In narratives it is largely limited to dialogue…". Here is an example of the zero UV construction in a Sama Bangingi' text (from Gault 1999:11):

(4)

Manjari,	*iyuk*	*si*	*Kuyya'*	*ma*	*si*	*Ba'uu*,
so-then	said	PM	monkey	OBL	PM	turtle

"***Bahagi'-ta***	*na*	*in*	*batang*	*saing*	*itu*."
UF.divide-1+2s	now	TM	trunk	banana	this

'So Monkey said to Turtle, "Let us divide this banana plant." ' (KB.6)

3.4 Potential demotion of the affixed UV agent

In Pangutaran Sama and in WC Bajau, the affixed non-AV construction involves the syntactic demotion of the UV agent to oblique status, and is analyzed as a passive. This is shown for Pangutaran Sama in (5a-c) below, where (5a) shows the AV construction, (5b) shows the zero UV construction (glossed as "UF"), and (5c) shows the affixed non-AV construction (glossed as "PASS-") where the oblique agent is marked by the preposition *uk*:

(5)

a.

N-b'lla	*d'nda*	*kiyakan.*
AF-cook	girl	food

'The girl cooked food.'

b.

Ø-b'lla	*d'nda*	*kiyakan*	*kami.*
UF-cook	girl	food	our

'The girl cooked our food.' (Walton 1986:117)

c.

-i-b'lla	*uk*	*d'nda*	*kiyakan*	*kami.*
-PASS-cook	by	girl	food	our

'Our food was cooked by the girl.' (Walton 1986:117)

c′.

-i-b'lla	*na*	*kiyakan*	*kami.*
-PASS-cook	now	food	our

'Our food is already cooked.' (Walton 1986:118)

A further set of examples is shown for West Coast Bajau, which uses the preposition (*o*)*le'* for the affixed non-AV construction. Commonly in WC Bajau the actor (*o*)*le'* phrase occurs following the undergoer, as in (6a). It may even occur after an adverbial oblique, as in (6b). Neither of these positions is possible for the actor argument of a zero UV clause, because in the zero UV clause the actor must follow directly after the verb (6c): (examples are from Miller 2007:169)

(6)	a.	*Ai*	*-in-kakan*	*moto*	*diing*	*e*	*le'*	*Kuzik*	*dilaw.*
		PERF	-PASS-eat	eye	fish	DEM	PREP	PN	yesterday
		'The fish eye was eaten by Kuzik yesterday.'							
	b.	*Ai*	*-in-kakan*	*moto*	*diing*	*e*	*dilaw*	*le'*	*Kuzik.*
	c.	**Ai*	*Ø-kakan*	*moto*	*diing*	*e*	*Kuzik*	*dilaw.*	
		PERF	UV-eat	eye	fish	DEM	PN	yesterday	
		'Kuzik ate the fish eye yesterday.'							

The multiple ordering possibilities associated with the (*o*)*le'* agentive phrase shown in (6) above are indicative of oblique behavior and contrast with the rigid verb-actor word order observed in the zero UV construction. The same pattern holds for the *oleh* agentive phrase in Indonesian, which (as an oblique prepositional phrase) is not required to immediately follow the *di-* affixed verb.

What about the status of the affixed non-AV actor in the other surveyed S-B languages? As in West Coast Bajau and Pangutaran Sama, it is possible in Southern Sama, Sama Bangingi', and Central Sama for the agentive phrase in the affixed non-AV construction to move to the right of any peripheral elements in the clause or to be omitted altogether, though some restrictions may apply. Note the following examples for Central Sama (Kemp Pallesen, p.c.), where syntactic demotion of the actor is possible when the actor is a full NP (7a) but not when the actor is a pronoun (7b-c):

(7)	a.	*Bay*	*-in-timbak*	*daing*	*di'ilaw*	*e'*	*si*	*Jiyani/ pulis.*
		past	-PASS(?)-dynamite	fish	yesterday	by	PM	Jiyani/ some police.
		'Jiyani/some police dynamited fish yesterday.'						
	b.	*Bay*	*-in-timbak*	*e'-ku*	*daing*	*di'ilaw.*		
		past	-PASS(?)-dynamite	by-me	fish	yesterday		
		'I dynamited fish yesterday.'						
	b′.	**Bay*	*-in-timbak*	*daing*	*di'ilaw*	*e'-ku*		
		past	-PASS(?)-dynamite	fish	yesterday	by-me		
		'I dynamited fish yesterday.'						

In Central Sama, Sama Bangingi', and Southern Sama, the affixed non-AV construction is somewhat perplexing. It might sometimes be analyzed as a passive, as suggested by the optional omission of the actor (though this does not in itself constitute a proof of syntactic passivity). Discourse considerations make it more resemble an active construction in many or most of its occurrences. In fact, Pallesen avoids the term 'passive' in describing the affixed non-AV construction in Central Sama. Brainard and Trick (unpublished paper)

posit for Southern Sama two different constructions: one is 'inverse' (an active construction) and the other is 'passive', which are often identical in their morphosyntax (though when the actor is omitted, the construction is clearly a passive). Gault (1999) refers to the affixed non-AV construction as 'narrative mode', noting that "it is the most frequently used construction in foreground material…. Once the setting has been given and the participants introduced in a Bangingi' narrative, nearly all transitive events are non-agent constructions in either the narrative or abilitative modes" (17). Gault notes also that "often the agent is not expressed, not because it has been demoted, but precisely because it is highly topical. So topical in fact that is can be safely elided and still be understood" (Joanne Gault, p.c.). Understandably, given the high discourse topicality of the affixed non-AV agent, Gault hesitates to apply the label "passive" to this construction. Even in WC Bajau, where the affixed non-AV construction is syntactically a passive, the frequency of this construction on the story line (40%) in a corpus of narrative material is much higher than might be expected for passives in discourse. (Compare the frequency of the *ka-* passive on the story line in Balinese: 10%, as reported by Pastika [1999:160]).

Interestingly, the *di-* passive in Classical Malay was often used for foreground material (Hopper 1983). I suggest that, from a discourse perspective, the *di-* construction(s) in Indonesian/ Malay are similar to the affixed non-AV construction S-B languages, in that both show a higher percentage of foregrounded material than would be typical of a 'passive' construction, and higher topicality of the actor argument. Recall too that Arka and Manning (1998) consider the agent in the [*di-* + verb + *-nya*] construction in Indonesian to be core rather than oblique, so that *di-* marking on the verb does not always correlate with a passive analysis. Similarly, at least for the Inner Sulu Sama subgroup of S-B languages, *-in-/-in* marking on the verb is not consistently analyzed as passive.

4 Conclusion

We have seen that S-B languages show several similarities with Indonesian/Malay in terms of the major voice oppositions (actor and undergoer voice). The S-B languages, like Indonesian, show a tripartite pattern of AV, zero UV, and affixed non-AV. The affixed non-AV construction obligatorily takes a preposition (*oleh* in Indonesian; *(l)e'* or *uk* in the S-B languages). Such agentive phrases are generally not found in other languages of the Philippines. The zero UV construction in both Indonesian and the S-B languages is characterized by obligatory expression of the actor and its tight pairing with the verb. For Indonesian and three of the surveyed S-B languages, the zero UV construction is limited to 1st and (sometimes) 2nd person pronouns, and a full NP must be expressed by the affixed UV construction. Finally, in all surveyed S-B languages and in Indonesian there is the potential for syntactic demotion of the affixed non-AV actor, but also the tendency for high actor topicality and occurrence of this construction in foreground material.

What are we to make of these similarities between Indonesian/Malay and the S-B languages? Apart from the major voice oppositions, S-B languages are also like Indonesian (and unlike most Philippine-type languages) in that they lack case marking particles on common nouns. On the other hand, as we saw earlier in Table 3.2, all major voice constructions in most S-B languages are verb initial. Most of the S-B languages (particularly the Sulu groups) have voice mappings to various semantic roles such as beneficiary, instrument, and location; and the AV construction in these languages is typically of lower transitivity than the non-AV constructions. These features are characteristic of Philippine-type voice systems (Arka and Ross 2005:7). It is therefore difficult to pinpoint just where the S-B languages belong in their typological classification, as they do not easily fit within either a Philippine-type or Indonesian-type voice system.

What is surprising, given their contact of several hundred years with languages of the southern Philippines (especially Tausug: see Pallesen 1986), is why the S-B languages should nevertheless show extensive similarities with Indonesian and other Indonesian-type languages. These observations support Kemp Pallesen's (1985) conclusion that Proto Sama-Bajaw (PSB), given its seven-vowel system and a marked agentive phrase, likely had an Indonesian origin (245). More specifically, Blust (2005) has put forth the hypothesis (based on lexical evidence) that Proto-S-B had its origins in southeast Kalimantan, with extensive exposure to the Sriwijayan Malay through trade.

In order to make further progress on the classification of the S-B languages, a more complete analysis of their voice systems is required. In particular, the morphosyntactic status of the affixed UV agent argument could be profitably analyzed using a set of 'core vs. oblique' tests, such as have been identified for Balinese and Indonesian by Arka (2005).

References

Arka, I Wayan, 2005, The core-oblique distinction and core index in some Austronesian languages of Indonesia. Paper presented at the International ALT VI (Association of Linguistic Typology) Conference, Padang, Indonesia.

Arka, I Wayan and Christopher D. Manning, 1998, Voice and grammatical relations in Indonesian: A new perspective. In Miriam Butt and Tracy Holloway King, eds *Proceedings of the LFG '98 conference*. Stanford: CSLI. URL http://www-csli.stanford.edu/publications.]

Arka, I Wayan and Malcom Ross, 2005, Introduction. In I WayanArka and Malcom Ross, eds *The many faces of Austronesian voice systems: Some new empirical Studies*, 1-15. Canberra: Pacific Linguistics.

Blust, Robert, 2005, The linguistic macrohistory of the Philippines: Some speculations. In Hsiu-chuan Liao and Carl R. Galvez Rubino, eds *Current Issues in Philippine Linguistics and Anthropology: Parangalkay Lawrence A. Reid*, 31-68. Manila: The Linguistic Society of the Philippines and SIL Philippines.

Brainard, Sherri and Phyllis Trick, The Southern Sinama affix *-an*: Valence increaser and classifier. Unpublished MS; Summer Institute of Linguistics.

Gault, JoAnn Marie, 1999, *An ergative description of Sama Bangingi'*. Manila: Linguistic Society of the Philippines.

Chung, Sandra, 1976, On the subject of two passives in Indonesian. In C.N. Li, ed *Subject and Topic.* New York: Academic Press.

Hopper, Paul J, 1983, Ergative, passive and active in Malay narrative. In Flora Klein-Andreu, ed *Discourse: Perspectives on Syntax*, 67-88. Paris: Academic Press.

Kana, Marit A, 1986, Grammatical relations in Bahasa Indonesia. Ithaca: Cornell University. PhD dissertation.

Lewis, M. Paul (ed.), 2009, *Ethnologue: Languages of the World*, Sixteenth edition. Dallas: SIL International. URL http://www.ethnologue.com

Miller, Mark, 2007, A grammar of West Coast Bajau. Arlington: The University of Texas

at Arlington. PhD dissertation. URL http://hdl.handle.net/10106/577

Pallesen, A. Kemp, 1985, *Culture contact and language convergence*. Manila: Linguistic Society of the Philippines.

Pastika, I Wayan, 1999, Voice selection in Balinese narrative discourse. Canberra: The Australian National University. PhD dissertation.

Ruhlen, Merritt, 1987, *A guide to the world's languages.* Stanford: Stanford University Press.

Sneddon, James Neil, 1996, *Indonesian: A comprehensive grammar.* London, New York: Routledge.

Walton, Charles, 1986, *Sama verbal semantics: classification, derivation and inflection.* Manila: Linguistic Society of the Philippines.

17 *Voice-related constructions in the Austronesian languages of Flores*

I WAYAN ARKA AND FAY WOUK

1 Introduction

This paper discusses findings of our research on voice in the Austronesian languages of Flores, Indonesia.[1] Research reported in this paper is part of a broader inquiry into voice and voice expressions in the languages of Nusa Tenggara. We begin with a brief introduction to the typology of Flores languages and their position within the Austronesian family. We then present data showing the types of voice constructions and their properties. We claim that, while typically lacking voice morphology, Flores languages show remarkably rich voice oppositions. Finally, we conclude with a discussion of issues relating to the historical question of the total loss of the Austronesian voice system in these languages and the problem of how to best analyse the data.

2 The Austronesian languages of Flores: an overview

While it is commonly accepted that Flores languages belong to the Central Malayo-Polynesian (CMP) subgroup, the precise details of the genealogical relationships between Flores languages and other languages outside Flores is far from clear. Esser includes Flores languages, together with Bimanese (spoken in Sumbawa), the languages of Sumba, and Sawu in his Bima-Sumba group. However, Blust (2008) argues against such a grouping, although he allows that there is limited evidence of a connection between a proposed Sumba-Hawu group and many of the languages of Western and Central Flores. Dyen (1965) assigns Sikka, a language in eastern Flores, to his Moluccan Linkage, but Blust (1993:263) (admitting the problem of reconstructing languages that form a linkage as defined in Ross (1988:8)) claims that there is evidence for internal grouping of CMP languages geographically into those in the lesser Sundas (which includes Flores) and those in the Moluccas. However, the evidence for the phylogenetic unity of the grouping labels of CMP and WMP (western Malayo-Polynesian) has been questioned (see Donohue and Grimes (2008) for detailed discussion, and Blust (2009) for counter-arguments.

Fernandes (1996) argues that Flores languages are legitimately grouped together as a subgroup within the CMP group.[2] We assume the family grouping shown in (1).[3]

[1] This research is funded by NSF grant BCS-0617198. The data on Rongga was also collected as part of Rongga Documentation Project funded by the ELDP grant (IPF0011) awarded to I Wayan Arka.

[2] Note that the so-called Flores languages in (1) also include non-Flores languages. The following languages are not (all) spoken on the island of Flores: Komodo is spoken on the island of Komodo (although it is also spoken in some areas in west Manggarai); Palu'e is spoken on the island of Palu'e,

Asia-Pacific Linguistics, 2014.

(1) The internal subgrouping of the languages of Flores, adapted from Fernandes (1996)

(Proto) Flores
- West Flores
  - Komodo-Manggarai-Rembong subgroup
    - Komodo
    - Manggarai-Rembong
      - Manggarai
      - Rembong
  - Ngadha-Lio-(Palu'e) subgroup
    - Ngadha-Lio
      - Rongga
      - Ngadha
      - Nage-Keo
      - Ende
      - Lio
    - Palu'e
- East Flores
  - Sikka
  - Lamaholot
  - Kedang

Typologically, Flores languages are generally SVO and isolating. However, they also vary considerably, e.g., Manggarai has relatively rich cliticisation and Lamaholot has verbal agreement while Lio and Sikka do not. Examples of basic constructions in these languages are given in section 4 below.

3 Voice defined in brief

In this paper voice is defined as a language-specific system of grammatical opposition pertaining to stages of event realization and the conceptual-pragmatic relevance of the participants of the event (cf. Shibatani 2006). The opposition may be coded by at least one of the following strategies: different verbal marking (morphological and/or periphrastic, e.g. English *be+V-en* for passive), different argument marking (morphological/prepositional case, e.g. English *by*+NP[agent], Manggarai prepositional clitic *le/l*=NP[agent] (example (11) below), and different linear order (i.e. analytic coding), e.g. example (16) from Palu'e. Voice alternation often, but not necessarily always, involves a change of grammatical relations. The active-middle opposition in Sanskrit given below, for example, involves no change in linking of the actor and undergoer arguments.

north of Flores; Lamaholot, while spoken in eastern Flores, is mainly spoken in the small islands of Adonara and Solor, east of Flores; Kedang is spoken on the island of Lomlen east of Solor. It should be also noted that there are other small undocumented languages in central Flores (Riung, Rajong, Waerana, Manus, and Kepo' not shown in (1) but see SIL International (2001) and Arka (2005a).

[3] Further historical linguistics research on Flores is needed for a precise internal subgrouping of Flores languages. This includes checking Fernandes' (1996) findings and gathering more data, especially about the small languages in central Flores, and verification of the inclusion of Kedang as part of the East Flores subgroup.

(2) a. *Katam* *karo-ti* (active) Sanskrit
mat.ACC make-3SG.ACT.PRES
'He makes a mat.'
b. *Katam* *karu-te* (middle)
mat.ACC make.3SG.MID.PRES
'He makes a mat (for himself).' (Shibatani 2007)

Some explanation of the keywords used in the definition of voice adopted in this paper is in order. The term 'event' covers actions, processes and states. Stages of event realization include linguistically important phases of initiation, development, and extension and/or termination.

Relevant to the initiation phase is agency or lack thereof. Many languages show a verbal opposition between volitional and spontaneous events, e.g. Japanese *omo-u* 'think (volitionally)' vs. *omow-are-ru* 'think (spontaneously)'. Split intransitive marking is known to be sensitive to the semantic opposition of volition, e.g. Actor Voice vs. Undergoer Voice verbal marking in Balinese is also used for certain intransitive verbs: *ng-enceh* 'AV.urinate (volitionally)' vs. *enceh-enceh* 'UV.urinate-RED (involuntarily)'.

Relevant to the development phase is the degree of affectedness or individuation of the patient. Languages often make a distinction as to whether or not affectedness involves a patient that is referentially distinct from the participant that initiates the action (i.e. the agent). This gives rise to the distinction between active and middle voice.

Related to the development phase is the extension and/or termination, whether an action develops in its 'normal' course, whether a change of state (i.e. result) is produced, and/or whether it involves and extends to an additional participant. This gives rise to the distinction between simple and benefactive/applicative constructions.

A particular construction may express or highlight one or more than one of these evolutionary phases of an event. For example, the Balinese stative passive *ma-* as in *ma-bukak* 'open(ed)' expresses the resultative end state without agency whereas the passive *-a, bukak-a* 'be opened by somebody' expresses termination with agency.

The conceptual and pragmatic relevance of participants in a given context is also a well-known factor in voice opposition. The conceptual and pragmatic prominence of agentivity and patientivity gives rise to an active/passive opposition.

While all languages may have ways to express the different phases of events just described, not all express those phases in a systematic way constituting a voice system. When languages do exhibit voice systems, they generally vary in how distinctions are coded. Common means of coding, as mentioned earlier, include differential verbal marking and/or argument coding.

An interesting question would be what happens when (verbal) voice coding is lost or is being lost. Our research on voice in Flores languages reported in this paper investigates this phenomenon.

We shall discuss the following points. First, the loss of (Austronesian) voice morphology in the isolating languages of Flores does not always mean the loss of voice oppositions. Second, voice-related oppositions can exist without making use of verbal marking. Third, if we compare these Flores languages with the Austronesian languages to the west, focusing on the island from Bali to Sumbawa, we will observe an interesting pattern of gradual attrition of the voice system (coding) from Balinese to Sasak, to Sumbawa and Bima, and Flores languages (Wouk 2002; Arka 2003c; Shibatani 2005). However, it is not immediately clear that this gradual picture of attrition reflects shared historical development of analytical voice in the Austronesian languages in this area.

4 Grammatical relations and Voice types in Flores Languages

All Flores languages are SVO(O) languages. Core arguments--Actor (A) and Undergoer (U)-- are NPs, and Obliques are generally PPs. Examples (3) through (7) below illustrate this basic construction for a variety of languages:

(3) *Ata sua situ kole kawe wua.* (Manggarai)
person two that again search rattan
'Those two men returned (from) looking (for) rattan.'

(4) *Ana ndau ka maki* (Rongga)
child that eat rice
'That child ate/eats rice'

(5) *Go sorõ na/Kewa muko* (Lewotobi Lamaholot)
1s give 3s/Kewa banana
'I gave him/Kewa (a) banana'

(6) *A'u toma au* (Sikka-Krowe)
1s found 2s
'I found you'

(7) *'Iné ti'i 'imu jawa.* (Nage Keo)
mum give 3sg corn (Baird 2002:76)
'Mum gave her corn.'

There is evidence for the existence of grammatical subject (SUBJ), or syntactic PIVOT[4], distinct from logical subject or agent in Manggarai and Rongga (Arka and Kosmas 2005; Arka, Kosmas, and Suparsa 2007) and also perhaps Palu'e (Donohue 2005). That is, the agent argument grammatically appear in different syntactic functions, e.g. as SUBJ in an active sentence and an oblique in its passive counterpart.

Grammatical subject in these languages is structurally preverbal, and has certain exclusive properties associated with relativisation and control. In terms of coding, SUBJ in Manggarai and Rongga is an NP. When the A argument appears as an Oblique it must be a PP. The examples in (8) from the Rego dialect of Manggarai (Arka and Kosmas 2005) show a contrast in relativisation. Sentence (8a) exemplifies subject relativisation, which is acceptable, and sentence (8b) shows object relativisation, which is unacceptable. Furthermore, the patient can be relativised through passivisation as in (8c), where it is made subject, and the agent must appear in an Oblique PP.

(8) a. *Ata molah [se __ita aku] ghitu rebao ngo gi* (Manggarai)
person girl REL see 1s that just.now go already
'The girl [who saw me] has just gone'

---

[4] The term subject has been used in the literature in different senses. It is also informally used to mean logical subject, i.e. agent. In this paper it is used as a syntactic function/role, represented in capital letters (SUBJ), which is equivalent to syntactic PIVOT in RRG (Foley and Van Valin 1984; Van Valin and LaPolla 1997), syntactic (e.g. S/A) alignment in Dixon's Basic Linguistic Theory (Dixon 1994, 2010), or syntactic Topic in Shibatani's (2009) description of the Philippine/Formosan languages. However, the term Topic itself is also used to in the literature in relation to pragmatic information structure. This discourse Topic NP is not necessarily grammatical SUBJ. To avoid confusion, if necessary, an explicit modifier will be used, e.g. grammatical subject as opposed to logical subject, or syntactic Topic as opposed to discourse Topic.

b. **Ata* *molah* *[se* *aku* *ita* *__]* *ghitu* *rebao* *ngo* *gi*
person girl REL 1s see that just.now go already
'The girl [that I saw] has just gone'

c. *Ata* *molah* *[se __* *ita* *l=aku]* *ghitu* *rebao* *ngo* *gi*
person girl REL see by=1s that just.now go already
'The girl [that I saw or that was seen by me] has just gone'

4.1 Passive Voice with prepositional obliques

An active-passive opposition with a change in structural coding, as shown in (9), is observed in Manggarai and Rongga. This change of structural coding, particularly the demotion of the Agent to Oblique, is typical for passivisation. However, the verb in these languages is not morphologically marked. As seen in these examples, the same verb forms, *pongga* (10) and *cero* in (11), are used in both active and passive sentences.

(9) a. NP_A V NP_U (active)

b. NP_U V (PP_A) (passive)

(10) a. *Ardi* *pongga* *ana* *ndau* (Rongga)
A hit child that
'Ardi hit the child'

b. *Ana* *ndau* *pongga* *ne* *Ardi*
child that hit by A
'The child was hit by Ardi'

(11) a. *Aku* *cero* *latung=k* (Manggarai)
1s fry corn=1s
'I fry/am frying corn'

b. *Latung* *hitu* *cero* *l=aku=i*
corn that fry by=1s=3s
'The corn is (being) fried by me'

As seen from example (11), in addition to agent demotion, the voice change in Manggarai is also encoded by a change in subject co-referential cliticisation, *=k* vs. *=i.* In the active construction (11a), the agent '1s' is subject and the clitic *=k* co-indexes the free NP *aku*. In the passive counterpart (11b), the patient *latung* 'corn' is subject, and the agent *aku* '1s' is an oblique, appearing with a prepositional clitic *l=*. Crucially, the co-indexing clitic is *=i*, in agreement with the NP *latung*.

When both the clitic and its free co-indexed NP are present, they are both associated with subject, but structurally the enclitic appears to be in grammatical subject position, and the free NP in the Topic position (see Arka and Kosmas 2005). That this co-indexed argument is subject is shown by the fact that it is the sole argument of the intransitive clause as exemplified in (12) below. Furthermore, the clitic itself can appear as the subject without a cross-referenced NP in intransitive and transitive clauses.

(12) a. ***Hia*** *pa'u eta mai bubung mbaru hitu=**i**.* (Manggarai)
3s fall above from top.roof house that=3s
'(S)he fell down from the roof top of the house.'

b. ***Hi*** ***Kode*** *ka'eng wa tana=**i**.*
ART monkey stay down ground=3s
'The monkey lives on the ground.'

(13) a. *Ongga aku=**i**.* (Manggarai)
hit 1s=3s
'He hit me.'

b. *Iti wa tanah ghitu=i.*
that down soil that=3s
'He is down there inside the ground.'

4.2 Atypical passive or undergoer voice

Languages of central and eastern Flores show a structural alternation of the type shown in (14). Examples from Sikka, Palu'e and Lio are given in (15)-(17). Note that the A argument in the non-active constructions (b) cannot be backgrounded. It is neither placed after the verb, nor does it appear as a PP, as occurs in the Manggarai/Rongga passive. The A argument is also often obligatorily present when the U is fronted.

(14) a. NP_A V NP_U (active)

b. NP_U NP_A V (passive or inverse)

(15) a. *Petrus gita ilin ia* (Sikka)
Petrus see mountain that
'Petrus saw the mountain'

b. *Ilin ia Petrus gita.*
mountain that Petrus see
'The mountain, Petrus saw'

c.* *Ilin ia toma gita é'i Petrus*
mountain that can see by Petrus

(16) a. *Ia cube vavi vaʔa.* (Palu'e)
3sg shoot pig that (Donohue 2005)[5]
'He shot that pig.'

b. *Vavi vaʔa ia cube.*
pig that 3sg shoot
'That pig, he shot (it).' ~ 'That pig was shot by him.'

(17) a. *Petrus tebo ana ghea* (Lio)
Petrus hit child that
'Petrus hit the child'

b. *Ana ghea Petrus tebo*
child that Petrus hit
'The child, P hit (him), or the child was hit by Petrus'

[5] Palu'e examples are presented in a phonemic transcription (see Donohue 2005).

There is evidence in Palu'e (Donohue 2005), Sikka (Sedeng 2000) and Lio (Sawardi 2000) that the structure with a fronted U as in the (b) sentences in examples (12)-(14) is a non-active voice construction. Donohue argues that it is a passive in Palu'e. Evidence comes from the fact that the U argument in (16) gains promotion to grammatical subject. For example, he demonstrates that it is structurally in subject position, not in preclausal topic position. It also acquires other subject properties such as simple quantifier float, conjunction reduction, and purposive clauses (see Donohue 2005 for details).

This passive appears to be an atypical passive. The A is still encoded as a preverbal NP and is obligatorily present. In Donohue's analysis, this A argument is an Oblique-like argument, rather than a core argument, mainly based on the evidence that this preverbal A cannot launch quantifier float associated with the simple quantifier *tetiʔón* 'all' as seen in (18b). Simple quantifier float is a property of core arguments (A/S/P) in Palue.

(18) *a.* *Konen bere somu vaʔa tetiʔón.*
3PL chop garlic that all
'They chopped all of that garlic.'
~ 'All of them chopped that garlic'

b. *Somu konen bere tetiʔón.*
garlic 3PL chop all
'They chopped all of the garlic.'
~ 'The garlic was all chopped by them.'
* 'All of them chopped the garlic.'

Applying a core-index analysis (Arka 2005b) to A of the P-A-Verb in Palu'e, we can determine the core-oblique status of A more precisely. Its core index is 0.87 (i.e. satisfying six out seven core properties). It is therefore classifiable as a core argument. The following core properties are identifiable in Palu'e on the basis of Donohue's description: (i) structurally in A-Verb-P, A is in subject position and P in object position, (ii) allowing simple quantifier float, (iii) ability to bind a core reflexive argument, (iv) typically NPs, (v) participate in analytic voice alternation, vi) subcategorised for, and vii) obligatory. The preverbal A definitely satisfies properties (iii), (iv), (v), (vi) and (vii). Structurally, the position of A in P-A-V is analysable as [Spec, VP] rather than an adjunct position; hence A is also a core position. Overall, A of the P-A-Verb structure is not significantly demoted to Oblique status. Its core index of 0.87 typologically signifies a high degree of coreness (i.e. a typical core argument or core type 1 in Arka's (2005b) typology). In this analysis, the P-A-Verb in Palu'e is more like the Undergoer or Objective Voice (UV/OV) construction in Indonesian (Arka and Manning 2008) or Balinese (Artawa 1994; Arka 2003a) as exemplified below. This is what Shibatani (2009) calls Patient Focus (PF). UV or PF is not Passive Voice, as the A argument is not an oblique.

(19) a. *Saya (mem-)beli buku itu* (Indonesian)
1s (AV-)buy book that
'I bought the book'

(20) b. *Buku itu saya beli*
book that 1s UV.buy
'The book, I bought'

(21) a. *Tiang numbas bawine* (Balinese)
1 AV.buy pig-DEF
'I bought the pig'

b. *Bawi-ne* *tumbas* *tiang*
pig-DEF UV.buy 1
'The pig, I bought'

Lamaholot also shows Undergoer fronting but differs from Lio, Sikka and Palu'e in that the fronting triggers pronominal agreement or co-indexation.[6] In a canonical SVO clause we find no agreement (18a). However, the bound pronoun *-roʔ* in (22) co-indexes the NP *Kewa*.

(22) a. *Go* *sorõ* *na/Kewa* *muko* (Lewotobi Lamaholot)
1s give 3s/K banana
'I gave him/Kewa a banana'
b. *Kewa* *go* *sorõ-roʔ* *muko*
Kewa 1s give-3s banana
'Kewa, I gave him a banana'

The acceptability of the structure with fronted U NP without *-roʔ* is degraded as seen in (23).

(23) ?* *Kewa,* *go* *sorõ* *muko*
Kewa 1s give banana
'Kewa, I gave him a banana'

However, the bound pronoun can appear without a fronted NP as seen in (24), where *-roʔ* is understood as the Benefactive argument.

(24) *go* *sorõ-roʔ* *muko*
1s *give-3s* *banana*
'I gave him banana.'

In short, the bound pronoun *-roʔ* must be associated with the U of a transitive verb, or the Benefactive of a ditransitive verb. In the ditransitive structure, the Benefactive is the U that gets the privilege to be (co-) indexed on the verb.[7] Importantly, the free co-indexed NP cannot come postverbally.[8] The evidence for this comes from the contrast in (25). Sentence (25a) is unacceptable because of an attempt to co-index *-roʔ* with the following Benefactive *Kewa*. For this sentence to be acceptable, *-roʔ* must be absent as seen in (25).

(25) a. * *Muko* *peʔe* *go* *sorõ-roʔ* *Kewa*
banana that 1s give-3s Kewa
FOR: 'The banana, I gave (to) Kewa'
b. *Muko* *peʔe* go *sorõ* *Kewa*
banana *that* 1s give Kewa
'The banana, I gave (it to) Kewa'

---

6 Some speakers regard pronominal co-indexation as obligatory in some contexts but not in others. Some speakers even regard it as always obligatory. Further research based on texts of real language use is needed to investigate its distribution.

7 This could be taken as evidence for prominence, further discussed in subsection 3.5 below.

8 This constraint must be pragmatically motivated: the bound pronoun on the verb anaphorically refers to an established referent in the discourse. The NP that introduced this referent must therefore precede the bound pronoun.

There seems to be a third person animacy restriction on the type of fronted U NP in Lamaholot, as noted by Nishiyama and Kelen (2007). Thus, the downgraded acceptability of (25a) might also be due to this animacy constraint of *-roʔ*; i.e., the fronted NP muko 'banana' is inanimate and is incompatible with *-roʔ*. However, this constraint does not hold across all Lamaholot dialects, or even among all speakers of the Lewoingu dialect (Nishiyama and Kelen 2007: chap 15). The following from the Nusa Tadon dialect of Lamaholot shows that *-roʔ* can co-index a fronted inanimate NP.

(26)	*Wataʔ*	*ne*	*goʔ*	*kã-roʔ*	(Nusa Tadon Lamaholot)
	corn	that	1s	eat-3s	(Japa 2000: 39)
	'The corn, I ate (it)'				

In addition, there is evidence that the structure with Undergoer fronting in Lamaholot changes the meaning. Nagaya (2009) reports a realis/past vs. irrealis/future contrast shown in (27). This would not be expected if the structure with Undergoer fronting in (27b) were simply a pragmatically marked variant of the structure in (27a).

(27)	a.	*Nius*	*n-enu*	*tuaʔ*	*teʔẽ*	(Lewotobi Lamaholot)
		Nius	3sg-drink	tuak	this.poss	(realis/past)
		'Nius drank this tuak.'				
	b.	*tuaʔ*	*teʔẽ,*	*Nius*	*n-enu*	(irrealis/future)
		tuak	this.poss	Nius	3sg-drink	
		'This tuak, Nius will drink (it).'				(Nagaya 2009)

That a voice/diathesis alternation may involve semantic contrast is common across languages. For example, active-middle opposition correlates with a contrast in reflexivity, and applicative-non applicative opposition correlates with a contrast in affectedness of the applied argument. In a language with more than one passive, different types of passive encode certain semantic contrasts, e.g., volitionality, as in Indonesian *di-* and *ter-* in (28).

(28)	a.	*Orang*	*itu*	*di-tembak*	(Indonesian)
		person	that	PASS-shoot	
		'the person was (deliberately) shot'			
	b.	*Orang*	*itu*	*ter-tembak*	
		person	that	PASS-shoot	
		'The person was (accidentally) shot'.			

Recall that the Palu'e non-active structure is analysed by Donohue (2005) as a passive, because the A is regarded as an Oblique. The equivalent structure in Lamaholot could be analysed in a similar way. However, there are two good reasons for not adopting a passive analysis for the Lamaholot case.

First, as is the case in Palu'e, when a core index analysis is applied to Lamaholot, the A argument in the P-A-Verb structure shows a high core index. On the available descriptions of Lamaholot (Japa 2000; Nishiyama and Kelen 2007; Grangé 2009; Nagaya 2009), core arguments in this language show the following properties: i) verbal prefix agreement (associated with A) or enclitic agreement (associated with P), ii) (core) argument positions (in A-Verb-P structure, A and P are in argument positions, and in P-A-Verb structure, A is analysable as being in argument position [Spec,VP] (Japa 2000) not in adjunct position), iii) typically NPs, v) subcategorised for, and vi) typically obligatory. In addition, Nagaya reports that coordination is associated with A, even in the P-A-Verb structure as seen in the examples (29b) below. Given all these characteristics of A, it can

be concluded that A of P-A-Verb is highly core (a maximum core index of 1.00). Thus, there is no evidence that the coreness of A is significantly downgraded (to oblique status to warrant a passive analysis) when A appears in the alternative structure of P-A-Verb.

(29) a. *na* *bəŋo* *go,* *kədiʔ* *gwali.*
3sg hit 1sg then return
'S/he hit me, and (s/he) returned.'
b. *go* *na* *bəŋo,* *kədiʔ* *gwali.*
1sg 3sg hit then return
'Me, s/he hit, and (s/he) returned.'

Second, from what we currently know of the Austronesian languages in eastern Indonesia (Klamer 2002; Arka 2003c; Himmelmann 2005), features such as grammatical subject/pivot, syntactic control and voice alternations of the kind encountered in the Austronesian languages of western Indonesia such as Indonesian and Balinese are not typical of this area. Lamaholot might be different, but more language-specific evidence is needed to confirm whether fronting of an Undergoer NP in this language may indeed involve some kind of change in grammatical status of the arguments and therefore a change in voice. Nagaya analyzes P of the P-A-Verb structure as a topicalised object. That is, the structural change of P does not involve a voice alternation.[9]

4.3 Intransitive-transitive alternation and applicativisation

In our conception of voice, applicativisation is part of the voice system (cf. Verhaar 1984; Davies 2005). Applicativisation typically adds or promotes a non-actor argument to core status in the argument structure of the base verb. It transitivises an intransitive base, to which the voice alternation that picks up the non-actor as subject can then be applied. Thus, non-active voice constructions with intransitive roots require applicativisation (e.g. in Balinese, Arka 2003: chapter 5). Consider the examples in (30) from Balinese. The verb *edot* 'want' is intransitive; it can take a complement clause with subject control as seen in (30a). However, the complement can be contrastively focused by fronting. The fronted complement grammatically functions as subject. The relativiser *ane* must be used (30b). Importantly, the verb must appear with the applicative suffix *-ang*; otherwise the structures are not acceptable (30c-d).

(30) a. *Cang* *edot* *[makaad* *joh]* (Balinese)
name want go far
'I wanted to go far away'
b. *Makaad* *joh* *ane* *edot-ang* *cang*
go far REL want 1
'Going far away is what I wanted'
c. * *Makaad* *joh* *ane* *cang* *edot*
go far REL 1 want
'Going far away is what I wanted'

[9] Certain tests such as control and clausal adverbial placement tests show that in Nusa Tado Lamaholot the fronted U NP acquires syntactic subject properties (Japa 2000). In this view, the P-A-verb structure is analyzed as having a voice type different from the A-Verb-P structure (Japa 2000). We have not been able to verify this. Whether this is indeed the case across dialects of Lamaholot is a matter for further research.

d.* *Makaad* *joh* *ane* *edot* *cang*
go far REL want 1
'Going far away is what I wanted'

Being of the isolating type, Flores languages have no applicative and no voice morphology. However, the structural effect of the intransitive-transitive alternation, which is expressed by the applicative and voice morphology in Balinese, is observed in Flores languages. As expected, the alternation is expressed by a change in construction type which utilises the voice structure given in (14). Consider the examples from Sikka in (31) below. The verb *deri* 'sit' is intransitive (31a), and cannot be used in an active transitive structure (31b) (cf. the unacceptability of Indonesian (30b) above). However, Sikka allows a transitive alternation with *deri* (31c). Crucially, this structure must have the U argument fronted to the subject position. This is arguably an instance of an Undergoer Voice construction without an applicative affix or any voice morphology on the verb.

(31) a. *Wae* *buang* *ia* ***deri*** *ei* *kadera* (Sikka)
face white that sit P chair
'The pretty girl sat on the chair'

b. * *Wae* *buang* *ia* ***deri*** *kadera*
face white that sit chair
'The pretty girl sat on the/a chair'

c. *Kadera* *ia* *wi* *wae* *buang* *ia* ***deri***
chair that REL face white that sit
'It is that chair which was sat on by the pretty girl'

4.4 Middle

Certain verbs allow an intransitive-transitive alternation which is achieved by placing the same verb form either in an intransitive frame of [NP V] or in a transitive frame of [NP V NP]. An agentive verb such as *zhio*[10] 'bathe' (32) in Rongga and *sebong* in Manus (33) in the intransitive frame gets a middle interpretation whereas in the transitive frame it gets an active interpretation. The same pattern is observed in other Flores languages in central and western Flores such as Ende, Nage/Keo, Waerana and Manggarai.

(32) a. *Ana* *ndau* *zhio* (Rongga)
child that bath
'the child took a bath'
(or the child bathed himself)'

b. *Ine* *zhio* *ana* *ndau*
mother bathe child that
'Mother bathed the child.'

(33) a. *Anak* *koe* *itu* *sebong* (Manus)
child small that bathe
'The child took a bath/bathed himself.'

b. *Ene* *itu* *sebong* *anak* *koe* *itu*
mother that bathe child child that
'The mother bathed the child.'

---

[10] In the orthography adopted by a group of Rongga speakers, *zh* represents the alveolar approximant [ɹ], which is phonemic in Rongga.

Lamaholot utilizes an Undergoer suffix on the verb to exhibit split intransitivity.[11] The patientive intransitive verbs take the undergoer suffix and agentive intransitives do not take the suffix, e.g. *goka-na* 'fall-' vs. *pla'e* 'run'. Some verbs, such as *buka* 'open' and həbo 'bathe' show a three-way contrast, between a suffixed (patientive/middle) intransitive (34a), a bare (middle) intransitive (34b), and an active transitive (34c).

(34) a. *Goe* *həbo=k* (intransitive) (Lamalera Lamaholot)
1sg bathe=1sg
'I took a bath/bathed myself.'

b. *Goe* *həbo* (intransitive)
1sg bathe
'I took a bath/bathed myself.'

c. *Goe* *həbo ana* *nawe* (transitive)
1sg bathe child that
'I bathed the child.'

4.5 Three-place predicates and dative shift alternation

Constructions similar to dative shift are widespread in Flores languages. Example (35), from Rongga, illustrates how the change from transitive to ditransitive occurs with no change in the form of the verb. This kind of transitive-ditransitive alternation generally would be encoded by applicative morphology in Indonesian-type languages.

(35) a. *Ardi* *indi* ***ndoi*** (***pe*** ***ndia*** ***ne*** ***ja'o***) (transitive)
name bring money to here with 1s (Rongga)
'Ardi brought money (here)(for/to me).'

b. *Ardi* *indi* ***jao*** *(ko)* ***ndoi*** (ditransitive)
name bring 1s ko money
'Ardi brought you money'

c.* *Ardi* *indi* *(ko)* ***ndoi*** ***jao***

On semantic grounds, the transitive construction with an undergoer/theme argument (i.e. the (a) sentences above) is more 'basic' than the ditransitive one. The ditransitive one encapsulates a complex event that includes the event conception of the transitive one. For example, the ditransitive structure 'bring (for)' (35b) implies the 'bringing' event of (35a) which extends to a benefactive argument, and this benefactive/goal argument is also given semantic/conceptual/grammatical prominence.

The new grammatical prominence is evidenced by the acquisition of first object structural/grammatical properties. The benefactive argument now must come in the first object position, and the theme argument appears as the second object. Swapping these objects is not permitted; hence the unacceptability of (35c). Note that the former direct object, now demoted to second object, is optionally marked with the particle *ko*, which marks weakened or demoted objects in Rongga.

[11] Verbal affixation in Lamaholot, including the semantics of the intransitive split, is remarkably complex and possibly varies across dialects (Arka 2000; Japa 2000; Nishiyama and Kelen 2007; Grangé 2009). In Dixon's (1994) terminology, the system in Lamaholot allows split-S as well as fluid-S (Arka 2000; Grangé 2009) .

Ditransitives are not generally restricted to those with benefactive/goal arguments. They may also include locatives, as exemplified in (36) with the verb *mula* 'plant', also from Rongga. (36a) is monotransitive; the locative is an adjunct PP and the theme is an object. (36b) is ditransitive; the locative is an object NP and the theme is the second object with *ko*. (36c) is monotransitive; the locative object has a particle *ko* and the theme is downgraded to an oblique, marked by *ne* 'with'. (36d) is monotransitive, like (36c), except that the locative (first) object comes without *ko*. (36d) is ditransitive with the second object without *ko* and its acceptability is degraded.

(36) a. *Ja'o mula (ko) nio (kana) **one uma ja'o*** (transitive)
1s plant ko coconut all in garden 1s (Rongga)
'I planted (all) coconuts in my garden.'

b. *Ja'o mula **uma ja'o** ko nio kana* (ditransitive)
1s plant garden 1s ko coconut all
'I planted the full area of my garden with coconuts'

c. *Ja'o mula **ko uma ja'o** ne nio lepa* (transitive)
1s plant ko garden 1s with coconut finished
'I planted the full area of my garden with coconuts'

d. *Ja'o mula uma ja'o ne nio lepa*. [without *ko* for the first object].

e. ?* *Ja'o mula uma ja'o nio kana.* [without *ko* for the second object]

The constraints on *ko* appear to reflect argument prominence. It has been recognized in the literature that argument relations, whether they are grammatical (subject, object, obliques; core and non-core) or semantic (agent, benefactive, goal, theme, etc.), are structured in terms of prominence or hierarchy (Keenan and Comrie 1977; Bresnan and Kanerva 1989; Bresnan 1995; Manning 1996; Bresnan 2001; Arka 2003a, among others). For example, Object is more prominent than Oblique and in case there are two objects direct/first object is more prominent than secondary object (Keenan and Comrie 1977). Thematically, it has been recognized that beneficiary/goal is more prominent than locative (Bresnan and Kanerva 1989).

Thus, we observed the following constraints on *ko* with objects in Rongga: (i) the object of a monotransitive structure is optionally marked by *ko*; (ii) the first object of a ditransitive structure cannot be weakened (i.e. marked with *ko*) as it must be superior to the second object; (iii) the second object of ditranstive generally takes *ko* as this is a natural way of marking a secondary object; (iv) the second object can be without *ko* so long as the first object is also thematically prominent (i.e. beneficiary/goal). Constraint (iv) accounts for why in the ditransitive structure with a locative first object as in (36b), *ko* is obligatory; otherwise the sentence is not quite acceptable (36d).

Semantically, the alternation also gives rise to a change of meaning. The promoted locative argument in (36), for instance, is understood to be 'fully' affected by the action of planting. As the translation shows, the ditransitive structure (36b) implies the area of the garden is fully covered with coconuts. This implication is not available for the transitive construction (36a).

The possible transitive alternations exemplified above for Rongga can be summarized in (37). (Whether one or two alternations are allowed depends on the verb; hence, it is lexically constrained.)

(37) a. $NP_{A:SUBJ}V$ (*ko*)$NP_{U:OBJ1}$ (PP_{OBL})

b. (i) $NP_{A:SUBJ}V$ $NP_{U(BEN/GOAL/LOC):OBJ1}$ *ko*=$NP_{U:OBJ2}$

(ii) $NP_{A:SUBJ}V$ $NP_{U(BEN/GOAL/LOC):OBJ1}$ (PP_{OBL})

There is also evidence that the argument prominence discussed above places a constraint on argument linking of the subject in transitive sentences. According to this constraint, in a non-symmetrical system, only the most prominent core argument participant can be subject. Thus, the A argument of a transitive verb is by default selected as Subject. In passive, where A is not available or demoted, the basic undergoer, or the promoted undergoer, is available as the most prominent argument and is selected as Subject. The demoted undergoer of a ditransitive structure, however, cannot be selected as subject.

Rongga shows a non-symmetrical system. The asymmetry in subject selection of non-actor arguments due to the prominence constraint is observed in this language. Thus, given a monotransitive of the type (37a) where the theme is OBJ, exemplified by (36a), the theme is Subject in the passive:

(38) a. *Nio* *ndau* *mula* *one* *uma* *ne* *ja'o* *(Rongga)*
Coconut that plant in garden by 1s
'The coconuts were planted in the garden by me'

However, in a ditransitive structure of the type (37b) where the theme is the second OBJ, exemplified by (36b), the theme cannot be the passive subject:

(39) * *Nio* *ndau* *mula* *uma* *ja'o* *ne* *ja'o*
coconut that plant garden 1s by 1s
'The coconuts were planted in the garden by me'

For the ditransitive verb, the first OBJ of the active structure alternates with passive subject. Thus, being the first OBJ of the ditransitive structure in (36b), a locative argument may become a passive subject, as seen in (40). Note that the theme appears with *ko*, without which the acceptability of this sentence is degraded.

(40) *Uma* *ja'o* *mula* *lepa* *ko* *nio* *ne* *ja'o*
garden 1s plant finish ko coconut by 1s
'The my garden was all covered with coconuts that I planted'

4.6 Causative and anticausative alternations

In the definition of voice adopted here, causative and anticausative alternations are also part of the voice system. These constructions express different conceptions of event development. For languages with voice morphology, there is often a restriction on the formation of certain voice constructions in relation to (anti)causativisation. For instance, in Balinese, only patientive or stative intrasitive bases which are expressed by bare forms (i.e. forms also used for undergoer voice) can undergo morphological causativisation with –*ang*. Thus *ulung* 'UV.fall' can take the -*ang* suffix to form *ulung-ang* 'fall-CAUS', but

ngeling 'AV.cry' cannot (**ngelingang* 'AV.cry-CAUS').[12] Also, only causative transitive verbs of the form of *AV.V-ang* can undergo anticausativisation in Balinese (Arka 2003b). Thus, *nyenik-ang* 'AV.small-CAUS' or 'make something small' (41a) can be used to mean 'become small' (41b) when the agent is unexpressed and the patient is in preverbal position. While the verb form remains formally causative, the meaning of the anticausative structure is distinct. The anticausative structure (41b) expresses the effect or the inchoative aspect/function of the category/concept of causation expressed by (41a).

(41) a. *Nyoman* ***nyenikang*** *api-ne*
N-cenik-ang
Name AV-small-CAUS fire-DEF
'Nyoman made the fire smaller'

b. *Api-ne* ***nyenikang***
N-cenik-ang
fire-DEF AV-small-CAUS
'The fire became small'

Similar instances of (anti)causative alternation are found in Flores languages. The following examples are from Rongga (Arka et al, 2007). The same verb form *nggoli* 'roll' expresses a different event conception depending on the construction it is used in. In the intransitive construction of [NP V], it may be conceptualised as patientive or passive as in (42a), or self-instigated or middle (42b). The animacy of the sole NP and/or the context is critical. In the transitive construction of [NP V NP] (42c), it is conceptualised as an active causative action. Given the absence of any morphology, it is unclear which structure could be analysed as basic. If *nggoli* is analysed as basically transitive, then intransitive structures (42a) and (42b) could be regarded as the anticausative/passive and middle alternations respectively. Alternatively, if *nggoli* is basically intransitive, then the transitive counterpart is its lexical causative counterpart.[13]

(42) a. *Watu ndau* ***nggoli*** (Rongga)
stone that roll
'The stone rolled (or was rolled)'

b. *Jao* ***nggoli***
1s roll
'I rolled (myself)'

c. *Ja'o* ***nggoli*** *watu* *ndau*
1s roll stone that
'I rolled the stone or I made the stone roll'

---

[12] Note that ngelingang is an acceptable form for an applicative verb as in the example below.
Ia *ngelingang* *meme-ne*
3SG AV.cry-APLL mother-3POSS
'S/he cried for his/her mother.'
*'S/he made his/her mother cry.'

[13] The question of 'basic-ness' is perhaps irrelevant if one assumes that this verb, like many others, has multiple subcategorisations at the lexical level, each of which is activated by a specific construction in use. We will not pursue this issue any further in this paper.

5 Concluding remarks

In this short paper we have presented data from the Austronesian languages of Flores Indonesia. In this last section, we provide a summary, followed by discussion of issues arising from the present study.

As noted at the beginning of this paper, typologically the salient feature of these languages is their isolating type. At some point during their history, all of these languages have lost the rich verbal voice morphology found in western Austronesian languages and reconstructed for Proto-Austronesian. To some extent, the loss of verbal morphology is compensated for by the development of bound pronominal forms in Lamaholot in the eastern part of Flores and pronominal clitics in Manggarai in the western part.

These bound forms, however, do not function in the same way in both languages. In Manggarai, the alternation of pronominal co-indexing is a diagnostic test for a change in grammatical relations, and therefore also clear diagnostic evidence for voice alternation. In Lamaholot, the situation is slightly different. The undergoer bound pronominal on the verb provides coding for split intransitivity, which is done by verbal voice morphology in other Austronesian languages such as Balinese. It is unclear, however, whether the pronominal suffix on the transitive verb which is often triggered by undergoer fronting encodes an undergoer voice alternation.

The other languages, mainly in central Flores, are at the far end of the isolating continuum.

However, the development of Flores languages into the isolating type with the loss of verbal voice morphology is not followed by a (total) loss of a voice system. Different kinds of voice (active voice, passive voice, undergoer voice, middle voice, (anti)causative, and dative shift) are attested in Flores languages. The distinctions are mainly coded constructionally. While further research on the semantics and pragmatics of voice in Flores languages is needed, our preliminary research has revealed the subtlety and richness of the meaning contrasts encoded by constructionally-coded voice distinctions in these languages.

In contrast with wesern Austronesian languages, the loss of verbal voice morphology in Flores languages leads to a less systematic contrast among voice types, and a reduced number of voice distinctions in the individual languages on this island. Grammatical relations and voice alternations are arguably harder to tease apart in these languages than in Indonesian-type or Philippine-type languages.

In section 3 we provided a broad conceptually-based definition of voice. Using this definition, the richness of alternative argument realisations typically featuring voice alternations can be compared across the languages of Nusa Tenggara. The alternations may employ reflexes of Proto-Austronesian morphological coding, other morphological coding, or in the case of the languages of central Flores, no verbal coding at all.

For example, cases of split-intransitivity in Lamaholot and Balinese are both analysed as part of the system of voice oppositions. The two languages differ, however, in their coding: Balinese marks the agentive intransitive employing the AV (Actor voice)/AF (Actor Focus) nasal prefix marking (*N*-), a marker traced back to PMP **maN*- or PAN *-*um*-, which signalled 'active' in the old systems (Blust 2002; Ross 2002 and the references therein). Lamaholot, however, marks the patientive intransitive, employing pronominal marking coding, typically used for the U argument in the transitive verb. The difference appears to reflect different historical development. Other Flores languages (Sikka and Palu'e in eastern Flores and Manggarai in western Flores) also have bound pronominal forms, possibly used on verbs. In contrast with Lamaholot, however, such pronominal coding is not employed to express split intransitivity. Languages of central Flores such as Rongga and Ngadha are extremely isolating. There is no bound pronominal marking in

these languages. And there is no intransitive split phenomenon in these languages expressed by other means either.

Other phenomena such as dative shift and (anti-)causativisation also fall into voice in our definition. We have observed that, in the absence of any verbal morphology, voice distinctions rely solely on analytic coding. The verb form remains the same in different types of voice. A passive construction, for example, may be simply marked on the A argument by a preposition as in Manggarai and Rongga (section 4.1).

Our findings clearly show that voice without voice morphology in the Austronesian languages of Flores is attested. A detailed study of passives without passive morphology in Manggarai is given in Arka and Kosmas (Arka and Kosmas 2005).

A number of issues arise from our study. The first is historical. The precise answer to the question of historical development of the complete disappearance of Austronesian voice morphology in Flores languages is a matter for further research.

The remaining issues are typological/theoretical. One important question is whether the 'non-active' constructions found in Flores languages are best analysed as passive, as inverse, as undergoer voice, or as undergoer focus. There has been some disagreement whether Focus systems in the Philippine/Formosan languages are Voice systems as found in languages like English, and if not to what extent they are different. Typologists such as Foley (2008), Himmelmann (2002a; 2002b; 2005), and Ross (2002; 2006) consider the Focus System a voice system. Formal linguistic analyses of different theoretical persuasions also treat the Focus system as a voice system.

Shibatani (2009), however, argues that the Focus system is not the same as a voice system. The focus alternation does not change clause-level grammatical relations whereas the voice system does. Thus, different Focus structures may belong to the same voice type. In his analysis, AF (Actor Focus) shows the alignment of subject/actor as (grammatical) Topic whereas PF (Patient Focus) shows the alignment of patient/object as Topic. In PF, the actor is in his analysis still subject. AF and PF are therefore both active voice.

It should be noted that Shibatani's notion of subject is not exactly the same as surface SUBJ (or Pivot) as used in this paper (see footnote 4). Rather, it is a logical subject (or Actor) that is a core argument (equivalent to what Manning (1996) and Arka (2003a) call argument-structure subject (a-subject).

It should also be noted that Shibatani's analysis is in line with other earlier analyses which conclude that UV is not passive (Foley 1998; Arka 2003a; Katagiri 2005; Foley 2008, among others). The term active voice should not be taken as the exact equivalent of actor voice (or vice versa). In this paper, we adopt the analysis that Austronesian Focus system is part of wider system of voice. It is a voice system because it regulates the mapping of arguments onto surface grammatical relations.

Returning to Flores languages, it has been argued that the languages on this island such as Manggarai (Arka and Kosmas 2005) and Palu'e (Donohue 2005) show cases of passive voice without passive morphology. The relevant examples from Manggarai are repeated below:

(43) a. *Aku* *cero* *latung=k* (Manggarai)
1s fry corn=1s
'I fry/am frying corn'

b. *Latung* *hitu* *cero* *l=aku=i*
corn that fry by=1s=3s
'The corn is (being) fried by me'

The alternation in (43) shows a change of the status of the actor (*aku*): it is subject (core) argument in (42a) but an oblique in (42b). Clear evidence for its oblique status is

observed from its structural position and PP coding. More convincing evidence, taking into account these structural and marking properties, comes from the investigation of its core index (i.e. degree of coreness).

Core properties in Manggarai show at least the following six properties: i) subcategorised for, ii) obligatory, iii) NPs rather than PPs (which are obliques), iv) structurally in core (S/A/P) argument positions, v) possibly realised as core clitics, and vi) binder of a reflexive. The A *aku* in (43a) (active) shows a prototypical core with a core index of 1.00 (i.e. satisfying all these six core properties), in addition to satisfying subject properties in Manggarai such as subject position, the only argument targeted in control in purposive and participial/nominalised (complement) clauses and relativisation (see Arka and Kosmas (2005) for details). Its counterpart in (43b), *l=aku*, shows a very low degree of coreness, a core index of 0.33 (i.e. a satisfying at best the first two out of six).

It should noted that the oblique A in (43) does not appear to be a typical oblique as it is not optional. However, a transitive verb in Manggarai may have a clitic, e.g. *=d* as in (44), to encode a passive structure. In this example, the Argument marked by *le (l=ise)* is optional. The optionality makes this prepositionally-marked actor a typical oblique.

(44) *Poli=s* *emi=d* *(l=ise)* *bao* *surak* *situ.*
already=3p take=POSS by=3p just.now letter that.p
'The letters have been taken (by them) (just now)'
(Lit. The letters were already, the taking (of them) just now (by them).'

In short, we have good evidence that Manggarai shows a passive alternation, possibly without voice morphology, and that the actor of the passive structure has a low degree of coreness (i.e. an Oblique).

However, the question is whether there is also evidence for a UV/PF construction in Manggarai. That is, do we have a structure where the undergoer is syntactic Subject/Pivot and the actor remains as a core. This is a difficult question given our current understanding of Manggarai grammar. However, a structure of the type in (45), quoted from the story Tura Kambu Lawang, Manggarai Timur, could be a UV/PF candidate. What is special about this structure is the expression of the Actor: it is realised as a sentence-initial NP but is also anaphorically cross-referenced by the PP *l=iha (i.e.,* obliquely marked). This A argument is arguably more core-like than the A counterpart in (43) above, where it is simply backgrounded as a PP. On this analysis, and on the analysis that the undergoer Empo Ete is a postposed surface SUBJ (or Topic in Shibatani's terminology), then sentence (45) could be an instance of UV/PF in Manggarai. More research is needed to investigate how widespread this kind of construction is.

(45) *Hi* *Kambu Lawang* *hitu* *kawe-kawe* *l=iha* *Empo Ete*
ART *Kambu Lawang* that search-REDUP by=3SG *Empo Ete*
'Kambu Lawang was looking for Empu Ete'

Finally, the notion of voice without voice morphology (or Focus without Focus morphology) as discussed in this paper is not trivial theoretically. It may pose a challenge to certain frameworks. This has been discussed in Arka and Kosmas (2005); Richards (2006) and Arka (2009), and will not be repeated here. Among the issues are the nature of the different structures and the processes by which they are derived; e.g. which one is basic and which one is derived, whether the derivation is purely lexical or syntactic, and the precise mechanism of derivation, or whether the notion of structural derivation itself is tenable. The precise and best analysis of voice without voice morphology is a matter of debate.

References

Arka, I Wayan. 2000. Beberapa aspek terpilah pada bahasa-bahasa Nusantara: sebuah analisis leksikal fungsional (Aspect of intransitivity split in the languages of Nusantara: a Lexical-functional approach). In *Kajian serba linguistik*, edited by B. Kaswanti Purwo, 423-456. Jakarta: Unika Atmajaya.

———. 2003a. *Balinese morphosyntax: a lexical-functional approach*. Canberra: Pacific Linguistics.

———. 2003b. Causative and anticausative N- -ang constructions in Balinese. In *Cakrawala baru: liber amicorum untuk Prof. Soenjono Dardjowidjojo*, edited by K. E. Sukamto, 3-17. Jakarta: Yayasan Obor Indonesia.

———. 2003c. Voice systems in the Austronesian languages of Nusantara: Typology, symmetricality and Undergoer orientation. *Linguistik Indonesia* 21 (1):113-139.

———. 2005a. Challenges and prospect of maintaining Rongga: a preliminary ethnographic report. In *Proceedings of the 2004 Conference of the Australian Linguistics Society*, edited by I. Mushin. http://deneb.library.usyd.edu.au:8080/handle/2123/138.

———. 2005b. The core-oblique distinction and core index in some Austronesian languages of Indonesia Paper read at Paper presented at International ALT VI (Association of Linguistic Typology conference, Padang Indonesia, July 2005.

———. 2009. On the zero (voice) prefix and bare verbs in Austronesian languages of Nusa Tenggara, Indonesia. In *Discovering history through language: papers in honour of Malcolm Ross* edited by B. Evans, 247-270. Canberra: Pacific Linguistics.

Arka, I Wayan, and Jeladu Kosmas. 2005. Passive without passive morphology? Evidence from Manggarai. In *The many faces of Austronesian voice systems: some new empirical studies*, edited by I. W. Arka and M. D. Ross, 87-117. . Canberra: Pacific Linguistics.

Arka, I Wayan, Jeladu Kosmas, and I Nyoman Suparsa. 2007. *Bahasa Rongga: tatabahasa acuan ringkas*. Canberra: Linguistics Department, RSPAS, ANU.

Arka, I Wayan, and Christopher Manning. 2008. Voice and grammatical relations in Indonesian: a new perspective. In *Voice and grammatical relations in Austronesian Languages*, edited by P. K. Austin and S. Musgrave, 45-69. Stanford: CSLI.

Artawa, I Ketut. 1994. Ergativity and Balinese syntax. PhD thesis, La Trobe University.

Baird, Louise 2002. A grammar of Keo: an Austronesian language of East Nusantara. PhD thesis, Linguistics Department, ANU, Canberra.

Blust, Robert. 1993. Central and central-eastern Malayo-Polynesian. *Oceanic Linguistics* 32 (2):241-293.

———. 2002. Notes on the history of 'focus' in Austronesian languages. In *The history and typology of western Austronesian voice systems*, edited by W. a. Ross, 63-80. Canberra: Pacific Linguistics.

Bresnan, Joan. 1995. Linear order, syntactic rank, and empty categories: on weak crossover. In *Formal issues in Lexical-Functional Grammar*, edited by R. M. K. M. Dalrymple, J.T. Maxwell III, and A. Zaenen, 241-74. Stanford.: CSLI Publications.

———. 2001. *Lexical functional syntax*. London: Blackwell.

Bresnan, Joan, and Jonni M. Kanerva. 1989. Locative inversion in Chichewa: a case study of factorization in grammar. *Linguistic Inquiry* 20:1-50.

Davies, William D. 2005. The richness of Madurese voice. In *Aspects of voice and marking in Austronesian languages*, edited by I. W. Arka and M. D. Ross, 197-220. Canberra: Pacific Linguistics.

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

———. 2010. *Basic Linguistic Theory*. Oxford: Oxford University Press.

Donohue, Mark. 2005. The Palu'e passive: from pragmatic construction to grammatical device. In *The many faces of Austronesian voice systems: some new empirical studies*, edited by I. W. Arka and M. D. Ross, 59-85. Canberra: Pacific Linguistics.

Dyen, Isidore. 1965. A lexicostatistical classification of the Austronesian languages. *International Journal of American Linguistics* (Memoir 19).

Foley, William A. 2008. The Place of Philippine languages in a typology of voice systems. In *Voice and grammatical relations in Austronesian languages*, edited by P. K. Austin and S. Musgrave, 22-44. Stanford: CLSI.

Foley, William A. 1998. Symmetrical Voice Systems and Precategoriality in Philippine Languages. Paper read at Workshop on voice and grammatical functions in Austronesian languages, LFG98, Sydney.

Foley, William A., and Robert D. Van Valin. 1984. *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.

Grangé, Philippe. 2009. Split intransitivity in Lamaholot (east Flores, Indonesia). In *11-ICAL*. Aussois, France, 22-26 June 2009.

Himmelmann, N.P. 2002a. Voice in two northern Sulawesi languages. In *The history and typology of western Austronesian voice systems*, edited by F. Wouk and M. D. Ross, 123-142. Canberra: Pacific Linguistics.

———. 2002b. Voice in western Austronesian: an update. In *The history and typology of western Austronesian voice systems*, edited by F. Wouk and M. D. Ross, 7-15. Canberra: Pacific Linguistics.

Himmelmann, Nikolaus P. 2005. The Austronesian languages of Asia and Madagascar: typological characteristics, edited by K. A. Adelaar and N. P. Himmelmann, 110-181. London: Routledge.

Japa, I W. 2000. Properti Argumen Inti, Interpretasi Tipologis dan Struktur Kausatif Bahasa Lamaholot Dialek Nusa Tadon, S2 Linguistik, Universitas Udayana, Denpasar.

Katagiri, Masumi. 2005. Voice, ergativity and transitivity in Tagalog and other Philippine languages: a typological perspective. In *The many faces of Austronesian voice systems: some new empirical studies*, edited by I. W. Arka and M. D. Ross. Canberra: Pacific Linguistics.

Keenan, Edward L., and Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8:63-99.

Klamer, Marian. 2002. Typical features of Austronesian languages in Central/Eastern Indonesia. *Oceanic Linguistics* (41):363-384.

Manning, Christopher D. 1996. *Ergativity: argument structure and grammatical relations*. Stanford: CSLI.

Nagaya, Naonori. 2009. Subject and topic in Lamaholot, eastern Flores. In *11-ICAL (International conference on Austronesian linguistics), Aussois, France, 22-25 June 2009*.

Nishiyama, Kunio, and Herman Kelen. 2007. A grammar of Lamaholot, eastern Indonesia: the morphology and syntax of the Lewoingu dialect. Muenchen: Lincom Europa.

Richards, Norvin W. 2006. The many faces of Austronesian voice systems: some new emprical studies. Book Review. *Oceanic Linguistics* 45 (2):501-504.

Ross, Malcolm D. 1988. *Proto Oceanic and the Austronesian languages of western Melanesia*. Canberra: Pacific Linguistics.

———. 2002. The history and transitivity of Western Austronesian voice and voice marking. In *The history and typology of Western Austronesian Voice Systems*, edited by F. Wouk and M. D. Ross, 17-62. Canberra: Pacific Linguistics.

———. 2006. The argument structure of undergoer voice clauses in Formosan and other Philippine-type languages. In *AFLA 13*. Taiwan.

Sawardi, F.X. 2000. Argumen kompleks bahasa Lio (Complex arguments in Lio). MA thesis, Linguistics, Udayana University, Denpasar.

Sedeng, I N. 2000. Kalimat Kompleks dan Relasi Gramatikal Bahasa Sikka. Master Thesis, Universitas Udayana, Denpasar.

Shibatani, Masayoshi. 2005. The attrition of the Austronesian focus system. Paper read at Proceedings of the Taiwan-Japan Joint Workshop on Austronesian Languages. June 23-24, 2005. Taipei.

———. 2006. On the conceptual framework for voice phenomena. *Linguistics* 44.

———. 2007. What is voice?: Handout, Rice University.

———. 2009. Focus constructions without focus morphology in the AN languages of Nusa Tenggara. In *11-ICAL (International Conference on Austronesian Linguistics), Aussois France 22-25 June 2005*.

SIL International, Indonesia Branch. 2001. *Languages of Indonesia*. Jakarta: SIL International, Indonesia Branch.

Van Valin, Robert D. , and Randy J. LaPolla. 1997. *Syntax. Structure, meaning and function*. Cambridge: Cambridge University Press.

Verhaar, John W. M. 1984. Affixation in Contemporary Indonesian. In *Towards a Description of Contemporary Indonesian: Part I*, edited by K. Purwo, 1-26. Jakarta: NUSA , Universitas Atma Jaya.

Wouk, Fay. 2002. Voice in the languages of Nusa Tenggara Barat. In *The history and typology of western Austronesian voice systems*, edited by F. Wouk and M. D. Ross, 285-309. Canberra: Pacific Linguistics.

18 *Maka and the left periphery of Classical Malay*

IKA NURHAYANI

1 Introduction

Classical Malay (CM) has a particle in the left periphery, *maka,* which has resisted previous analysis. *Maka* has been analyzed as (i) the final element of the precore of a clause (Ajamiseba 1983), (ii) a linker indicating temporal or logical sequencing (Cumming, 1991) and (iii) a clause boundary marker (Muller-Gotama 1996). It can also be argued that the highly productive *maka* is a coordinating conjunction commonly found in the beginning of most sentences in oral literature. In this paper, I will show that *maka*, as well as other left periphery particles in CM, finds a natural interpretation in the analysis of clausal structure proposed in Rizzi (1997). Rizzi (1997) proposes that the structure of complementizer phrase (CP) is split into several discourse-related categories arranged in a hierarchical fashion. The split CP consists of *Force, Topic, Focus* and *Finiteness* heads. The Force head encodes the illocutionary force of the sentence as a question, a declarative, an exclamative, a relative, a comparative, or an adjunct of a certain kind, and can be selected by a higher selector. The head of the Force phrase is sometimes expressed by overt morphological encoding (Rizzi, 1997: 283).

I base my argument on the fact that *maka* can only appear in the beginning of a declarative clause with a topic or a focus.

(1) *Maka Paya Tun Kerub Maharajana pun ber-anak*
PART Paya Tun Kerub Mahajarana PART INTR-child
'Then Paya Tun Kerub Maharajana had a child'
(Teeuw and Wyatt 1970: 68)

(1b) *Maka anak itu lah yang di-namakan bangsa Megat*
PART boy that PART REL PASS-name people Megat
'That boy was known as the Megat people.'
(Adat Raja Melayu C57:4)

(1c) **Maka baginda kah yang datang?*
PART majesty PART REL come
'Was it his majesty who came?'

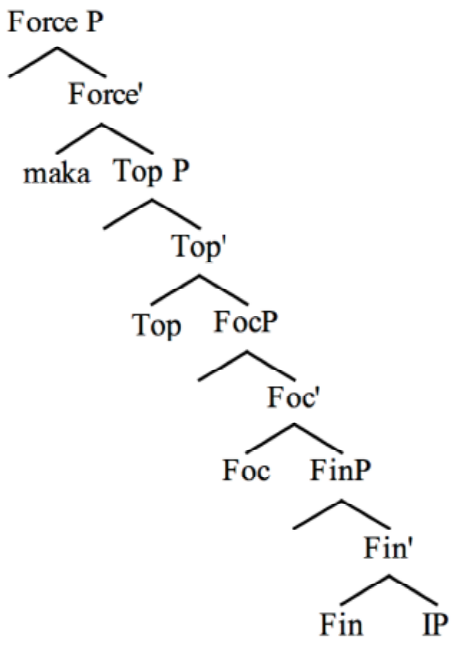
In (1a), *maka* heads a declarative clause containing a topicalized *NP Paya Tun Kerub Maharajana* marked with *pun,* while in (1b) it heads a declarative clause with a focalized noun phrase (NP), *anak itu* 'that boy'. However, the presence of *maka* in front of an

Asia-Pacific Linguistics, 2014.

interrogative sentence marked with a focus question particle *kah* is hypothetically ungrammatical in (1c) since the data from CM literature shows that the particle cannot occur in a declarative sentence.

Hence, I argue that *maka* represents overt morphological encoding on the Force head, which expresses that a sentence is declarative.

(2) The CP projection of *maka*



Moreover, *maka* poses a challenge to the split CP approach, since certain sentences appear to have two overt Force markers. For instance, sentence (3), has two *maka*.

(3) *Maka* *daripada* *geram* *Laksamana itu*
PART because anger admiral that

maka *balai* *bendahara* *pun* *bergoncang*
PART house treasurer PART shake
'The office of the treasurer was shaken because of the anger of the admiral.'
(Hang Tuah, 310: 31)

The goal of this paper is to prove that Rizzi's (1997) split CP approach provides the best framework to account for *maka* and similar left periphery particles in CM. I conclude by discussing the syntax of the double *maka* construction.

This paper is organized as follows. Following this introduction, which is section 1, section 2 provides background on CM, section 3 surveys previous analyses of *maka*, section 4 discusses the analysis of *maka* under the split CP approach, and section 5 concludes the discussion. In addition, the data is taken from CM tales: The Story of Patani (Hikayat Patani), Hikayat Hang Tuah, Cerita Kutai, Adat Raja Melayu, and Hikayat Abdullah bin Abdul Kadir.

2 Background on Classical Malay

2.1 The history of Malay

Malay is the name given to a range of language varieties, which have been spoken over a wide geographical area in the coastal areas of the Malay peninsula and Kalimantan, in South and Southeast Sumatra and in nearly all major trade centers of the Indonesian archipelago (Cumming, 1991: 8). The oldest inscription in Malay, found in Southern Sumatra, was dated from the second half of seventh century (Cumming, 1991: 10). CM is the name of the literary language of the court of the Malacca Empire, dated from the 17th to the 19th century, whose capital was in the Southern part of the Malay peninsula (Cumming, 1991: 1). It is the historical precursor of modern Malay and Indonesian (Cumming, 1991: 11). The existing text of CM includes legal texts, poetry and a prose genre known as *hikayat*. Most linguistic research on the language has concentrated on the latter narratives. Although *hikayat* are in written language, they were most likely recited orally in front of an audience (Muller-Gotama 1996:1).

2.2 The basic word order of Classical Malay

The word order of CM has been a subject of an ongoing debate. As Cumming (1991: 151) points out, 'CM is one of those languages which bring into question the validity of the constituent order typologies themselves; the most frequent order varies wildly according to clause type, and this variation itself can be seen deriving from the functional differences between the orders'. However, Cumming (1991) claims that CM basic word order is VSO. She further states that CM is verb-initial based on her observation that elements with *pun* (see sub-section 2.3) have a status similar to the left-dislocated NP in Italian or English, and that they should be considered as external to the clause. She observes that 80% of CM clauses exhibit a verb-initial word order and therefore concludes that CM has a dominant verb-initial order, and an alternate verb-medial order (Cumming 1991: 152-153).

As far as I know, Cumming's (1991) claim about CM word order has never been supported by other studies of CM syntax. However, it should be noted that current Malay dialects considered more conservative in terms of preserving older forms of Malay, such as Brunei Malay (BM), have a VSO basic word order, as pointed out by Poedjosoedarmo (1995: 227).

(4) *Bakaraja ku di malam sama ading-ku di yumah*
Work I last night with brother-my at house
' I worked with my brother last night at home'
(Poedjosoedarmo 1995: 231)

It can be assumed that BM reflects an intermediate form between CM and modern Malay (MM). In contrast with BM, MM, as shown in Indonesian and Malaysian, has a fixed SVO word order.

(5) *Aku bekerja bersama adik-ku di rumah tadi malam*
I work with brother-my at house last night
'I worked with my brother at home last night'
(Poedjosoedarmo 1995: 231)

By assuming that BM reflects an intermediate stage between CM and MM, we can speculate that CM basic word order is closer to VSO since usually basic word order changes from VSO to SVO rather than the opposite. Therefore, in this paper, I adopt Cumming's (1991) claim about the basic word order of Malay. Moreover, I further posit that CM has the possibility to have V2, or V3 patterns where one or more elements can precede the verb. These elements must be focused or topicalized with *pun* or *lah*. These particles will be discussed further in 2.3.

2.3 The other particles of Classical Malay: *pun* and *lah*

Maka is one of several left-periphery markers in CM. Other important markers are *pun* and *lah*. Adjamiseba (1983) claims that *pun* marks topics, which represent old information and *lah* marks new information. *Pun* is considered external to the clause in Adjamiseba's (1983) framework. This tradition is followed by Cumming (1991) and later by Muller-Gotama (1998). Furthermore, Cumming (1991) states that the *pun*-marked NP is similar in status to a left-dislocated NP in languages like English or Italian.

With the VSO basic word order, the NP is topicalized with *pun* in front of the verb. The topicalized elements are always a subject NP in an active or a passive sentence.

(6) *Maka baginda pun keluar lah ke balairung.*
PART his majesty PART come out PART to big hall
'His majesty came out to the big hall'

(7) *Maka anak raja itu pun di-bawanya kembali ke rumah-nya*
PART child king that PART PASS-bring back to house- HIS
'They brought the child's king back to their house'
(Cumming 1991: 150)

In (6), the topicalized NP *Paya Tun Kerub Maharajana* is the subject of an active sentence, while in (7), the topicalized NP *anak raja itu* 'that king's child' is the subject of a passive sentence. In addition, it can be observed that only a definite subject NP can be topicalized while an indefinite subject has to occur in existential construction.

(8) *Yang diibaratkan jika ada lah seseorang misalnya itu*
REL supposed if exist PART someone as example that

maka beroleh putera dua orang
PART have someone two people
'Supposedly, if there is one that has two sons…'
(Adat Raja Melayu C56:17)

It should also be noted that particle *pun* is only compatible for declarative clause. For a question, the particle *kah* is used instead.

Other elements besides the subject NP, such as time adverbials, can also be topicalized. However, unlike the subject NP, the time adverbial can occur internally without *pun*.

(9) *Hatta* *antara setahun lamanya*
Meanwhile round a year long

maka *raja itu* *pun* *sakit pula*
PART king that PART sick also
'After around a year, the king fell ill also'
(Teeuw and Wyatt 1991: 73).

In (9), the time adverbial *antara setahun lamanya* 'after around a year' is topicalized without *pun* and followed by the second topic *raja itu* 'that king' marked with *pun*.

Now I turn the discussion to the second particle in CM, the particle *lah,* which serves to focalize an NP in (1b) or a verb in (10).

(10) *Maka* *baginda* *pun* *turun* *lah* *dari atas* *gajah* *nya*
PART majesty PART descend PART from up elephantHIS
'Then his majesty descended from his elephant.'
(Muller-Gotama 1995: 148)

In (10), the particle *lah* marks the focalized verb *keluar* 'come out'.

Moreover, the topic and the focus can coexist in one sentence. The focus is always located after the topic.

(11) *Maka* *keluar* *lah* *api* *me-mancar-mancar*
PART come out PART fire TRANS-spout-REDUP
'Flames came pouring out'
(Cumming 1991: 74)

As seen in (11), the topic *baginda* 'majesty' marked with *pun* precedes the focus *turun* 'descend' marked with *lah.*

In addition to its function as a focus particle, *lah* can also mark the imperative.

(12) *Bunuh* *lah* *aku* *dahulu*
Kill PART I first
'Kill me first!'
(Teeuw and Wyatt 1970: 96)

In sum, the word order constraint in the left periphery of CM is as follows:

(13) (*maka) - topic 1 - maka - topic 2 - pun -focus -lah*

3 Previous Analyses on Maka

There have been three previous claims on the identity of *maka* by Ajamiseba (1983), Cumming (1991) and Muller-Gotama (1996). Ajamiseba (1983) divides Malay sentences into precore, core and elaboration and *maka* is a final element of the precore.

(14) *Setelah* *baginda* *mendengar* *sembah* *orang* *itu*
After his majesty hear homage person that
PRECORE

maka *baginda* *pun* *berangkat berjalan* *kepada* *tempat* *itu*
PART his majesty PART depart walk to place that
CORE ELABORATION
'After he heard the man's report, he departed for that place'
(Muller-Gotama 1991: 91)

In (14), Adjamiseba (1983) claims that *maka* is the final element of the precore *setelah baginda mendengar sembah orang itu* 'after his majesty heard the man's report' which precedes the particle. However, it can be observed that *maka* forms a constituent with the core *bagindapun berangkat berjalan kepada tempat itu* 'he departed from the place' following the particle, and not with the precore.

The second claim on *maka* is made by Cumming (1991). She identifies *maka* as a linker indicating temporal or logical sequencing.

(15) *Setelah* *Raja Gongga Syah* *mendengar* *Raja Suran* *datang* *itu*
After king Gongga Syah hear king Suran come that

maka *baginda* *pun* *menyuruh menghimpunkan* *rakyat-nya*
PART his majesty PART order assemble people-HIS
'After King Gongga Syah Johan heard about the arrival of King Suran, his majesty then ordered his people to assemble.'
(Cumming 1991: 126).

According to Cumming (1991), in (15), the linker *maka* is used to link *King Suran*'s previous action to *King Gonggang Syah Johan*'s subsequent one (1991: 126). However, she misses the fact that *maka* can manifest twice in a main clause as seen in (3). It is possible to argue that the first *maka* indicates a logical sequencing with the previous sentence. However, the second *maka* inside the clause surely does not indicate any logical or temporal sequencing with another sentence or clause.

Moreover, *maka* can co-exist with other 'linker' in the clause.

(16) *Syahadan* *maka* *heran lah* *hati raja Iskandar* *melihat*
LINK PART startled PART heart king Iskandar see

akan *rupa* *puteri Shahru l-Bariyah* *itu*
for look princess Shahru l-Bariyah that
'King Iskandar was amazed to behold the countenance of Princess Shahrul-Bariyah'
(Cumming 1991: 92).

In (16), *maka* is preceded by *syahadan*, which Cumming (1991) claims as a linker. If *maka* is also a linker, it immediately poses a question on why a clause needs two adjacent linkers.

Next, Muller-Gotama (1996) describes *maka* as a clause-initial punctuation word which indicates a sentence or clause boundary.

(17) *Maka* *setelah* *keesokan hari-nya* *maka* *jaring dan jerat pun*
PART after tomorrow day-DEF PART net and trapPART

di-tahan *orang* *lah*
PASS-set person PART
'The following morning, nets and snares were set'
(Muller-Gotama, 1996: 146)

In (17), based on Muller-Gotama (1996), the punctuation word *maka* is followed by the time adverbial *setelah keesokan harinya* 'the following morning'. We can see in (17) that the clause has a second instance of *maka* before the topic NP *jaring and jerat pun* 'the nets and snares'. Muller-Gotama (1996) claims that this second *maka* introduces the topic constituent. It can be observed that there is a significant problem in Muller-Gotama's (1996) framework: not only can *maka* occur twice in a clause, but the second *maka* has a totally different function from the first one. If *maka* is a punctuation word indicating a clause boundary, why can it also serve to introduce a topic constituent inside the clause? Furthermore, why does *jaring dan jerat* 'nets and snares' need *maka* to mark it as a topic constituent, given it is already marked with the particle *pun*?

Lastly, I discuss the possibility that *maka* is a coordinating conjunction commonly found in oral literature. Walter Ong (1982) observes that oral literature has an additive nature and therefore is marked with succession of coordinating conjunctions. *Maka* appears to have similar additive function in CM.

However, the data shows that *maka* has more specific distribution than a typical coordinating conjunction. As mentioned in section 1, *maka* can only be found in front of a declarative sentence. If *maka* is an ordinary coordinating conjunction like *then* and *and*, we can expect to find *maka* in front of any types of sentences, such as in interrogative or comparative. Nevertheless, this is not the case, as seen in (1c) with an interrogative and below with a comparative.

(18) *Anak-ku* *itu* *banyak ilmunya* *lebih* *daripada* *aku*
Child- MY that many knowledge more than see
'My child has more knowledge than me.'
(Hang Tuah 296:12)

It should also be noted that it is possible for a particle to have two different functions as both a discourse and a grammatical marker. Thereby, it can be argued that *maka* serves two purposes in CM: (i) to mark the beginning of a sentence in oral literature and (ii) to specify the sentence type as declarative.

4 Analysis of *maka* under the split CP approach

4.1 Rizzi's (1997) framework

Rizzi (1997) proposes the split CPhypothesis, which holds that the left periphery has a fine structure, as follows:

(19) Force < Topic < Focus < Fin < IP

(Rizzi, 1997: 288)

The highest projection is Force, which contains the clausal type information indicating whether the clause is a question, a declarative, an exclamative, or an adverbial of certain kinds, etc. The lower layer before the IP is Finiteness, which express the tense and other inflectional specifications of the verbal system. The morphological realization of Finiteness varies from one language to another language. Finite forms can manifest mood distinctions, tense, and subject agreement. It can be concluded that the Force and the Finiteness involves selectional relations between a C system and the immediately higher and lower structural system. The C system also has a topic-focus system, which is independent of selectional constraints. The system expresses the topic-comment and focus-presupposition articulations.

4.2 The two types of *maka*

Before I begin with the analysis of *maka*, I present the two types of *maka* in CM sentences: single *maka* and double *maka*.

4.2.1 Single *maka*

Single *maka* exists in front of a main clause with topic marked with *pun*, in front of a main clause with focus marked with *lah*, in front of a main clause with a topic and focus, and after a clausal conjunction.

(20a) In front of a main clause with *pun*

Maka	*baginda kedua pun*		*masuk lah*		*membawa*	*anakda itu*
PART	majesty two	PART	enter	PART	bring	child that

'Their Majesties entered and brought the child with them'
(Adat Raja Melayu, 59:13)

(20b) In front of a main clause with *lah*

Maka	*datang lah*		*ia*	*ke kebun*	*itu*
PART	come	PART	she	to garden	that

'She came into the garden'
(Cumming, 1991: 90)

(20c) In front of a main clause with *pun* and *lah*

Maka	*baginda*	*pun*	*naik*	*lah*	*ke atas kendaraan*	*baginda…*
PART	majesty	PART	mount	PART	on top mount	majesty

'His majesty got on top of his mount…'
(Cumming, 1991: 90).

(20d) After a conjunction

Syahdan	*maka*	*Paya Tu Antara pun*		*kerajaan*	*lah*
Then	PART	Paya Tu Antara	PART	become-king	PART

menggantikan	*ayahanda*	*baginda*	*itu*
change	father	majesty	that

'Then Paya Tu Antara became king succeeding his father'
(Teeuw and Wyatt, 1970: 68)

4.2.2 Double *maka*

There can also be two *maka* in the main clause, one at the initial position of the clause, and the other in front of the subject NP topic marked with *pun*.

(21) *Maka* *di-pukul* *orang lah* *rebab rebana*
PART PASS hit people PART rebab rebana

maka *biduan* *yang baik suara pun* *menyanyi lah*
PART PASS hit REL good voice PART sing PART
'As the rebab and rebana were played, a singer with a beautiful voice started to sing a sad song'
(Adat Raja Melayu 61:2)

The higher *maka* is optional when the clause is preceded by a conjunction.

(22) *Hatta* *antara* *dua bulan* *lamanya* *maka* *negeri itu* *pun* *sudah lah*
PART around two month long PART country that PART finish PART
'Meanwhile after around two months, the construction of the country finished'
(Teeuw and Wyatt, 1970: 70).

As seen in (22), the higher *maka* is optional if the clause is already preceded by a conjunction, *hatta* 'therefore'.

4.3 *Maka* with Rizzi's (1997) split CP approach

It can be concluded that the previous analyses within non-generative frameworks from Ajamiseba (1983), Cumming (1991) and Muller-Gotama (1996) are not adequate to describe the function of *maka*.

Therefore, in this subsection, I analyze *maka* using the framework of generative grammar, specifically the split CP approach of Rizzi (1997), arguing that the CM left periphery supports Rizzi's (1997) multiple functional projections.

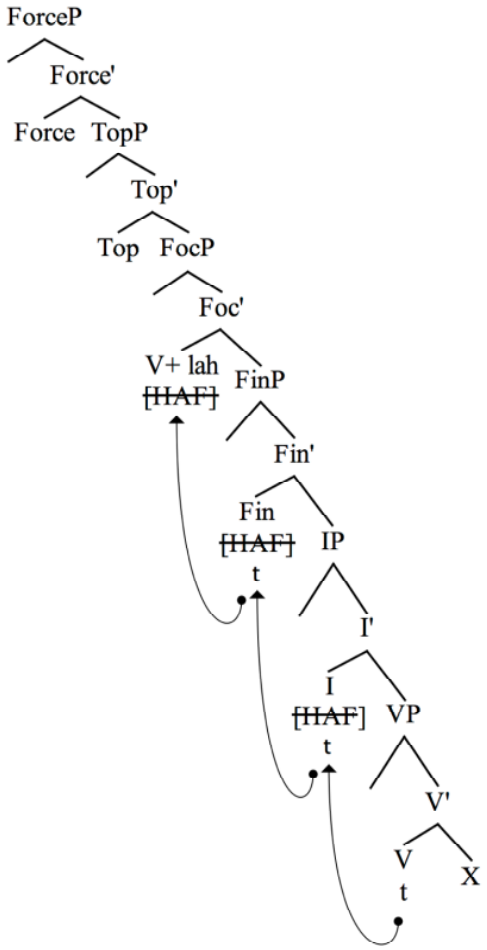
4.3.1 The Particle *pun* and *lah*

I now discuss the particles *pun* and *lah* in the left periphery of CM. I start first with the focus particle *lah.* Both NP (see example 1b) and verb can be focalized in CM. Focalization of a verb involves head movement from the verb head. This raises the following questions: (i) what motivates extraction of the verb to the Spec of Focus (Spec, Foc)?,(ii) how does the derivation select between an NP and a verb? A direct extraction of verb from the Verb phrase (VP)to the Focus phrase (FocP) will violate the head movement constraint (Travis, 1984) since it will skip the Inflectional (I) head and the Finiteness (Fin) head.

To solve the first problem, I follow Pomino (2005) in suggesting that the Focus head can have a head attraction feature (HAF) that can only be satisfied by moving the verb to FocP. For the second problem, I posit that the verb has to land first at intermediate landing positions, the I head and Fin head. This is possible because head movement can be reiterated. For this purpose, I propose that the I head and Fin head are also equipped with HAF. After landing at the Fin head, the verb can continue to raise to the Foc head, to check the latter head's HAF. The verb then adjoins to the particle *lah* at the Foc head in the

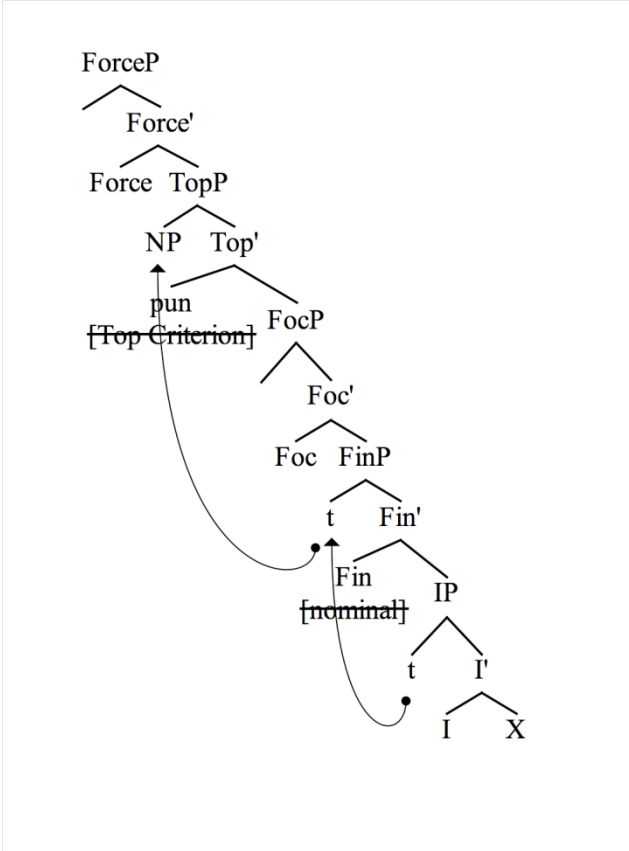
correct order, according to Kayne's (1991, 1994) analysis of head-adjunction as left-adjunction.

(23) The merger of *lah*



Next, I discuss the merger of the topic particle *pun*. The topic particle is merged at the Topic (Top) head located above the Foc head. The particle *pun* marks a topic NP. I assume that a nominative Case feature is checked in Spec, IP. In Rizzi and Shlonsky's (2006) framework, the subject is merged at the Spec of the Subject Phrase (Spec, Subj) above I. The merger of an NP at the Subject Phrase (SubjP) should be able to satisfy the subject criterion as well as the nominative Case features on I. However, the subject NP is prohibited from moving directly to the Spec of Topic (Spec, Top) from the Spec, Subject or Spec, I position, because of criterial freezing (Rizzi, 2003). Based on Rizzi (2003) when a phrase meets a Criterion, it undergoes a criterial freezing, such that it cannot undergo further movement in the derivation. Hence, the subject NP freezes at Spec, Subj and cannot be topicalized in the left periphery. This can be solved by landing the subject at the Spec, Fin before reaching its final destination at the Spec, Top. This is possible because the nominal feature of Fin is uninterpretable and can be checked by passing a nominal expression at the Spec, Fin (Rizzi and Shlonsky, 2006). After passing through the Spec, Fin, the nominal can continue to its final destination at the Spec, Top to satisfy the topic criterion at the Top head.

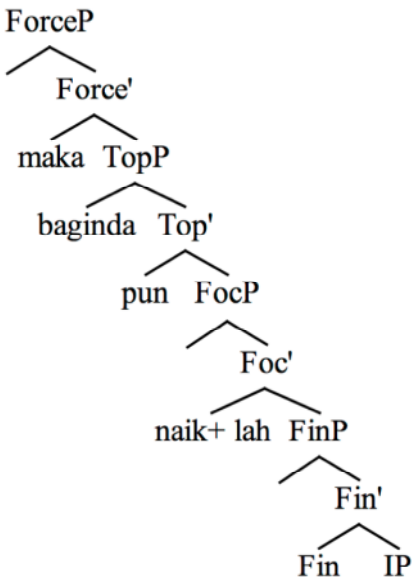
(24) The Merger of *pun*



4.3.2 The single maka

I argue that the single *maka* is a declarative typing particle which occupies the Force head. Based on Rizzi (1997), Force is sometimes expressed by overt morphological encoding on the head (special C morphology for declarative, question or relative). In CM, the C morphology for declarative is expressed by *maka*.

(25) *Maka* *baginda* *pun* naik lah…
PART majesty PART mount PART
'His majesty then mounted (his ride)'



I base my claim on the following evidence. First, *maka* only appears in main clauses, which is characteristic of a pure typing particle. Second, as a pure typing particle, it only marks a specific type of clause, a declarative clause. It does not appear in questions.

(26a) *Sungguh kah* *tuan hamba* *bercakap* *mengobati* *penyakit ..*
really PART master slave country heal illness..
'Are you really saying that you will heal my illness..?'
(Teeuw and Wyatt, 1970: 72)

(26b) **Maka* *sungguh kah* *tuan hamba* *bercakap* *mengobati* *penyakit ..*
PART really PART master slave country heal illness..
'Are you really saying that you will heal my illness..?'

In (26a), the interrogative question is marked with particle *kah* after the focused element *sungguh* 'really'. In (26b), the addition of *maka* in the question hypothetically renders the sentence ungrammatical.

4.3.3 The double *maka*

I now discuss the double *maka*. The higher *maka* is prompted by the merger of the second topic to the left periphery. Therefore I start my discussion with the types of possible topics merged at the Topic Field of the left periphery. Beninca and Poletto (2004) observe that there are four possible topics in Italian: Hanging Topic, Scene-Setting, Left Dislocation, and List Interpretation. Hence, the structure of the topic field in Italian can be described as follows:

(27) Hanging Topic [Scene Setting [Left Dislocation [List Interpretation [… Topic Field…]]]]].

CM has a simpler topic system. Besides the fronted NP marked with *pun*, the CM clause can have a second topic which resembles the Scene-Setting topic in Italian.

(28) *Maka* *di-lihat-nya* *Baginda* *segala perbuatan-nya* *itu*
PART PART-see-3RDSING majesty all doing-3RDSING that

maka *Baginda pun* *tersenyum* *kepada hari* *itu*
PART majesty PART smile to day that
When his majesty saw all that he did, he smiled on that day'
(Adat Raja Melayu, C76:13)

Hence, the CM topic structure in the double *maka* construction can be described as follows.
(29) [Scene Setting [… Topic Field…]].

Next, I posit that the double *maka* constitutes an instance of CP recursion, in which a C takes a CP as its complement. CP recursion can be observed in several languages, among them Welsh. The language has two focus particles *a* and *y*.

(30a) *Y* *dynion* *a* *werthodd* *y* *ci*
The men PART sold the dog
'It's the men who have sold the dog'
(Tallerman 1996: 103)

(30b) *Ym* *mangor* *y* *siaradis* *i* *llynedd*
In Bangor PART spoke I last year
'It was in Bangor I spoke last year'
(Tallerman 1996: 100)

When it is embedded, another particle, *mai*, is merged.

(31) *Dywedais* *i* *mai* *['r* *dynion* *a* *werthith* *y* *ci]*
Said I PART the man PART will-sell the dog
'I said that it's the men who will sell the dog'
(Tallerman 1996: 108)

Moreover, an adverb can be inserted between the focused constituent and the particle *a* as seen in (32). Tallerman (1996) also observes that this is the unmarked order for adverb insertion in the left periphery of Welsh.

(32) *Dywedais* *i* *mai* *['r* *dynion* *fel arfer*
Said I PART the man as usual

a *werthith* *y* *ci*
PART will-sell the dog
'I said that it's the men as usual who will sell the dog'
(Roberts, 2010:128)

Hence, Tallerman (1996) concludes that the position of the particle *a* is in the Fin head, while the adverb occupies Spec, Fin. Based on this data, Roberts(2010) argues that *mai* is merged in Force. He concludes that Force and Fin can be filled by different

complementizers simultaneously. Additionally, Harbert (2010) concludes that the embedding of a clause prompts the merger of particle *mai* to the clause so that declarative features will be close enough to satisfy the selectional requirements of the verb. Harbert (2010) further claims that the particle *y* is merged at the Fin head, which at the same time can also function as a Force Head. However, when the sentence is embedded, the typing particle *y* becomes too deeply embedded in the clause and thereby is not accessible to type the clause anymore for selectional purposes. Therefore, another particle, *mai* is needed in the Force Head.

I agree with Tallerman's (1996) framework in that I argue that CM double *maka* reflects a CP recursion in the language. I also agree with Harbert (2010) in that the higher *maka* is merged after a presence of a second topic at the left periphery, since the lower *maka* is too deeply embedded in the sentence and not visible to type the sentence as declarative. However, I suggest different merging positions for the two *maka*. It can be observed that the Fin head in CM is occupied as an intermediate landing position for the focused verb before reaching the Focus head. The movement of the verb obviously leaves a trace and therefore is not accessible for the lower *maka* to merge. Therefore, I posit that the lower *maka* is base-generated at the Force head and not at the Fin head.

It should be noted that although the lower and the higher *maka* are phonologically identical, they have different functions in the derivation: the lower *maka* types the sentence as declarative, and the higher *maka* extends the scope of the lower one. In Romani and Frisian, a higher *wh*-word can be copied to extend the scope of the lower one. Hence, when a second topic is merged in the languages, the *wh*-word moves to a higher position and leaves behind a copy pronounced at PF.

Romani

(33) a. *Kas misline kas o Demiri dikhl a t ?*
Whom you think whom that Demiri saw
'Who do you think Demir saw?'

Frisian

(33) b. *Wer tinke jo wer' t Jan wennet*
Whom think you whom that Demiri
'Who do you think Demir saw?'
(Kimper, 2008)

Chomsky (1995) proposes that movement is an operation of copying. The copy left behind by the movement may or may not be pronounced at PF. The copy and the moved particle can all be spelt out at PF if both of them are seen as distinct by the linearization algorithm in line with the LCA (Kayne, 1994). Hence multiple copy spell out is possible if (i) both copies are phonologically distinct (Grohmann and Nevins, 2004), or (ii) both copies are semantically distinct (Kimper, 2008). It is clear that the two copies of the *wh*-words in Romani and Frisian are not phonologically distinct; however, they do have distinct function in the derivation.

We can assume that *maka* is raised to the higher head from the Force head when there is a second topic, and that this movement leaves behind a pronounced trace. The two copies can be linearized in the clause since they have different functions. The lower copy is a Force marker, while the higher copy is a scope extender of the Force marker.

So what is the head that hosts the higher *maka*? I refer to Haegeman (2006, 2012) to answer this question. Haegeman (2006, 2012) observes that the Force head coincides with the head hosting the subordinating conjunction. Therefore, Haegeman (2006) proposes a Sub head, hosting the conjunction. The head serves to subordinate the clause and to make

it available for categorial selection independent from the Force head (Rizzi, 1997). As evidence, the merger of the Sub head causes the presence of the higher *maka* to be optional in (34a and b).

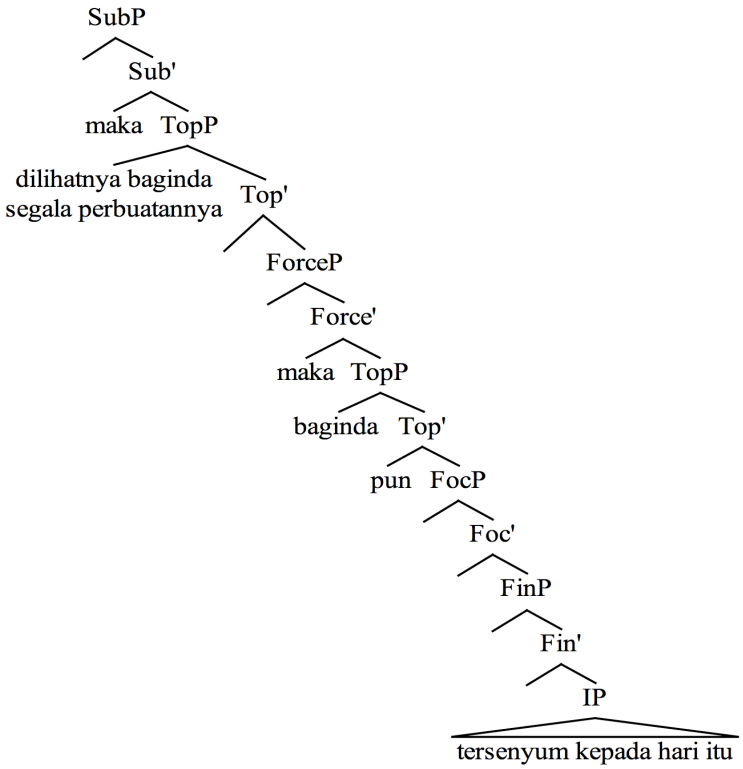
(34a) *Hatta* *beberapa lamanya* *maka* *sampai* *lah* *umur*
After some time PART arrive PART age

anak *baginda* *itu* *akil balig*
child majesty that maturity
'After years gone by, the king's child reached maturity'
(Adat Raja Melayu, 19:3)

(34b) *Setelah* *tawar* *cuka* *itu* *maka* *ambil* *raksa* *itu*
After dissolved vinegar that arrive take mercury that
'After the vinegar is dissolved, take that mercury …'
(Adat Raja Melayu, C71:15)

Hence, I posit the following derivation for the double *maka* with example (28) repeated here, in which the higher *maka* is merged at a Sub head.

(35) The merger of the double *maka*



5 Conclusions

It can be concluded that the CM left-periphery manifests Rizzi's (1997) split-CP projections, despite the double *maka* construction. The lower *maka* is merged at the Force head. When a second topic, a Scene-Setting topic, is introduced to the left periphery, the lower *maka* becomes too deeply embedded and can no longer type the clause as declarative. Therefore, *maka* has to undergo a copy movement to a Sub head above the Scene-Setting topic. The movement leaves behind a pronounced trace at the Force head. The double *maka* can be linearized since they have different functions in the derivation. The lower *maka* serves as a Force marker, while the higher *maka* extends the scope of the Force marker.

Reference

Ajamiseba, Danielo C, 1983,A Classical Malay text grammar: insights into a non-Western text tradition, in *Pacific Linguistics D-56*. Canberra: Australian National University, Department of Linguistics.

Benincà, Paola and Cecilia Poletto, 2004,Topic, focus, and V2: defining the CP Sublayers, In Rizzi (ed.), *The structure of CP and IP*, pp52-75. New York: Oxford University.

Chomsky, Noam, 1995,*The minimalist program*, Cambridge, MA: MIT Press

Cumming, Susanna, 1991, *Functional change: the case of Malay constituent order*, Berlin: Mouton de Gruyter.

Grohmann, Kleanthes and Andrew Nevins, 2004,Echo reduplication: when too-local movement requires PF-Distinctness. University of Maryland Working Papers in Linguistics 13: 84-108.

Haegeman, Liliane, 2006, Argument fronting in English, Romance CCLD and the left periphery, Rafaella Zanuttini, Hector Campos, Elena Herburger, and Paul Portner (eds.), In *cross-linguistic research in syntax and semantics: negation, tense and clausal architecture*, pp 27-52. Washington D.C.: Georgetown University Press.

Haegeman, Liliane. 2012. *Adverbial clauses, main clause phenomena, and the composition of the left periphery*. New York: Oxford University Press.

Harbert, Wayne, 2010,The left periphery of Welsh. In class notes. LING 7712 Syntax Seminar. Cornell University.

Kassim Ahmad (ed.), 1975, Hikayat Hang Tuah, Kuala Lumpur: Dewan Bahasa dan Pustaka.

Kayne, Richard, 1994,*The antisymmetry of syntax*, Cambridge: MIT Press

Kimper, Wendell,2008,Syntactic reduplication and the spellout of movement chains. Unpublished manuscript. University of Massachusetts, Amherst.

Mees, Constantinus A. (ed.), 1935,*De kroniek van Kutai (Cerita Kutai),* proefschrift Universiteit te Leiden, Santpoort.

Muller-Gotama, Franz. 1996, Clause structure in Classical Malay, in *General Linguistics*, Vol.36, No 1/2 p. 89-99. North Carolina: Pegasus Press.

Ong, Walter, J, 1982,*Orality and literacy: The technologizing of the word.* London: Methuen.

Poedjosoedarmo, Soepomo, 1995, Syntactic changes in Malay, In Peter Martin (ed), *Shifting pattern of the language use in Borneo: papers from the second biennial international conference, Kota Kinabalu, Sabah, Malaysia, July, 1992.* Virginia: Borneo Research Council.Inc.

Rizzi, Luigi, 1997, The fine structure of the left periphery, In L. Haegeman (ed), *Elements of grammar*, 281-337. Dordrecht: Kluwer.

Rizzi, Luigi and Ur Shlonsky, 2006, Satisfying the Subject Criterion by a Non Subject: English Locative Inversion and Heavy NP Shift, InM. Frascarelli, (ed),*Phases of Interpretation*, pp. 341-361. Berlin: Mouton de Gruyter.

Roberts, Ian, 2010, Agreement and Head Movement: Clitics, Incorporation, and Defective Goals. Cambridge, MA: MIT Press.

Sudjiman, Panuti H.M. (ed.), 1983,*Adat Raja-Raja Melayu*, Jakarta: Penerbit Universitas Indonesia.

Sweeney, Amin (ed.), 2006,*Karya Lengkap Abdullah bin Abdul Kadir Munsyi*, *jilid 3*: Hikayat Abdullah. Jakarta: Kepustakaan Populer Gramedia &École française dExtrême-Orient.

Travis, Lisa, (1984) *Parameters and Effects of Word Order Variation*. PhD Dissertation, MIT.

Tallerman, Maggie, 1996, Fronting constructions in Welsh, In: Borsley, Robert D. & Roberts, Ian G, ed, *The syntax of the Celtic languages: a comparative perspective*. pp.97-124, Cambridge: Cambridge University Press.

Teeuw, Andries and David K. Wyatt, 1970, *The Story of Patani*. The Hague: KITLV.

19 *The typological perspective of the Balinese serial verb constructions*

NI LUH KETUT MAS INDRAWATI

1 Introduction

Balinese belongs to the Western Malayo-Polynesian family, a subgroup of Austronesian languages (see: Blust1999, Arka2003, Artawa 2004, Mas Indrawati 2012). It has a rich morphological system for marking the valency of the verb. It has no marking for tense, nor any agreement system. In addition to the verbal predicate, it also has a noun phrase, adjectival phrase, prepositional phrase or numerative phrase as predicate, and it has SVO word order.

A serial verb construction (hereafter,'SVC')[1] is a construction which consists of more than one verb without any overt markers of subordination or coordination, and in which all the verbs in the construction are fully inflected. The SVC phenomenon was first observed in African languages as early as the 19th century. Senft (2008:2) mentions that Christaller (1875) discusses this phenomenon in his Twi grammar, and in 1907 Westermann publishedan Ewe grammar containing a discussion on this phenomenon. Years after, studies on SVCs are attracting more attention from linguists, but there are only a few studies on SVCs in Austronesian and Papuan languages. However, the situation is gradually changing. Lynch, Ross and Crowley's (2002) volume, *The Oceanic languages,* provides a section on verb serialisation in Oceanic, and Aikhenvald and Dixon's (2006) volume, *Serial verb constructions: a cross-linguistic typology* includes chapters on Papuan and Austronesian languages. Most recently, Senft's (2008) volume, *Serial verb constructions in Austronesian and Papuan languages,* provides nine chapters discussing SVCs in those languages. Miriam van Staden and Reesink (Senft ed. 2008: 21) mention that now most languages have SVCs.

Indeed, SVCs are of theoretical and typological interest, as their characteristics are language-specific. Theoretically, the suggested properties of SVCs are not all necessarily instantiated in particular languages, as mentioned by Ansaldo (2006:261). He also argues that defining SVCs is problematic, and a satisfying definition of SVCs can be achieved only at the language-specific level. As Durie (1997) suggests, the serialization phenomenon is likely to elude a syntactic analysis, and may be better understood as being based on culture-specific constructions of eventhood and fusion in complex argument

[1] Ansaldo (2006) describes SVC as a clause containing two or more predicates showing properties (denoting a single complex event; as usual the secondary predicate expresses an extension of the basic predicate) that suggest treating them as one structural unit. Other typical characteristics of SVC include: describing a complex event; having only one subject marked for the main predicate; more arguments of the clause are shared; having a shared tense modality, aspect, and polarity markers; perceived as one intonational unit. (cf. Givon,1990;Durie,1997; Payne, 1997)

Asia-Pacific Linguistics, 2014.

structure (cf. Alsina *et al*., 1997). Hence, typologically, different languages exhibit certain language- specific properties of SVCs.

Referring to the language-specific properties of SVCs, this article attempts to provide a description, from a typological perspective,[2]of the Balinese SVC, including its prosodic patterns, its morphosyntax, and its semantic typology. Prior to the discussion of the main topics of this article, a short description on the structure of the Balinese monoverbal clauses is illustrated in the following section.

2 The Structure of the Balinese Monoverbal Clauses

Before discussing the Balinese SVC, it is important to give a short description of Balinese monoverbal clauses. What I refer to as monoverbal clauses are clauses in which verbs have an internal structure consisting of one verbal morpheme or word, and hence constitute one structural unit. The significance of discussing monoverbal clauses is also to show morphosyntactic markers or the full inflections of the verbs in these clauses.

Crosslinguistically, the structure of monoverbal clauses can be categorised into two classes, namely clauses consisting of monovalent verbs, and clauses consisting of two- or three-valence verbs. The first types are referred to as intransitive clauses, and the second types, as transtitive clauses.

Morphologically, intransitive verbs in Balinese can be distinguished into (i) the unmarked forms, also referred to as the base forms or zero intransitives; and (ii) the marked forms or derivational forms, which are prefixed with *ma-* and *N-* (see: Granoka1994; Kardana 2005:111, Sedeng 2007:99), as exemplified below.

(1) a. *I* *Darta* ***luas*** *ke* *pasih.*
ART[3] Name Intr. go prep. beach
'Darta went to the beach.'

b. *Punyan poh-e* *suba* ***ma-bunga****.*
Tree manggo-Def Asp. Intr. flower
'The manggo tree has got flowers'.

c. *I* *Rai* ***ma-gending*** *di* *kamar mandi.*
ART Nama Intr. sing prep. bath room
'Rai sang in the bath room'.

---

[2] This paper is part of my dissertation entitled "Konstruksi Verba Beruntun Bahasa Bali: Kajian Sintaktik dan Semantik . This topic has also been presented in 12 ICAL in 2012.

[3] Abbreviations used in the glossing.

ART=article	Intr= intransitive	2sing.=secondperson singular
AV.=actor	PV.= passive voice	1sing.=first person singular
voiceApl.=aplicative	Pos=possessive	3sing.=third person singular
Asp=aspect	Prep.=preposition	Sta=stative
Caus=causative	Refl= reflexive	
Def.=Definite	UV=Undergoer voice	
Dem=Demonstrative		

d. *I* *Nyoman* ***ny.ebeng*** *dogen.*
ART Name Intr.bad.mood always.
'Nyoman is always in a bad mood'.

e. *Cicing-e* ***ngesil*** *beten* *kursi-ne.*
dog –Def Intr.hide prep. chair-Def
'The dog hid under the chair'.

The clauses in examples (1a) to (1e) have intransitive verbs. Example (1a) shows that the verb is in the base form, while (1b) and (1c) show that the verbs are derived from noun bases, *bunga* 'flower' and *gending* 'sing', with derivational prefix *ma-*. Examples (1d) and (1e) show that the intransitive verbs are morphologically marked with prefix *N-*, specifically their allomorphs *ny-* and *ng-*. The verb *nyebeng* 'be in a bad mood' is derived from the noun base *sebeng* 'bad mood', in which the initial phoneme is dropped when the prefix *N-* is added. The form *ngesil* 'to hide' is derived from the bound root[4] *kesil* 'hide' in which the initial phoneme is dropped when prefixed by *N-*.

The transitive verbs, on the other hand, can be morphologically distinguished into three classes: (i) unmarked forms, i.e., the base form, or zero-transitive; (ii) marked forms with *-a* and *ka-*, which are refered to as passive voice, and (iii) marked forms with *N-*,which are also referred to as active voice. Transitive verbs in (i) and (ii) are categorised into passive voice by Kersten (1984:93-100), while Arka (2003) refers to it (i) as Object voice (OV) or recently undergoer voice (UV), (ii) as Passive Voice (PV), and (iii) as Agentive Voice (AV). Each of these types is exemplified below.

(2) a. *I* *Sura* ***tangkep*** *polisi.*
ART Name UV.catch police
'Sura was caught by the police'.

b. *Pis cang-e* ***jemak*** *cai.*
Money 1sing.-Def UV.ambil 2sing.
'You took my money/ my money was taken by you'.

c. *I* *Bapa* ***nglempag*** *cicing.*
ART father AV.hit dog
'Father hit the dog'

d. *I* *Sura* ***nyemak*** *pipis* *tiang-e.*
ART Name AV.take money 1sing.-Pos
'Sura took my money.'

e. *Cicing-e* ***lempag-a*** *baan* *I* *Bapa.*
Dog-Def PV.hit prep ART father.
'The dog was hit by father.'

f. *Jinah-e* ***ka-ambil*** *antuk* *biang-ne*
money-Def PV-take prep. Mother-3sing.Pos.
'The money was taken his/her mother.'

---

[4] Some linguists consider that this type of bound lexical morphemes as precatergorial, however in this paper I adopt Clynes's (2010:335-336) who has given evidence that bound lexical bases in Balinese are not precategorials.

The verbs *tangkep* 'catch', and *jemak* 'take' in (2a) and (2b) are morphologically unmarked, i.e., in base form. The NPs *I Sura* and *pis cange* 'my money' are the patient arguments which precede the verbs as the subjects of the clauses. The NPs *polisi* 'police' and *cai* are the agent arguments which follow the verbs. Both the NP arguments preceding the verbs and following the verbs are core arguments. The verbs *nglempag* and *nyemak* in examples (2c) and (2d) are morphologically marked with prefix *N*-. The NP arguments, *I Bapa* and *I Sura* are the agent arguments which precede the verbs as functional Subjects, and *cicing* and *pipis tiange* are the patient arguments following the verbs as the Objects. All theNPs are core arguments. This supports Artawa's (2004:44) and Arka's (2003: 5) arguments that in the zero construction (i.e., Objective Voice), it is the patient which behaves as a grammatical subject, not the Agent, while in the *N*-construction, it is the Agent that behaves as a grammatical subject.

The verbs *lempaga* and *kaambil* are morphologically marked with suffix *-a* and prefix *ka*-, which are considered passive affixes. The NPs *cicinge* and *jinahe* are the patient arguments preceding the verbs, while *I Bapa* and *biangne* are the agentive arguments following the verbs, preceded by the prepositions *baan* and *antuk*[5]. The passive marker *ka*- is used for high level Balinese, and the other lexical items used in the clause are also used in high-level Balinese.

Causativisation in Balinese can be morphologically expressed through affixes *-ang*, *-in*, *pa-ang*, and *pa-in* (Artawa, 2004:58). The base of the causative verb marked by such affixes can be a verb, an adjective, or a bound root. This is shown in the following examples.

(3) a. *Tiang* ***ng-ulung-ang*** *gelas-e* *ento*
1sing. AV-fall-Caus glass-Def. that
'I made that glass fall.'

b. *Ia* ***ng-usak-ang*** *sepedan* *tiang-e*
3sing. AV-break-Caus bicycle 1sing.-POS
'He made my bicycle break'.

c. *I* *Karta* ***n-tegak-ang*** *pianak-ne* *di* *korsi-ne.*
ART Name AV-sit-Caus child- 3sing.Pos prep. chair-Def.
'Karta made his child sit on that chair'.

Applicative verbs can be derived from bound roots, intransitive and transitive verbs with suffixes *–ang* and *-in*. This can be seen in the following examples.

(4) a. *Ia* ***n-tegak-in*** *korsi-ne.*
3sing. AV-sit-Apl. chair-Def
'He sat on the chair'.

b. *Tiang* ***ny-jemak-ang*** *I* *Kardi* *baju* *di* *tukang jahit-e.*
1sing. AV-take-Apl. ART Name shirt prep. tailor-Def
'I took the shirt in that tailor for Kardi'.

---

[5] Balinese has two speech levels they are High and Low levels. Pastika (1999:1) and Artawa (2004:2) state that the difference between the two levels is in the lexical choice and it is not significant in terms of morphology and syntax.

3 Cross-linguistic Characteristics of SVCs

Durie (1997:291), Kroeger (2004:229), Aikhenvald and Dixon (2006), and Senft (2008) outline a range of cross-linguistic characteristics of SVCs. They are as follows:

a. A prototypical SVC consists of two or more morphologically independent verbs within the same clause, neither of which is an auxiliary.
b. There are no conjunctions or other overt subordinators or coordinators seperating the two verbs.
c. The serial verbs have a single intonation contour, with no pause separating them.
d. The entire SVC refers to a single event.
e. An SVC shares tense, aspect, modality, negation, etc.
f. The two verbs in the SVC share at least one semantic argument.
g. Obligatory non-coreference: a true SVC will not contain two overt NPs which refer to the same argument.
h. A prototypical SVC contains only one grammatical subject.
i. Constructions analysed as serialisation include those used to express motion, manner, comitative and instrument, aspect and mood, etc.

These typological characteristics can be classified into three, they are: prosodic (characteristic: c), morphosyntactic (characteristics: a, b, d, e, f, g, h) and semantic characteristics (characteristic:, d and i).

The Balinese SVC in this article will be assessed in terms of these characteristics.

4 Typological Perspective of the Balinese SVCs

4.1 Prosodic Perspective

In terms of the prosodic feature of SVCs, the Balinese SVC falls under one intonation contour. There is no pause between the verbs in the series. A break in intonation, either in the form of a marked rise or fall on a non-final verb in a series, indicates that the verbs belong to different clauses (non-SVC). Evidence from Balinese adheres to the observation by Aikhenvald and Dixon (2006) and other linguists that SVCs have intonations similar to the intonations of monoverbal clauses. Consider (5a) where the string *teka* and *ngalih* are parts of an SVC, whereas the same verbs (*teka* and *ngalih*) in (5b) are separately head predicates of their clauses in a coordinate structure. They have different prosodic patterns, as indicated by the lines.

(5) a. *Bli Ketut* ***teka*** ***ng-alih*** *tiang* *tur ...*
Brother name Intr.come AV-search 1sing. and
'Brother Ketut came looking for me and.…'

b. *Bli Ketut* ***teka*** *lan* ***ng-alih*** *tiang* *tur ...*
Brother name Intr.come and AV-search 1sing. and
'Brother Ketut came and looking for me and ….'

Example (5a) shows that it has one intonation. The act of *teka* 'come' and *ngalih* 'search' is a single contour and cannot be separated. However, in example (5b), the insertion of the coordinator *lan* 'and' brings about a pause on the verb *teka*, followed by a

level intonation, and ends in a falling intonation. The prosodic pattern in (5b) is similar to the prosodic pattern of biclausal constructions. Therefore it can be stated that prosodically the Balinese SVC falls under one intonation contour without a pause.

The insertion of negation *tusing* or *nenten* 'not' in the Balinese SVC can also be used as evidence[6] that the Balinese SVC falls under one intonation contour. This is revealed through examples (6a) and (6b).

(6) a.	*Bli Ketut*	***tusing/sing***	***teka***	***ng-alih***	*Tiang tur…*
	Brother name	neg.	Intr.come	AV-search	1sing.
	'Brother Ketut did not come looking for me'				

b.	*Bli Ketut*	***teka***	***tusing/sing***	***ng-alih***	*tiang.*
	Brother name	neg.	Intr.come	AV-search	1sing.
	'Brother Ketut came not looking for me'				

In (6a) the negation marker *tusing/sing* 'not' is inserted before the two serialised verbs. This negative polarity is shared by both verbs, and the clause falls under one intonation contour, hence it constitutes an SVC. When the negation marker is inserted in between the two verbs, the clause does not fall under one intonation contour and the two verbs express separate events, as shown in example (6b).

4.2 Morphosyntactic Perspective

Van Staden and Reesink (2008) distinguish four types of SVCs based on the morphosyntactic properties of the SVCs. They are: (i) *Independent SVC*s, where all the verbs in the construction have the complete range of verbal inflectional morphology that single verbs in simple clauses would have; (ii) *Dependent SVCs,* where only one of the verbs carries all the inflections, while the others are given either in their bare forms or in a stripped down form;.(iii) *Co-dependent SVCs*, which are characterized by a shared argument, and as such, the parts of the construction depend on each other. The object of the first clause is the subject of the second; and (iv) *Complex SVCs*, where two or more verbs have one set of affixes: prefixes attached to the first verb, and suffixes to the last verb, in sequence.

Furthermore, following Van Staden and Reesink (2008), Balinese SVCs can be categorized, on the basis of their morphosyntax, into two types: independent and co-dependent.

---

[6] The other test that can be used to prove that the Balinese SVCs have one intonation is by applying relativisation. The serialized verbs in SVC come under one relative clause and hence fall under one intonation contour, as shown in the following example:

Bli Ketut	*ane*	*teka*	*ngalih*	*Tiang,……*
brother name	Rel. Pron	Intr.come	AV.search	1sing.
Bli Ketut	Who came to look for me,…..			

4.2.1 Independent SVCs

The verbs in the SVCs have morphological markers similar to the markers of the verbs when they occur in monoverbal clauses (or what is also referred to as simple clauses). This can be seen in the following examples.

(7) a. *...I* *Nyoman* ***majalan*** ***nyemak*** *Agus* *panak tiang-e.*
...ART Name Intr. go AV.take Name child 1T.poss
...Nyoman walked to take Agus, my child.

b. *Tiang* *uli* *cerik* *suba* ***malajah*** ***ngigel.***
1sing. prep. little Asp. Intr.learn Intr.dance
Since I was a child I have learnt dancing.

Examples (7a) and (7b) show that the verbs in the SVC have similar morphological markers to the verbs in monoverbal clauses, as seen in the following:

(8) a. *...I* *Nyoman* ***majalan***
...ART Name Intr.walk
...Nyoman walked

b. *...I* *Nyoman* ***nyemak*** *Agus* *panak tiang-e.*
...ART Name AV.take Name child 1 sing..poss.
...Nyoman took his child.

c. *Tiang* *suba* ***malajah***
1sing. Asp. Intr.study
I have studied.

d. *Tiang* *uli* *cerik* ***ngigel.***
1sing. prep. Little Intr.dance
Since I was a child I have danced.

Example (8a) – (8d) show that the verbs in SVCs can occur by themselves in monoverbal clauses. The verb *mejalan* 'walk' is an intransitive verb with prefix *ma-*, *nyemak* 'take' is a transitive verb (with the base form *jemak* and prefix *N-*). The verb markers of these verbs do not change when forming the SVC in (7a). The verbs *melajah* 'study/learn' is an intransitive verb prefixed with *ma-* and its base form is from the bound root *ajah*. The verb *ngigel* 'dance' also has a bound root marked with *N-*. When they occur in SVCs in (7a) both verbs are fully inflected.

4.2.2 Co-dependent SVCs

Co-dependent SVCs are characterized by argument sharing. The object of the first verb (V1) is the subject of the second verb (V2). This type of SVC is illustrated by Van Staden and Reesink as (9a) and exemplified in (9b).

(9) a. (NP) V NPobj=su V (NP)

b. N=babas welik N=mot Do
3SG=bite pig 3SG=die REAL
'It bit the pig dead.' (Bowden 2008:81, Van Staden and Reesink 2008:25)

Example (9b) shows that the Object of V1 is the Subject of V2. This type of construction is referred to as object-sharing by Baker (1998) and in the literature such a construction is also called a causative or resultative construction, in which the first verb expresses the cause, and the second one expresses the result.

This type of SVC in Balinese can be seen in the following examples.

(10) a. *Dadong Sengol* ***ngajakin*** *tiang* ***maturan***.
Grandma Name AV.invite 1sing. Intr.pray
'Grandma Senggol invited me to pray'

b. *...luh-luh-e* ***nyait*** *kebaya* ***anggo-na*** *di* *Usaban-e.*
...women-Def AV.sew clothes PV.wear prep. A kind of ritual day-Def
'...the women sewed kebaya to be worn at Usaba day'.

c. *Luh Rai* ***ngae*** *jaja* ***adepa***.
Name AV.make cake PV.sell
'Luh Rai made cakes to be sold'.

Examples (10a) – (10c) show that the object arguments of V1 is the subject argument of V2. In (10a), V1, *ngajakin* 'ask/invite' semantically needs two arguments. The agent *Dadong Senggol,* syntactically functions as subject, and the patient *tiang* 'I' functions as the object. However, V2 *maturan* 'pray' needs only one argument, therefore the agent functions as the subject. Since the object of V1 and the subject of V2 refer to the same argument, this construction can be categorised as a co-dependent SVC.

Examples (10b) - (10c) show the same thing, however the verbal marking is different: one is active and the other one is passive.

Examples (10a) – (10c) also support Van Staden and Reesink's findingsthat in East Nusantara languages, co-dependent SVCs also involve full inflection, as seen in the independent SVCs, although Balinese does not belong to the East Nusantara languages.

The Balinese verbs that exhibit the morphosyntactic pattern of object-sharing SVCs, such as those in (10a)-(10c), have been discussed in detail by Arka (2003:18) and characterised as control constructions. Arka (2003:18) argues that in many languages, only the syntactic subject can be controlled. However, similar to that of English, Balinese control picks the grammatical subject (GF-Subj). As a consequence, voice marking on the embedded verb must be switched in certain circumstances, so that the controllee is properly controlled. Under the control analysis, V2 is dependent on V1; it is the argument of V1. By contrast, I see the constructions in (10a)-(10c) from the SVC point of view: that is, their prosodic, morphosyntactic and semantic properties cohere with those of SVCs. In terms of prosody, the verbs in constructions (10a)-(10c) fall under one intonation contour; in terms of morphosyntactic properties, all the verbs in the constructions are fully inflected and involved argument sharing; and semantically, (10a) expresses comitative, and (10b) – (10c) express purpose. The V1 in a co-dependent SVC should be in active voice since only active verbs need objects, and these objects are the subject arguments of V2; hence the form of V2 is dependent on the semantic role (e.g., actor, patient, or agent) played by its subject argument.

4.3 The Meanings of Balinese SVCs

Kroeger (2004:227) states that the verbs in SVCs usually function together to express a single event[7]; however, since both verbs contribute to the meaning of the clause, the resulting expression has a more complex[8] meaning compared to the meaning of the individual verbs. The meanings mentioned involve: instrument, manner, goal or direction, and aspect.

Baird (2008:59) describes six types of verb serialization in Keo, namely: benefective/purposive, causative, cause-effect, synonymic, manner and motion serializations.

Van Staden and Reesink (Senft, ed., 2008:36-46) discusses six semantic types of SVCsin the languages of East Nusantara in detail. They are: motion, direction, state change, comitative and instrument, manner, and aspect and mood. All the semantic types which Van Staden and Reesink propose are found in Balinese SVCs.The semantic types found in Balinese SVCs are described in the following sections.

4.3.1 Motion

The motion verbs such as *megedi* 'go', *melaib* 'run', *luas* 'go', *mulih* 'go home', *pesu* 'exit', and *teka* 'come' are very productive as V1, followed by other actions expressed in V2. This type of SVCs hasco-referent subjects. These constructions show that the Agents moves in a certain direction to carry out the events expressed in V2. Crosslinguistically, this type of SVCis very common, as noted by Durie (1997:310):

> "Every serializing languages I have encountered includes a category of motion serialisation, where a verb of motion is combined with some other verb in such a way that the motion verb comes first and the moving argument is the Agent of the second verb".

The Balinese SVCs expressing motion can be seen in the following:

(11) a. *Tiang* *padaduan* ***ma-jalan*** ***ma-ulih***.
1pl both Intr-walk Intr.go home
'We both walked home'.

b. *I* *Conan* ***ma-laib*** ***ng-alih*** *meme-ne.*
ART Nama Intr-run trans.search mother-3poss
'Conan ran searching his mother'.

c. *I* *meme* ***pesu*** ***meli*** *Baju.*
ART mother Intr.exit AV.buy clothes
'Mother went out buying clothes'

---

[7] Bowden (2008:80) argues that it has been noted by people writing on verb serialization that SVC fulfill a function in serializing languages similar to that of the individual verbs in non serializing languages. Hence SVC describes what native speakers conceptualise as single events with the individual verbs referring to subcomponents of those events. Van Stadent and Reesink (2008:28) claim that SVCs fulfill two different functions that is; to compose 'macro-events' out of smaller units which they refer to as 'sub-events' and to compose larger event complexes out of macro events, and they refer to these functions as component serialization and narrative serialization respectively. Following Van Staden and Reesink, in this article I only discuss the function of SVC that is to compose single 'macro-event' out of sub-events (one as framing event and the other co-event).

[8] One of the verbs in SVCs takes adverbial function such as: instrument, manner, goal or direction, etc.

Examples (11a) – (11c) show that V1 *majalan* 'walk', *melaib* 'run', and *pesu* 'exit' are motion verbs. The agents *tiang* 'we', *I Conan* 'Conan', and *I meme* 'mother' move to carry out the events *mulih* 'go home', *ngalih* 'search', and *meli* 'buy' expressed by V2.They all have co-referent subjects. Examples (11a) – (11c) can also be interpreted as having the semantic relation of *purpose* and *manner*. The motion events expressed in V1s are conducted in order to carry out the following events expressed inV2s, or the events expressed in V2s are carried out with the manner of the events expressed in V1. Artawa (2004:138-140) analyses constructions (11a) – (11c) as purposive clauses which are adjuncts to intransitive clauses; hence, they are not SVCs. I argue that the verbs in these construction act as single predicates. One of the verbs is the semantic head, which encodes the event, and the other verb is a modifier, which expresses the manner or purpose in which the event takes place.

A similar observation is made by Foley and Olson (1985:19). They notice that the second verb in an SVC is always in some sense a further development, result or goal of the first verb in the construction, they also argue that SVCs express a single event, where the actions of verbs are simultaneously related and they are conceived by the speaker to be one event. However, in a coordinated or subordinated expression, the verbs are perceived by the speaker to be two different events with different references in time, and the verbs exhibit some morphological dependency.

4.3.2 Directions

Van Stadent and Reesink (in Senft, 2008:38) state that the semantic relation of direction expressed in the SVC is defined by a sequence in which the V2, the 'direction verb', indicates the path or location of a motion or action event. Baird (in Senft, ed., 2008:68) states that a closed class of words expressing spatial relations and directions, like 'seawards', 'inland', 'above' and 'below', are commonly found in both Austronesian and non-Austronesian languages in East Nusantara and the Pacific. The manner or circumstance is given by V1. This can be seen in example (12).

(12) a. *...I* *meme* ***majalan*** ***nganginang***.
…ART mother Intr.walk Intr.east
'…mother moved to the east.'

b. *Pesu uling gang delodan* *, I* *Ketut* ***malaib*** ***ngkajanang***, ...
exit prepalley south ART Name Intr.run Intr.north
'exiting from the south alley, Ketut ran to thenorth….'

In example (12a) and (12b), V2 *nganginang* 'to the east' and *ngajanang* 'to the north' are verbs indicating directions. These verbs are the main verbs of the constructions, while the other verbs V1, *majalan* 'walk' and *malaib* 'run', are verbs of motion which show the manner of performing the actions expressing direction. This type of SVC in Balinese can also be categorised as having semantics of motion.

4.3.3 Manner

Some languages use adverbs or adjectives to indicate manner. English, for example, uses adverbs of manner, such as 'well', 'quickly', and 'slowly'. Indonesian, on the other hand, uses adjectives preceded by the preposition *dengan* 'with' or reduplication, such as *cepat-cepat* 'quickly', *pelan-pelan* 'slowly', *baik-baik* 'well'. Balinese has two ways for expressing manner, which are with the use of adverbs in the form of reduplication such as

enggal-enggal 'quickly', *adeng-adeng* 'slowly', *dueg-dueg* 'cleverly', *melah-melah* 'well', and through SVCs. This supports what is stated by Van Staden dan Reesink (in Senft, ed., 2008:44) that most languages use adverbs or adjectives to express manner, but some languages have more than one way in which to do this. An example of the Balinese SVC with the semantic relation of manner is shown in Example 13.

(13) a. *Bapa Dharma* ***ngenggalang*** ***nyemak*** *baju batik-e.*
Father Name AV.hasten AV. take shirt batik-Def.
'Uncle Dharma hastened to take the batik shirt quickly'

b. *Lampun-ne* ***ngendih*** ***makebyar*** *buka* *ibi puane.*
Lamp- Def Intr.light. Intr.bright like in the past
'the lamp lit brightly like before.'

In Example (13a), V1, *ngenggalang* 'hasten', is an active verb marked with *N*-, formed from the adjectival base *enggal* 'quick'. The V2 verb, *nyemak* 'take', is a transitive verb marked with *N*-. *Bapa Dharma* 'uncle Dharma' is the subject of V1 and also the subject of V2. In this example, *nyemak* 'take' acts as the main verb, while *ngenggalang* expresses the manner.

Example (13b) shows that V1, *ngendih* 'light', is an intransitive verb marked with *N*- with the bound root *endih*. The V2 verb, *makebyar* 'bright', is an intransitive verb marked with *ma*-. *Lampune* 'the lamp' is the subject of V1 and also the subject of V2. In (13a), *ngendih* acts as the main verb, while *makebyar* expresses the manner.

4.3.4 Instrument

The Balinese SVCs exhibiting the semantic relation of instrument are constructions in which one of the verbs expresses the instrument used in performing the action, while another action is specified by the other counterpart in the SVC. They can be composed of an intransitive verb marked with circumfix *ma-an* with noun base. Example (14a) shown below specifies a noun base that can be described as the means of *transportation*, specifically *sepeda* 'bicycle'. This noun can be replaced with *montor* 'car', *bis* 'bus' in the V2 position which indicates the means of transportation used in performing the action specified by V1. Example (14b) shows the verb *nganggo* 'use' followed by the instrument used.

(14) a. *I* *Bapa* ***teka*** ***masepedaan.***
ART father Intr.come Intr.bicycle
'Father came by bicycle'.

b. *I* *meme* ***nganggo*** *tiuk* ***nebih*** *poh.*
ART mother AV.use knife AV.cut manggo.
'Mother used knife to cut manggo'.

In this type of Balinese SVC, the order of verbs filling V1 and V2 is not fixed: note, for instance, that *masepedan teka* and *nebih poh nganggo tiuk* are also fine.

4.3.5 Comitative

Kersten (1984:127) mentions that the comitative in Balinese can be expressed using comitative adverbs *bareng* 'together', *bareng teken* 'together with', and also with the use

of verbs such as *milu* 'join', *ajak /ngajak* 'together with' or *iring /ngiring* 'accompany', (high level Balinese). As a verb forming an SVC, the verb *ngajak* expresses a comitative interpretation, as shown in example (15a).

(15) a.	*...I*	*meme*	*masi*	***teka***	***ngajak***	*I*	*bapa.*
	...ART	mother	also	Intr.come	AV.join	ART	father

'...mothe ralso came with father.'

b.	*I*	*Nyoman*	***milu***	***macelep***	*ka*	*kamar-ne*
	ART	Name	Intr.join	Intr.enter	prep	room-3sing.Pos.

'He entered his room with Nyoman' .

Examples (15a) and (15b) both show that the V2 in these respective examples, *ngajak* 'join' and *milu* 'join, follow', form comitative SVCs. In (15a), the agent of *teka* 'come' is associated with the agent of V2, whose activity is correlated with the agent of V1. In (15b), the agent of *macelep* 'enter' is associated with the agent of V1. In languages without SVCs, like English, the same meaning is expressed using preposition, as illustrated in the translation.

The verbs *ngajak* and *milu* always form SVCs, and the order of their occurrence is not always fixed, as shown in (15c),(15d), and (15e). In (15d) and (15e), the verbs *ajak* (UV) and *ajaka* (PV) show that the subject of V1 is not always the agent of V1.

(15) c.	*...I*	*meme*	***ngajak***	*I*	*bapa*	***teka***
	...ART	mother	AV.join	ART	father	Intr.come

'...mother together with father came here.'

d.	*...I*	*meme*	***ajak***	*I*	*bapa*	***teka***
	...ART	mother	UV.join	ART	father	Intr.come

'...mother and father came here.'

e.	*...I*	*meme*	***ajaka***	***teka***	*baan*	*I*	*bapa*
	...ART	mother	PV.join	Intr.come	Prep.	ART	father

'...mother and father came here.'

This type of SVC in Balinese does not necessarily have an agentive V1, as also shown in (15f) and (15g).

f.	*I*	*Nyoman*	***milu***	***celep-ang-a***	*ka*	*kamar-ne*
	ART	Name	Intr.join	enter-Apl-PV.	prep	room-3sing.Pos.

'Nyoman was put into his room'.

g.	*I*	*Nyoman*	***milu***	***ulung***	*ka*	*kamar-ne*
	ART	Name	Intr.join	Intr.fall	prep	room-3sing.Pos.

'Nyoman was put into his room'.

4.3.6 Aspect and Mood

Aspect and mood in Balinese can be expressed by both modal and aspectual SVCs. The various types of aspectual and mood SVCs expressed in Balinese are: inception, completion, continuation, ability and possibility. Some verbs like *nagih* 'ask for', *nyak*

'will', *nyidayang* 'can', and *maan* 'get'can occur as V1 in Balinese for these kinds of SVCs.

(16) a. *Adin tiang-e* *sing* ***nyak*** ***ngigel***.
Younger Sister I-Pos neg Intr.will Intr.dance.
'My younger sister will not dance'.

b. *Tiang* *sing* ***nyidayang*** ***nganggo*** *komputer-e.*
1 sing. Neg Intr.able Trans.use computer-Def
'I cannot use the computer'.

c. *Tiang* ***maan*** ***nelokin*** *I* *Nyoman.*
1sing. AV. get AV.visit ART Name.
'I have visited Nyoman'

d. *I* *Bapa* ***negak*** ***mebalih*** *TV*
ART Faher Intr.sit AV.watch TV
'Father is sitting watching TV'

In (16a) and (16b), V1 *nyak* 'will', and *nyidayang* 'can' are modal verbs referring to desirability and ability. In (16c), V1, *maan* 'get' is an aspectual verb referring to the completion of the action specified by V2 *nelokin* 'visit/see'. Example (16d) shows that the posture verb *negak* 'sit' implies that the action specified by V2 is still in progress: the use of posture verbs can imply progressive aspect. In this type of SVC in Balinese, the word order of V1 and V2 is fixed.

4.3.7 Causative or Resultative

In Balinese SVCs expressing a causative meaning, the object of V1 is the subject of V2, and V1 is always in the active voice. Therefore this type of SVC is classified as a co-dependent SVC.

(17) a. *I* *Made* ***nglempag*** *I* *Ketut* ***niwang.***
ART Name AV.hit ART Name Sta.faint.
'Made hit Ketut unconscious.'

b. *I* *Mbok* ***ngulung-ang*** *gelas-e* ***belah***.
ART older sister AV.Caus.fall glass-Def Intr.break.
'Older sister made the glass fall so it was broken.'

Data (17a) and (17b) show that the action expressed by V1 causes a change in the state of the patient arguments. Voice alternations of V1 are possible in this type of SVC, as shown in (17c) and (17d); however, the change of voice results in the change of morphosyntactic types of SVC.

(17) c. *I* *Ketut* ***lempaga*** ***niwang.***
ART Name PV.hit Sta.faint.
'Ketut was hit and as a result he was unconscious.'

b. *Gelas-e* ***ulung-ang-a*** ***belah***.
glass-Def fall-Caus-PV Intr.break
'The glass was dropped and as a result it was broken' .

It is also possible to change the order of V1 and V2 in the SVCs (17c) and (17d), as shown in (17c) and (17d).

(17) c. *I* *Ketut* ***niwang*** ***lempag-a.***
ART Name Sta.faint hit-PV
'Ketut was unconscious because of being hit.'

d. *Gelas-e* ***belah*** ***ulung-ang-a***.
glass-Def Intr.break fall-Caus-PV
'The glass was broken because it was dropped'.

4.3.8 Permission

Balinese SVCs expressing permission are composed of V1 *baang* 'give'. This kind of SVC is very productive in Balinese, and the V1 can occur with V2 in the form of passive, active, or objective voice or undergoer voice, and intransitive. The verb *baang* 'give'can also be in active, passive, or objective voice. This can be seen in the following data.

(18) a. *...,mirib* *meme* *sing* ***maang*** ***nganten***.
…,seem mother Neg. AV.give Intr.get.married
'…it seemed that mother did not let (you) get married'.

b. *Jani* ***baang*** *meme* ***ngajak*** *I* *beli* ***ng-keneh-ang*** *ene* makejang
now imp.give mother AV.join ART brother AV-think-Caus this
all
'Now let mother together with brother think of this all'.

c. *"Eda* ***baang-a*** ***maulih*** *buin".*
Don't PV.give Intr.go.home again
'don't let (him/her) go home again'.

The verb *baang* 'give' as V1, in the context of the clauses in (18), expresses permission. The verb *baang* with the meaning of permission is clasified as a verb of influence. Pollard (1994:286) states that this type of verb requires three semantic arguments, which are *influencer*, the *influencee* and *SOA (state of affairs argument)*. In (18a), the influencer is *meme* 'mother', which is the subject of V1; the influencee is the unexpressed object of V1 (the object can be understood from the context of the clause); and *nganten* 'get married' is the SOA argument.

The clause in (18b) is an imperative construction, where V1, *baang* 'let' is analysed as a verb of influence that needs three participants or semantic arguments. The influencer in (18b) is the unexpressed listener. The influencee is *meme ngajak I beli* 'mother together with brother' and *ngenehang ene makejang* 'think of this all' is the SOA argument.

The V1 verb, *baanga* in example (18c), is a negative imperative passive verb with the suffix-*a*. It is preceded by the imperative negation *eda* 'do not'. The influencee is unexpressed and the influencer is the listener, which is also unexpressed, and *mulih* 'go home' is the SOA argument. I here argue that the constructions in (18a) – (18c) belong to

SVCs since the verbs in the constructions act as single predicates, they are expressed under one intonation contour similar to the intonation of a single clause, they are fully inflected, and share the same time and polarity specification.

5 Conclusion

Based on the typological characteristics discussed in the literature, the characteristics of the Balinese SVC can be stated as follows: in terms of prosodic features, the Balinese SVCs have a single intonation contour, with no pause separating them.There are no conjunctions or other overt subordinator or coordinator separating the two verbs. The verbs in the Balinese SVCs refer to a single event. They share negation, time adverbial. The verbs in the SVC share at least one semantic argument and contain only one grammatical subject. The order of verbs in the Balinese SVCs is not always fixed, depending on the semantic of the verbs in series. Following Van Staden and Reesink, who offer four different types of SVCs based on morphosyntactic properties, that is, independent, co-dependent, complex, and dependent, the Balinese SVCs fall under independent and co-dependent ones. In terms of SVC semantics, the Balinese include all the semantics observed by Van Staden and Reesink such as: motion, direction, manner, purpose, comitative, instrument, mood and aspect, and state change. In addition, the Balinese SVC can express permission.

References

Aikhenvald, Alexandra. 2006. Serial Verb Construction in Typological Perspective. In Alexandra Aikhenvald and Robert M.W. Dixon , eds *Serial Verb Constructions: a Cross-linguistic Typology*, 1-87. Oxford Universty Press.

Ansaldo, U.2006. Serial Verb Constructions. In *Serbian-Croatian-Bosnian Linguistic Complex,* 260-264.The Netherlands: Elsevier Ltd

Arka, I Wayan. 2003. *Balinese Morphosyntax: A Lexical-Functional Approach*. Canberra, Australia: Pasific Linguistics Research School of Pasific and Asian Studies The Australian National University.

Arka, I Wayan and Malcolm Ross. 2005. *The Many Faces of Austronesian Voice Systems.* Canberra, Australia.

Artawa, Ketut. 1998. *Ergativity and Balinese Syntax*. Part I, II, III. Dalam *NUSA* Volume 42, 43, 44. Jakarta: Badan Penyelenggara Seri Nusa Universitas Katolik Indonesia Atma Jaya.

Clynes, Adrian. 2010. Bound roots in Balinese and Indonesian -- precategorials or verbs? In *A journey through Austronesian and Papuan linguistic and cultural space: papers in honour of Andrew Pawley*, 333-356. Canberra: Pacific Linguistics

Baker, Mark C. *Object Sharing and Projection in Serial Verb Constructions* Source: Linguistic Inquiry, Vol. 20, No. 4 (Autumn, 1989), pp. 513-553 Published by: The MIT Press Stable URL: http://www.jstor.org/stable/4178644 Accessed: 08/12/2009 00:01

Blust, RA (1999) Subgrouping Circularity and Extinction: Some issues in Austronesian Comparative Linguistics in E. Zeitoun and P.J.K Li (eds), 31-94.

Foley W and Olson M (1985) 'Clausehood and verb serialization'. In Nichols J & Woodbury A C (eds.) *Grammar inside and outside the clause: some approaches to theory from the field.* Cambridge, UK: Cambridge University Press. 17-60.

Durie, Mark. 1997. Grammatical Structures in Verb Serialization. Dalam Alsina Alex, Joan Bresnan and Peter Sells, eds Complex predicater, 289-354. Stanford: CSLI Publications.

Granoka dkk.1984/1985. *Tata Bahasa Bali*. Jakarta: Proyek Penelitian Bahasa dan Sastra Indonesia dan Daerah, Pusat Pembinaan dan Pengembangan Kebudayaan, Departemen Pendidikan dan Kebudayaan.

Kardana, I Nyoman. 2004. *Konstruksi Refleksif dan Diatesis Medial Bahasa Bali*(disertasi). Denpasar: Program Pascasarjana Universitas Udayana.

Kroeger, Paul R. 1993. *Phrase Structure and Grammatical* Relation *in Tagalog*. Stanford, California: CSLI

Mas Indrawati, Ni Luh Ketut. 2012. *Konstruksi Verba Beruntun Bahasa Bali (Kajian sintaktis dan Semantis) (*disertasi). Denpasar: Program Pascasarjana Universitas Udayana.

astika, I Wayan. 2002. *Pemilihan Diatesis dalam Wacana Narasi Bahasa Bali*. DalamBambang Kaswanti Purwo (Peny.). *PELBBA 15*: 181 – 228. Jakarta: Pusat Kajian Bahasa dan Budaya, Unika Atma Jaya.

Pollard, Carl and Ivan A. Sag.1994.*Head-Driven Phrase Structure Grammar*.Studies in contemporary linguistics. Chicago/Stanford, California: University of Chicago Press/ CSLI Publications

Sedeng, I Nyoman. 2007. *Morfosintaksis Bahasa Bali Dialek Sembiran. Analisis Tatabahasa Peran dan Acuan* (disertasi) Denpasar: Program Pascasarjana Universitas Udayana Denpasar.

Senft, Gunter. 2008.*Serial Verb Construction in Austronesian and Papuan Languages*. Australia: Pacific Linguistics Research School of Pacific and Asian Studies.

20 Describing Madurese crossed control

WILLIAM D. DAVIES

1 Introduction

The general topic of this paper is the Madurese instantiation of what has been dubbed 'funny' or 'crossed' control in Indonesian/Malay Gil (2002), Polinsky & Potsdam (2008), Nomoto (2008, 2011). A Madurese example is given in (1).[1]

(1) *Motor sè anyar terro è-belli-yâ moso anom.* crossed control
car REL new want OV-buy-IRR by uncle
'Uncle wants to buy a new car.'

The funny or crossed nature is in comparison with a canonical-looking control structure, as in the synonymous (2).[2]

(2) *Anom terro mellè-ya motor sè anyar.* standard control
uncle want AV.buy-IRR car REL new
'Uncle wants to buy a new car.'

In the standard control structure (2), the matrix nominal *anom* 'uncle' controls the reference of the null subject of the complement and the direct object of the embedded predicate *motor sè anyar* 'new car' behaves as expected, occurring in postverbal position. *Anom* bears two thematic roles, the experiencer of 'want' and the agent of 'buy'. The object *motor sè anyar* is the theme of 'buy'. By contrast, in the control structure in (1), the controller of (2), *anom*, occurs in the complement, yet it is still the experiencer of 'want' and the agent of 'buy'. This is the funny or crossed nature of the structure.

The structure has been the topic of some recent work by Gil (2002), Polinsky & Potsdam (2008), Nomoto (2008, 2011), Sato & Kitada (2012), and others. It is generally agreed that this crossed control structure is distinct from standard control in having a verb phrase complement rather than a larger complement. This accounts for a number of

[1] Although the form identified with the *è-* prefix is frequently referred to as 'passive' like the *di-* form in Indonesian/Malay, the term 'object voice' (OV) is used here. First, there is only a two-way voice distinction in Madurese and the structure in Madurese has grammatical properties characteristic of the object voice in Indonesian/Malay and Balinese, conspicuously the ability of the agent to bind reflexives and variable pronouns in subject position (as illustrated in sections 2.3, 2.4 and 4).

[2] I will refer to constructions such as (2) as 'standard control'. This will designate a control structure with a crossed control verb which appears to be a canonical control structure inasmuch as the matrix subject receives a thematic role from the matrix verb. I will reserve the term 'canonical control' for control structures with predicates that never admit a crossed interpretation, structures familiar from languages such as English, which in Madurese include predicates such as *paksa* 'force', *lantor* 'allow', *oman* 'persuade' and others.

Asia-Pacific Linguistics, 2014.

properties of the crossed control construction to be described below. The purpose of this paper is to describe the basic structure in Madurese, in as much as the phenomenon has not really been examined in much detail in other languages, although that is changing to some degree, cf. Natarina (2012) for Balinese and Kurniawan (2013) for Sundanese. I will be supporting the general type of analysis that has been laid out for Indonesian/Malay, and will briefly consider the ability of existing analyses to account for the data.

An oft-cited sentence in the literature (from Polinsky & Potsdam 2008) illustrates a unique aspect of the structure, the potential for ambiguity:

(3) *Anak itu mau/ingin di-cium oleh ibu.* (Polinsky& Potsdam 2008)
child that want PASS-kiss by mother
i. 'The child wants to be kissed by the mother.'
ii. 'The mother wants to kiss the child.'

The second of the two readings is the crossed control reading, the first being like a standard control type of structure. As is common with ambiguous sentences, not all speakers easily see the ambiguity; however, with appropriate context speakers of Madurese can get either of the readings available in Indonesian for the corresponding sentence. This is illustrated in (4) and (5).[3]

(4) *Kana' rowa è-pa-semma' bâkto* ***terro è-sèyom-a moso embu'***.
child that OV-CS-close time want OV-kiss-IRR by mother
'The child was brought close when the mother wanted to kiss it.'

(5) *Kana' rowa nyemma' polana* ***terro è-sèyom-a moso embu'***.
child that AV.close because want OV-kiss-IRR by mother
'The child came close because he wanted to be kissed by mother.'

In (4), the most salient interpretation of the string *terro èsèyoma moso embu'* is that the mother desires to kiss the child, and the child is being brought to her for that reason. This is the crossed control reading. In (5), the standard control interpretation is most salient, that in which the child is the experiencer of want, because the child is actively seeking out the mother. For the most part, however, there is a preference among the speakers that I have consulted to have an inanimate subject in the crossed control sentences, in which case there is no ambiguity. This is likely because many of the crossed control verbs require a human experiencer, which is most naturally the subject, and so it is difficult to dissociate the experiencer role from an animate subject. Additional examples of sentences with a crossed control interpretation are given in the (b)-examples in (6-7).[4]

---

[3] Kurniawan (2013) reports that even with a rich context this ambiguity does not obtain in Sundanese. Speakers consistently assign a canonical control interpretation when the matrix subject is animate.

[4] Because the crossed control interpretation is difficult for speakers to get when there is an animate matrix subject, the majority of examples considered here have inanimate subjects, which unambiguously induces the crossed control interpretation. The purpose for doing this is to focus on crossed control and identify the grammatical properties associated with it. However, I wish to assert that the crossed control interpretation and the canonical interpretation are derived from the same structure with predicates such as *terro* 'want', *molaè* 'start', *ambu* 'stop', and others. This point is made explicitly in the conclusion in section 6, where I demonstrate that the PP-fronting property of the crossed control structure is exhibited by the same verbs with a canonical control interpretation. This point is also made for Sundanese by Kurniawan (2013) and again in Davies, Kurniawan and Natarina (2013), which considers Madurese, Sundanese, and Balinese.

(6) a. *Bu Yus nyacak ma-pessa kaca ng-angghuy bâto.*[5]
Bu Yus AV.try AV.CS-break glass AV-use rock
'Bu Yus tried to break the glass with a rock.'

b. *Kaca rèya è-cacak è-pa-pessa bi' bu Yus ng-angghuy bâto.*
glass this OV-try OV-CS-break by bu Yus AV-use rock
'Bu Yus tried to break the glass with a rock.'

(7) a. *Ina molaè nyèmprot bengko-na bi' cèt.*
Ina start AV.spray house-DEF with paint
'Ina began to spray her house with paint.'

b. *Bengko-na molaè è-sèmprot bi' cèt moso Ina.*
house-DEF start OV-spray with paint by Ina
'Ina began to spray her house with paint.'

A partial list of predicates that can occur as the main verb in crossed control structures includes *ambu* 'stop', *asèl* 'succeed', *cobâ*/*cacak* 'try', *loppa* 'forget', *molaè* 'start', *nèyat* 'have intention', *potossaghi* 'decide', *satuju* 'agree', *terro* 'want', *terros*/*terrossaghi* 'continue', *tolak* 'refuse'.

2 Principal properties of the Madurese crossed control structure

We now examine a few of the defining characteristics of the crossed control structure, indicating the properties which distinguish it from standard control and standard complement structures.

2.1 No temporal auxiliary in embedded constituent

Crossed control structures do not allow temporal or aspectual auxiliaries in their complements. Some standard control structures do, as do run-of-the-mill complement clauses.

(8) *Motor sè anyar (*bhâkal) è-belli-yâ moso anom*
car REL new will OV-buy-IRR by uncle
'Uncle wants to buy a new car.'

(9) *Kaca rèya è-cacak (*bhâkal) è-pa-pessa bi' bu Yus ng-angghuy bâto.*
glass this OV-try will OV-CS-break by bu Yus AV-use rock
'Bu Yus tried to break the glass with a rock.'

---

[5] While some of the behavior is variable, it appears that when the verbs *cobâ'/cacak* 'try' are in the actor voice, as in (6a), the crossed control interpretation is not possible. This is illustrated in the following sections when illustrating the grammatical properties of crossed control structures, where the 'try' verbs in the actor voice are used to illustrate the behavior of canonical control in opposition to the crossed control structures. There is some equivocation among some speakers regarding the PP-fronting diagnostic discussed in section 2.2 for some speakers in some contexts.

(10) *Ali nyobâ' **bhâkal** ma-beccè' sapèdhâ motor-ra bi' obhing.*
Ali AV.try will AV.CS-good motorcycle-DEF with screwdriver
'Ali tried to fix the motorcycle with a screwdriver.'

(11) *Pa' Salim ngèra anom mellè-ya motor sè anyar.*
Pak Salim AV.think uncle AV.buy-IRR car REL new
'Pak Salim thinks Uncle will buy a new car.'

The structures in (8-9) are crossed control structures. The inclusion of the future auxiliary *bhâkal* 'will' renders the otherwise grammatical sentences ungrammatical. Standard control (10) structures and declarative clausal complements (11) are compatible with the future auxiliary, and thus these sentences are perfectly grammatical.[6]

2.2 Fronting of embedded PP

An embedded PP agent in crossed control can be fronted to the beginning of the sentence, as in (12-13).

(12) ***Bi' polisi**$_i$, malèng sè ghilâ è-cacak è-tangkep* ___$_i$
by police thief REL crazy OV-try OV-catch
'The police tried to catch the crazy thief.'

(13) ***Moso anom**$_i$, motor sè anyar rowa terro è-belli-yâ* ___$_i$
by uncle car REL new that want OV-buy-IRR
'Uncle wants to buy that new car.'

In (12), the agent *bi' polisi* 'by the police' has been front from complement (indicated by the coindexed underscore); in (13) the agent PP *moso anom* 'by uncle' has been fronted. This type of fronting is not possible in canonical control or standard complement structures.

(14) ***Bi' polisi**$_i$, malèng sè ghilâ nyacak è-tangkep* ___$_i$
by police thief REL crazy AV.try OV-catch
(By the police, the crazy thief tried to be caught.)

(15) ***Bi' Dayat**$_i$, Ina abâlâ bèca'-eng è-pa-beccè'* ___$_i$
by Dayat Ina AV.say becak-DEF OV-CS-good
(By Dayat, Ina said that the becak was fixed.)

(14) is a canonical control structure, with the matrix subject *malèng sè ghilâ* 'the crazy thief' controlling the reference of the embedded subject, and fronting of the embedded PP agent *bi' polisi* is disallowed, just as it is in (15), a standard complement structure with the embedded agent *bi' Dayat* fronted. Fronting the agent PP results in ungrammaticality in both cases.

---

[6] Arka (2000) reports that the future auxiliary *akan* cannot occur in control complements in Indonesian, which he takes as an indication that control complements are infinitival. However, Kurniawan (2013) shows that, as in Madurese, the future auxiliary can occur in complements of subject control structures.

2.3 Reflexive in matrix subject can be bound by PP agent

Another property that distinguishes the crossed control structure from canonical control and standard complementation structures is reflexive binding. A reflexive possessor in the subject position of a crossed control structure can be bound by the complement agent (16), whereas the subject of a standard control structure (17) or a typical complement structure (18) cannot be bound by the complement's agent.

(16) *Buku-na* *dhibi'*$_i$ *è-cacak* *è-bâca* *bi' Siti*$_i$.
book-DEF self OV-try OV-read by Siti
'Siti tried to read her own book.'

(17) **Ana'-na* *dhibi'*$_i$ *nyobâ'* *è-sèyom* *bi' Siti*$_i$.
child-DEF self AV.try OV-kiss by Siti
(Her own child tried to be kissed by Siti.)

(18) **Ana'-na* *dhibi'*$_i$ *ngèra* *pèssè-na* *è-kèco'* *bi' Siti*$_i$.
child-DEF self AV.think money-DEF OV-steal by Siti
(Her own child thought the money was stolen by Siti.)

2.4 Bound variable pronouns

Pronominals can be bound by the quantifier *bhân-sabbhân* 'each', resulting in a pair-list reading.

(19) *Bhân-sabbhân* *embu'* *terro* *a-berri'-â* *pèssè* *ḍâ' ana'-na.*
RED-each mother want AV-give-IRR money to child-DEF
'Each mother wanted to give money to her (own) child.'

The sentence in (19) is ambiguous. However, the most salient reading of (19) is that in which every mother wants to give money only to her own child. In this interpretation a mother Siti gives money to Dayat, Ina gives money to Wati, Marlena gives money to Ita, and so on. Of course, depending on the discourse context, (19) could also mean 'Each mother gave money to the child', where one child receives all of the money. It is the bound variable reading that distinguishes crossed control from the other embedding constructions, however.

As the sentence in (20) illustrates, this type of pair-list reading is available for crossed control structures.

(20) *Ana'-na* *terro* *è-berri'-â* *pèssè* *bi'* *bhân-sabbhân* *embu'.*
child-DEF want OV-give-IRR money by RED-each mother
'Each mother wanted to give money to her (own) child.'

While the bound variable interpretation is possible in (20), it should be noted that in this sentence the more salient reading is the one in which there is a single child. This is due to the animacy of the subject. As noted above, for many of these crossed control type predicates, speakers avoid ambiguity and strongly prefer the standard control reading with an animate subject. However, when the subject is inanimate, as in (21), the pair-list reading is the preferred interpretation. So in (21), Ita fixes her own computer, Marlena fixes her own computer, and so on.

(21) *Komputer-ra molaè è-pa-teppa' moso bhân-sabbhân morèd.*
computer-DEF start OV-CS-right by RED-each student
'Each student started to fix her own computer.'

This interpretation is not, however, possible with regular control (22) or standard complementation (23) structures.

(22) *Ana'-na nyacak è-berri' pèssè bi' bhân-sabbhân embu'.*
child-DEF AV.try OV-give money by RED-each mother
'His/her child tried to be given money by each mother.'

(23) *Ana'-na a-bâlâ bengko-na è-eccèt bi' bhân-sabbhân embu'.*
child-DEF AV-say house-DEF OV-paint by RED-each mother
'His/her child said that the house was painted by each mother.'
or 'His/her child said that each mother painted her (own) house.'

In (22), there is some child who tries to be given money by each of the mothers under consideration. In (23), the child is saying either that there is a house that each of the mothers helped to paint, or that each mother painted her own house (the bound variable reading, but only applied to the complement clause).

The inadmissibility of temporal auxiliaries in the complement, the possibility of fronting the complement agent PP, reflexive binding, and bound variable pronouns all distinguish the crossed control structure from the canonical control and standard complement structures.

3 Crossed control and raising

The four grammatical characteristics just outlined also distinguish the crossed control structures from structures that look quite similar, that is, structures that have been analyzed as Raising in the literature. A canonical raising verb is *kèra* 'think, guess'. Raising the subject of an embedded clause is claimed to derive (25) from a structure like (24).

(24) *Ali ngèra anom mellè motor sè anyar.*
Ali AV.think uncle AV.buy car REL new
'Ali thinks Uncle bought a new car.'

(25) *Anom$_i$ è-kèra ___$_i$ mellè motor sè anyar.*
uncle OV-think AV.buy car REL new
'Uncle is thought to have bought a new car.'

In the raising structure, *anom* has been raised from the embedded clause to the matrix. Some raising structures appear to be isomorphic to the crossed control structure. One example is (26b), which would be derived from a form like (26a).

(26) a. *È-kèra anom mellè motor sè anyar.*
OV-think uncle AV.buy car REL new
'It is thought that Uncle a new car.'

b. *Motor sè anyar è-kèra è-belli anom.*
car REL new OV-think OV-buy uncle
'A new car is thought to have been bought by Uncle.'

(26b) appears to have the same surface form as the crossed control clauses, that is, the order of elements in the sentence is subject-verb-verb-agent, and like the examples with any of the verbs meaning 'try', the verbs are both in the object voice. Crucially, unlike the crossed control sentences, the embedded agent can never bear a semantic relation to the matrix verb *kèra.* The raising structure has different grammatical properties as well.

First, temporal auxiliaries are freely admissible in the complement, as in (27).

(27) *Motor sè anyar è-kèra* ***bhâkal*** *è-belli anom.*
car REL new OV-think will OV-buy uncle
*'A new car is thought to will be bought by Uncle.'

Of course, the English translation is ungrammatical, but would better be translated as something like 'It is thought that a new car will be bought by Uncle'.

Second, embedded PP agents cannot be fronted in raising structures, thus the unacceptability of (28).

(28) **Moso Hasan*$_i$, *sapèdhâ motor-ra è-kèra è-pa-beccè'* ___$_i$.*
by Hasan motorcycle-DEF OV-think OV-CS-good
(By Hasan, the motorcycle was thought to have been repaired.)

The string in (28) is fine, however, as long as *moso Hasan* is not construed as the agent of *èpabeccè'* 'repair' but is taken to be a sort of comitative the referent of which was with or at least agrees with the one doing the thinking, as for example *Ita* in (29).

(29) *Moso Hasan, sapèdhâ motor-ra è-kèra Ita è-pa-beccè'.*
by Hasan motorcycle-DEF OV-think Ita OV-CS-good
'With Hasan, the motorcycle was thought by Ita to have been repaired.'

Third, it is not possible for a reflexive contained in the matrix subject to be coreferential with the embedded agent.

(30) **Bengko-na dhibi'*$_i$ *è-kèra è-eccèt Ali*$_i$.*
house-DEF self OV-think OV-paint Ali
(His own house was thought to have been painted by Ali.)

Of course, if *dhibi'* and *Ali* are not coreferential, then the sentence is acceptable with a meaning like 'His house was thought to have been painted by Ali', where the *dhibi'* has a discourse referent.

Finally, pronominal elements in the matrix cannot be bound by quantified NPs in the embedded clause.

(31) *Ana'-na è-yakènnè pa' Salim è-berri' pèssè moso*
child-DEF OV-sure pak Salim OV-give money with
bhân-sabbhân embu'.
RED-each mother
'The child was thought by Pak Salim to have been given money by each mother.'
≠'It was thought by Pak Salim that each mother gave money to her own child.'

4 The monoclausal nature of crossed control

The properties of crossed control that distinguish the structure from the other complement structures are, in fact, characteristic of simple clauses. For example, agents of object voice structures can be fronted (32).

(32) *Moso anom, motor sè anyar è-belli.*
by uncle car REL new OV-buy
'Uncle bought a new car.'

Also, the agent of an object voice structure can bind a reflexive in subject position. In (33), *Siti* binds the reflexive possessor.

(33) *Buku-na dhibi' è-bâca (moso) Siti.*
book-DEF self OV-read by Siti
'Siti read her own book.'

Finally, quantified agents can bind pronouns in subject position.

(34) *Ana'-na è-berri' pèssè bi' bhân-sabbhân embu'.*
child-DEF OV-give money by RED-each mother
'Each mother gives money to her (own) child.'

So, despite the presence of two verbs, unlike other complementation structures, crossed control has the syntactic characteristics of a single clause. Consistent with this, all recent analyses of cross control in Indonesian/Malay have proposed that the complement is a verbal projection of some sort, rather than the sentential/clausal constituents (i.e. CP or TP/IP) of other complementation structures. Three recent analyses have been proposed for crossed control in Indonesian/Malay, all of which are couched in the Minimalist Program. Nomoto (2008, 2011) proposes a *v*P complement in his analysis, and Sato & Kitada (2012) either a VP or *v*P depending on whether the second verb is in passive or object voice. Polinsky & Potsdam (2008) represent the complement as VP, although it appears that this is simply an indication of a verbal projection of some description rather than being a commitment to VP rather than *v*P.[7] The absence of a CP or TP/IP projection accounts for the monoclausal characteristics of crossed control as well as the impossibility of a temporal auxiliary in the complement: as there is only a *v*P/VP node dominating the complement, the projection is too small to accommodate a temporal auxiliary. Therefore, within a general generative theory of syntax, positing some sort of VP complement seems correct.

At the same time, despite the single clause characteristics of crossed control, each does, in Polinsky & Potsdam's (2008) characterization, have an 'independent syntactic existence'. Each can be negated separately, (35).

(35) a. *Motor sè anyar* ***ta'*** *terro è-belli-yâ moso anom.*
car REL new not want OV-buy-IRR by uncle
'Uncle doesn't want to buy a new car.'

[7] Legate (2012) has proposed that in Acehnese the verbs *ci* 'try' and *cuba* 'try' are restructuring verbs in the sense of Wurmbrand (2001), arguing explicitly that the complement is a VP and not *v*P. But the details of Acehnese are significantly different as there is no voice morphology as in Madurese and Indonesian/Malay but an agreement morpheme instead. In Legate's analysis, the agreement morpheme heads the *v*P, but cannot occur on the lower verb in the structure in question.

b. *Motor sè anyar terro* ***ta'*** *è-belli-yâ moso anom.*
car REL new want not OV-buy-IRR by uncle
'Uncle wants to not buy a new car.'

Additionally, the complement VP can be deleted in VP ellipsis, (36).

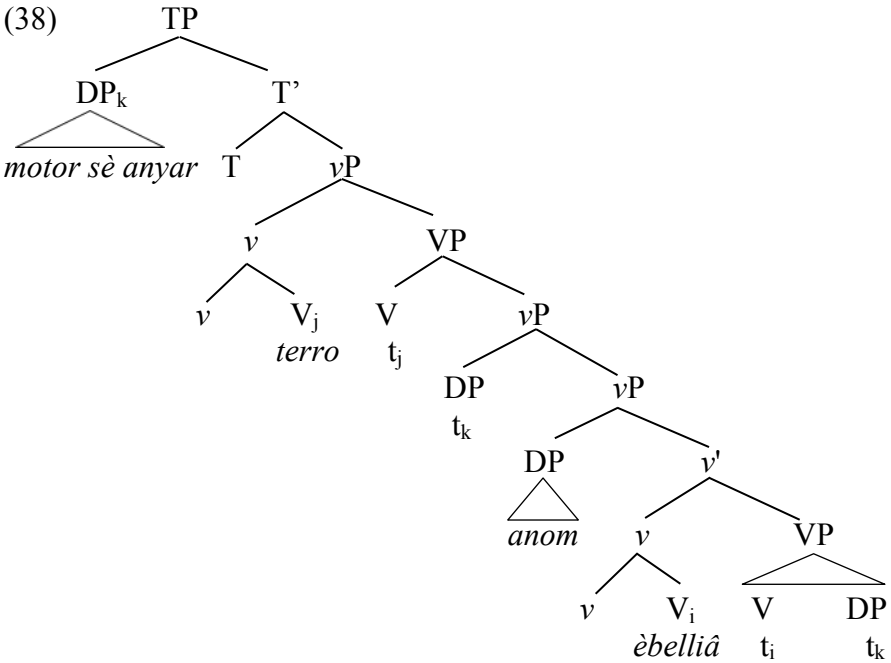
(36) *Motor rowa terro è-juwâl-lâ moso Ali, bân sapèdhâ motor rèya terro ~~è-juwâl-lâ~~ ~~moso Ali~~ kèya.*
car that want OV-sell-IRR by Ali and motorcycle this want OV-sell-IRR by Ali also
'Ali wants to sell that car and (wants to sell) that motorcycle, too.'

Thus an analysis in terms of two distinct verbal heads is indicated.

As stated above, the three recent proposals differ in detail regarding the precise nature of the verbal complement. The differences are attributable to the assumptions made about (i) the analysis of the voice marking (specifically whether it is a *v* head or simply morphology on the verb) and (ii) the assignment of thematic roles. As no one analysis can account for all of the Madurese facts and a precise analysis will not be proposed here, I illustrate one potential structure applying Nomoto's (2008, 2011) analysis to the Madurese sentence in (1), repeated here as (37). (Note that Nomoto treats the object voice morpheme as the head of *v*P, a fairly common assumption in Minimalist analyses of Indonesian-type languages. See also Son & Cole 2008, Cole et al. 2008, and others.)

(37) *Motor sè anyar terro è-belli-yâ anom.*
car REL new want OV-buy-IRR uncle
'Uncle wants to buy a new car.'

(38)

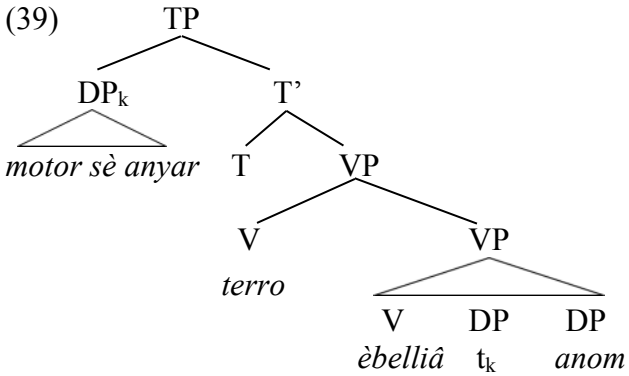


In (38), the complement *v*P is represented as it would be for the simple clause *Motor sè anyar terro èbelliyâ anom* 'Uncle wants to buy a new car'. In the structure, the V heads, *belliâ* 'buy' and *terro* 'want', raise and adjoin to the *v* heads of their respective *v*Ps and *motor sè anyar* 'new car' moves to eventually reside in subject position of the entire sentence. (The movement of the object of the lower VP to the subject of the sentence is a

common feature of all three analyses.) I should note that one positive aspect of Nomoto's analysis is the fact that the agent of the complement *v*P (*anom* 'uncle' in this case) c-commands the object in the same clause. This accounts nicely for the binding facts noted in sections 2.3 and 2.4. On the negative side, however, in (38) the agent remains to the left of the embedded verb, which is not the surface word order. It is necessary to stipulate the order of verb-agent (as Nomoto, in fact, must for Malay). Although illustrated here, Sato & Kitada's (2012) analysis shares these features with Nomoto 2008.

Polinsky & Potsdam's (2008) analysis has the opposite features. Under their analysis, (37) would have a structure something like that in (39).

(39)



Note that in this structure, the proper post-verbal position of the agent is ensured, but the binding facts are not accounted for.

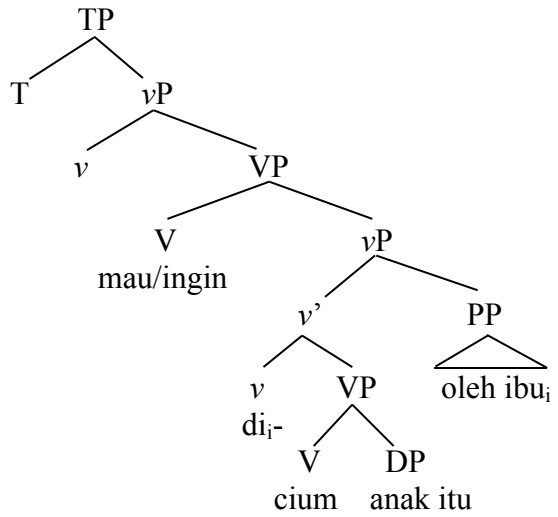
From a theoretical standpoint, the challenge of the crossed control structure has been the assignment of thematic roles—in particular, how the theta role of the matrix predicate, in this case the experiencer role associated with *terro* 'want', can be assigned to an element in the complement—a seemingly non-local relationship. Nomoto (2008), Polinsky & Potsdam (2008), and Sato & Kitada (2012) all propose separate mechanisms for doing so, each of which is stipulative to some degree. All, though, depend on assumptions of Minimalism. However, it is unclear that analyses in other frameworks would be able to any more successfully account for all of the properties of crossed control structures than these.

5 A wrong turn that yields an interesting parallel

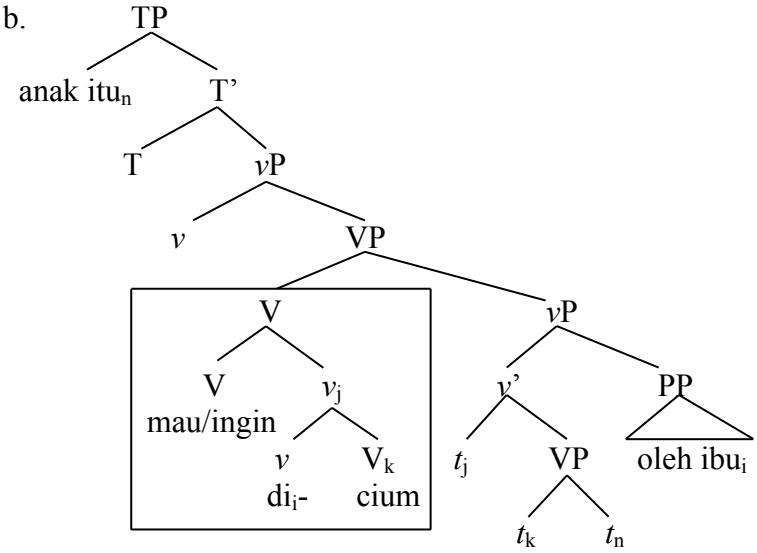
Sato (2010) proposes an analysis, again within the Minimalist framework, in which the *di-* passive prefix in Indonesian is assumed to be a clitic that doubles the 3rd person agent—that is, in (3), the *di-* prefix and *ibu* corefer. As in the other analyses, the complement is a verbal projection (*v*P). But in this analysis, *di-cium*, the lower *v*, raises to adjoin to *mau*/*ingin*, as in (40b).

(3) *Anak itu mau/ingin di$_i$-cium oleh ibu$_i$.*
child that want PASS-kiss by mother
i. 'The child wants to be kissed by the mother.'
ii. 'The mother wants to kiss the child.'

(40) a.



b.



This coindexation and adjunction of the lower verb to the matrix (in the box in (40b)) not only allows long-distance theta-role assignment but crucially accounts for the word order of the lower verb and its agent (which is one of the aspects of Nomoto's analysis that requires stipulation). While this ultimately does not appear to be a very promising analysis for Madurese, it points to an interesting, and perhaps unexpected, parallel between Madurese and Indonesian morphosyntax.

Important to the motivation of this analysis are some language-particular facts of Indonesian which Sato cites as justification for the proposal:

1. *di-* passives in Standard Indonesian are restricted to 3rd person agents—as is well-known, 1st and 2nd person agents of passives occur as clitic pronouns on the verb, as in (A-B).

(41) *Anak itu* *ku-cium/*di-cium* *oleh saya.*
child that 1SG-kiss/ PASS-kiss by me
'I kissed the child.'

(42) *Anak itu* *kau-cium/*di-cium* *oleh anda.*
child that 2SG-kiss/ PASS-kiss by you
'You kissed the child.'

2. *di-* is a clitic form of the 3rd person pronoun *dia*, a somewhat common assumption (Cartier 1979; Shibatani 1985; Guilfoyle, Hung & Travis 1992, among others).

3. The *di-* clitic doubles the 3rd person passive agent—that is, they are coindexed and refer to the same entity.

What is important at this point is the motivation of the analysis rather than the nuts and bolts of adjunction of the lower verb to the higher verb.

It is notable that although the Madurese facts appear to basically parallel those of Indonesian with respect to the crossed control structure, the properties of Madurese object voice/passives are quite distinct in two important ways:

1. Passives with 1st and 2nd person agents are neither morphologically nor syntactically distinct from those with 3rd persons.

(43) *Dayat è-pokol bi' sèngko'/ba'na/aba'na.*
Dayat OV-hit by 1st 2nd 3rd
'I/you/he/she hit Dayat.'

All of the sentences in (43) are perfectly grammatical regardless of the person of the agent.

2. The *è-* prefix bears not even a faint resemblance to any pronominal forms in Madurese, let alone 3rd person pronouns.

However, the Madurese structure does bear a striking resemblance to Indonesian given a distinctly different hypothesis regarding the *di-* morpheme. Gil (2002) has proposed a relationship not between the passive prefix and the third person pronoun but between the passive morpheme and the locative preposition *di*. Specifically, he proposes that the morphemes are one and the same in Riau Indonesian and some other local varieties. Under Gil's analysis, *di* signals the presence/requirement for either a patient or a locative theme, what he takes to be different instantiations of the same superordinate role that he refers to as 'generalized patient'. He likens the generalized patient role to the notion 'actor', which is superordinate to the more specific roles of agent and experiencer.

If one were inclined, one could extend such a hypothesis to Madurese. In Madurese there is a general locative preposition *è*, as in (44) and (45).

(44) *Rina a-temmo Wati è pasar.*
Rina AV-meet Wati at market
'Rina met Wati at the market.'

(45) *Kè Pegkè bâḍâ è dhisa Lembung.*
Ke Pegke exist at town Lembung
'Ke Pegke lived in the town of Lembung.'

The locative preposition is a clitic, and speakers are inconsistent about whether it should be written as a separate word or as part of the object (not unlike what one finds regarding *di* even among educated speakers of Indonesian). Additionally, both passive *è-* and locative *è* undergo the same distinct morphophonological process. When either occurs in the environment right before a word-initial *è*, a palatal glide is inserted (Stephens 1968, Davies 2010).

(46)	*è* + *ènom* OV 'drink'	→	*èyènom* 'be drunk'	/ɛ + ɛnɔm/	→	[ɛjɛnɔm]
	è + *èlong* at 'nose'	→	*èyèlong* 'in the nose'	/ɛ + ɛlɔŋ/	→	[ɛjɛlɔŋ]

That this is a special morphophonological rule is clear in that in all other environments, a sequence of two *è*'s induces glottal insertion, this is true word-internally and with suffixation.

(47)	matè + è 'die' APPL	→	matè'è 'kill'	/matɛ + ɛ/	→	[matɛʔɛ]
	bersè + è 'clean' APPL	→	bersè'è '(make) clean'	/bərsɛ + ɛ/	→	[bərsɛʔɛ]
	lèèr 'neck'	→	lè'èr 'neck'	/lɛɛr/	→	[lɛʔɛr]

While I am not willing at this point to claim that *è*- has the same function as both the object voice morpheme and locative preposition, the parallel with Indonesian (and Sundanese, in fact) is intriguing, and it seems a far less tenuous relationship than the object voice morpheme as an instantiation of pronominal features.

While it is certainly not impossible to propose that *è*- is a clitic that doubles the agent of a passive (and thus adopt Sato's analysis), there is certainly no motivation of the type cited for Indonesian which can be applied to Madurese (or to Sundanese or Javanese for that matter). And the motivation provided by Sato is certainly not wholly uncontroversial in any event.

6 Conclusion

This paper set out a modest goal, describing some of the important grammatical characteristics of the crossed control structure in Madurese. For the most part, Madurese exhibits the same properties as those identified in descriptions of Indonesian/Malay crossed control, those set out most comprehensively in Polinsky& Potsdam (2008). Additionally, however, I have outlined some key monoclausal properties of Madurese crossed control, properties that have not (to my knowledge) been noted in previous descriptions of crossed control: (i) the fronting of PP dependents of the complement VP, (ii) agents in object voice binding reflexives in subject position, and (iii) agents in object voice binding variable pronouns in subject position.

One property in particular is important for the analysis of the 'crossed control' predicates when they occur in what appear to be canonical control structures. Recall that alongside the crossed control structure in (1) is the structure in (2).

(1)	*Motor*	*sè*	*anyar*	*terro*	*è-belli-yâ*	*moso*	*anom.*	crossed control
	car	REL	new	want	OV-buy-IRR	by	uncle	

'Uncle wants to buy a new car.'

(2) *Anom terro mellè-ya motor sè anyar.* standard control
uncle want AV.buy-IRR car REL new
'Uncle wants to buy a new car.'

Polinsky & Potsdam (2008) leave as an open question whether or not the Indonesian analogue of (2) should be subsumed under their analysis of *mau*/*ingin* while Nomoto (2008) assumes the Malay counterpart should but offers no particular evidence. What would be crucial in both cases is that the main verb in (2) take a verbal complement just as it does in (1), and if it does, then the structure in (2) should evince the monoclausal properties of crossed control. Because of c-command restrictions on binding, the reflexive and variable pronoun properties would not obtain. However, fronting of complement PPs should be possible. As the data in (48b) and (49b) indicate, such PPs can indeed be fronted.

(48) a. *Anom molaè ng-eccèt bengko-na bi' kuas.*
uncle start AV-paint house-DEF with brush
'Uncle started painting the house with a brush.'

b. *Bi' kuas, anom molaè ng-eccèt bengko-na.*
with brush uncle start AV-paint house-DEF
'With a brush, uncle started painting the house.'

(49) a. *Embu' terro a-berri'-â kakanan kana'-kana' sè lapar.*
mother want AV-give-IRR food to RED-child REL hungry
'Mother wants to give money to the hungry child.'

b. *Kana'-kana' sè lapar, embu' a-berri'-â kakanan.*
to RED-child REL hungry mother AV-give-IRR food
'To the hungry child, mother wants to give money.'

The fact that *bi' kuas* 'with a brush' and *kana'-kana' sè lapar* 'to the hungry child' can be front in (48) and (49), respectively, provides evidence that rather than the CP complement of canonical control, there is a VP complement in these structures. Such a diagnostic may prove useful in identifying bona fide crossed control predicates.

References

Arka, I Wayan. 2000. Control and argument structure: Explaining control into subjects in Indonesian. Ms.

Cole, Peter, Gabriela Hermon &Yanti. 2008. Voice in Malay/Indonesian. *Lingua* 118:1500-1553.

Davies, William D., Eri Kurniawan & Ari Natarina. 2013. Toward a unified analysis of crossed control. Presented at the 4th International Symposium on the Languages of Java.Universitas Bung Hatta, Padang, Indonesia.

Gil, David. 2002. The prefixes *di*- and *N*- in Malay/Indonesian dialects. In *The History and Typology of Western Austronesian Voice Systems, ed. by Fay Wouk and Malcolm Ross,* 241-283. Canberra: Pacific Linguistics.

Kurniawan, Eri. 2013. *Complementation in Sundanese*. PhD diss, University of Iowa.

Legate, Julie A. 2012.Subjects in Acehnese and the nature of the passive.*Language* 88:495-525.

Natarina, Ari. 2012. Balinese crossed control construction. Ms., University of Iowa.

Nomoto, Hiroki. 2008. A unified analysis of funny control. Presented at ISMIL 12.

Nomoto, Hiroki. 2011. Analisis seragam bagi kawalan lucu [A unified analysis of funny control]. In *Isamu Shoho: Tinta Kenangan "Kumpulan Esei Bahasa dan Linguistik"*, ed. by Hiroki Nomoto, et al. Kuala Lumpur: Dewan Bahasa dan Pustaka.

Polinsky, Maria & Eric Potsdam. 2008. The syntax and semantics of wanting in Indonesian. *Lingua* 118:1617-1639.

Sato, Yosuke. 2010. The crossed-control construction and the syntactic role of passive morphology in Standard Indonesian. lingbuzz/001177 Accessed 16 June 2013.

Sato, Yosuke, & Shin-Ichi Kitada. 2012. Successive feature inheritance, Ɵ-features and the crossed-control construction in Standard Indonesian. http://ling.auf.net/lingbuzz/001523

Son, Minjeong, & Peter Cole. 2008. An event-based account of *–kan* constructions in Standard Indonesian. *Language* 84:120-160.

Wurmbrand, Susi. 2001. *Infinitives: Restructuring and clause structure*. Berlin: Mouton de Gruyter.

21 *Serial Verb Constructions in Amis*[1]

YI-TING CHEN

1 Introduction

In the surface structure of a serial verb construction (SVC), a series of verbs appear without any intervention in between, denoting one single event (e.g., Stahlke 1970, Sebba 1987, Baker 1989, Crowley 2002, Aikhenvald 2006, inter alia). SVCs are reported to exist in Oceanic and Austronesian languages (e.g., Crowley 2002); specifically, some are so-called adverbial verb constructions (AVCs). SVCs are said to exist in Amis, the largest Formosan language, as well.

Two pioneering studies into Amis SVCs are those of Wu (1995) and Huang (1997). Their studies suggest that, in Amis, the surface structures Verb1-Verb2-Actor1 and Verb1-Actor1-Verb2 constitute SVCs. These structures can be found with phrasal verbs, manner verbs, and "to give away to" in the pivotal construction. Example (1) is considered an SVC.

(1) [2]*mi-lingatu* *mi-nanum* *ø-ci* *aki* *t-u* *caciyaw* *n-u* *amis*
AV-begin AV-learn NOM-PPN Aki DAT-CM language GEN-CM Amis
'Aki began to learn the language of Amis.' (Wu1995, p. 34)

Liu (2003) criticizes Wu's classification and argues "phrasal verbs and manner verbs should not play the same grammatical roles as other first verb(s) do in SVCs" (p. 41). Nevertheless, she does not directly argue these structures are not SVCs. Tsai (2008) and Chen (2008) maintain that "no intervening linker" is the one criterion which defines an SVC, and suggest that what Huang (1997) and Wu (1995) claim to be SVCs in Amis are not true SVCs. On this account, most SVCs in Huang (1997) and Wu (1995) are instead analysed as control constructions (e.g., phrasal verbs "to begin" and "to finish").

In addition, unlike many other Formosan languages, in which an adverbial verb construction is an SVC, Amis adverbial verbs often precede another verb with an

[1] The data presented in this paper was collected from several field trips during 2011 and 2012 with the support of the National Science Council grant (NSC100-2410-H-160-009) for which I am grateful. I also would like to thank my major informant, Mr. Samuhuwan Chen, born in 1933, who resides in Chiapin tribe, Taitung, Taiwan, for kindly providing the data. The data was also cross checked by Mr. Zheng and Mr. Sheng, who both reside in Kaohsiung and currently teach Amis.

[2] The gloss of this paper follows Wu (2006). Here are the list of abbreviations: AV: actor-voice; UV: undergoer-voice; NEUT: neuter voice; NOM: nominative case; DAT: dative case; GEN: genitive case; 1st: first person; 2nd: second person; 3rd: third person; SG: singular; PL: plural; PPN: person noun marker; CM: common noun marker; FAC: factuality; ASP: aspectual marker; COMP: complementizer; PREP: preposition; NEG: negator.

intervening linker. For example, according to Yeh and Huang (2009), one of four common strategies for two-verb serialization is to combine a manner verb with an action/motion verb (p. 90). Nevertheless, Liu (2003) shows that in Amis, the manner adverb, such as *harakat* 'fast', becomes a verb, preceding the other lexical verb with an optional linker, *a*.

(2) *harakat* ø-ci *aki* *(a)* *k<um>aen* *t-u* *hemaj*
fast<AV> NOM-PPN Aki COMP eat<AV> DAT-CM meal
'Aki eats meals very fast.' (Liu 2003, p. 66)

In permitting the occurrence of the linker, as shown in (2), the AVCs of Amis differ from those of other Formosan languages. In other words, Amis does not follow the common verb-serialization strategies found in the other Formosan languages. Chang (2010) thus argues that Amis, as well as Paiwan and Marinax Atayal, are not serializing languages, due to the presence of a linker. If so, are there any structures in Amis where more than one verb exists in a clause without a linker?

This paper suggests such a structure does exist in Amis—that there are SVCs, specifically resultative SVCs and sequential SVCs, in the language. Section 2 provides a full description of these two kinds of SVCs. Section 3 analyses the syntactic structures of these two types. Finally, typological issues are discussed in section 4.

2 Two types of SVCs in Amis

There are two functions of SVCs in Amis: they either denote a resultative state (hereafter, RSVC, as in (3) and (4)) or a sequence of actions (hereafter, SSVC, as in (5)). As shown below in (3)-(5), the occurrence of a linker renders each sentence ungrammatical.

(3) *mi-kucakuc-ay* *ø–ci* *aki* *t-u* *kilang* *(*a)* *tahatuluc-ay*
AV-climb-FAC NOM-PPN Aki DAT-CM tree (COMP) top- FAC
tu
ASP
'Aki (has) climbed to the top of the tree.'

(4) *mi-sulikul-ay* *ø–ci* *aki* *(*a)* *mi-siayaw-ay* *ci* *panay-an*
AV-turn-FAC NOM-PPN Aki (LNK) AV-face-FAC PPN Panay-DAT
'Aki turned and faced Panay.'

(5) *mi-sanga-an* *n-i* *aki* *k-u* *liku'* *(*a)*
UV-make clothes-UV GEN-PPN Aki NOM-CM clothes (COMP)

teli-an *i* *palikian*
put-UV PREP wardrobe
'Aki made the clothes and put (them) in the wardrobe.

Examples (6)-(9) show equivalent sentences with the linker *a*. In each counterpart with the linker, the second verb must be actor-voiced (an 'AV-only' restriction), and cannot bear any TMA marking (an 'a temporal' restriction). The occurrence of the linker thus forces the following verb to have the default voice and TMA specifications.

(6) *mi-kucakuc-ay* *ø–ci* *Aki* *t-u* *kilang* *(a)* *tahatuluc*
AV-climb-FAC NOM-PPN Aki DAT-CM tree (COMP) top
'Aki (has) climbed to the top of the tree.'

(7) [3]*tahatuluc-ay* *tu* *n-u* *kilang* *ø-ci* *Aki*
top-FAC ASP GEN-CM tree NOM-PPN Aki
(a) *mi-kucakuc*
(COMP) AV-climb
'Aki (has) climbed to the top of the tree.'/'Aki was on the top of the tree because of climbing.'

(8) *mi-sulikul-ay* *ø–ci* *aki* *(a)* *mi-siayaw* *ci* *panay-an*
AV-turn-FAC NOM-PPN Aki (COMP) AV-face PPN Panay-DAT
'Aki turned and faced Panay.'

(9) *mi-sanga-an* *n-i* *aki* *k-u* *liku'* *(a)*
UV-make clothes-UV GEN-PPN Aki NOM-CM clothes (COMP)
mi-teli I *palikian*
AV- put PREP wardrobe
'Aki made the clothes and put (them) in the wardrobe.'

Note that, in (7), the positions of two verbs are switched, corresponding to Liu's (2003) finding for resultative expressions in Amis.

Good evidence that structures with and without the linker are of two different kinds may be found in the different response of each structure to negation. Structures with the linker *a* do allow the presence of negation, either in the sentence-initial position or preceding the second verb. This is exemplified in (10) and (11).

(10) *mi-lenac-an* *n-i* *aki* *k-u* *pu'ut* *(a)* *caho* *citalem*
UV-grind-UV GEN-PPN Aki NOM-CM knife (COMP) NEG sharp
'Aki grinded the knife but it was not yet sharpened.'

(11) *cuwa* *pi-sulikul* *ø–ci* *aki* *(a)* *mi-siayaw* *ci*
NEG PI-turn NOM-PPN Aki (COMP) AV-face PPN
panay-an
Panay- DAT
'Aki did not turn to face Panay.'

On the other hand, in SVCs, negation is not allowed to precede the second verb, which constitutes evidence of monoclausality. Interestingly, some SVCs are also ungrammatical even if the negator is placed in sentence-initial position[4]. Examples are shown in (12) and (13).

(12) [5]**mi-lenac-an* *n-i* *aki* *k-u* *pu'ut* *caho* *citalem-ay* *tu*
UV-grind-UV GEN-PPN Aki NOM-CM knife NEG sharp- FAC ASP

---

[3] *Tahatuluc* itself can be considered a verb/predicate, as it can take a NP phrase alone, as in *tahatuluc tuciakiikilang* 'Aki was up to the top of the tree.'

[4] To make an SVC negative, the structure must be clefted or pseudo-clefted.

[5] A possible explanation for the ungrammaticality here is that in Amis,*–ay* is a factuality marker (Wu, 2006). The appearance of a negator contradicts the nature of *–ay (tu)*.

(13) **cuwa* *mi-sulikul-ay* *ø–ci* *aki* *mi-siayaw-ay* *ci*
NEG AV-turn- FAC NOM-PPN Aki AV-face- FAC PPN
panay-an
Panay- DAT

Although these structures appear to be an instance of covert coordination, there is good evidence that this is not the case. In Amis, canonical coordination is subject to the Coordinate Structure Constraint, under which a conjoined constituent cannot be extracted, for instance in a pseudo-cleft or cleft construction:

(14) *r<um>atiw* *(*a)* *ma-sakeru* *ø–ci* *aki*
sing<AV> (*COMP) AV–dance NOM-PPN Aki
'Aki is singing and dancing.'(Liu 2003, p. 562)

(15) **ma-sakeru* *ku/u* *r<um>atiw* *ø–ci* *aki*
AV–dance KU/U sing<AV> NOM-PPN Aki

On the other hand, for structures such as (16) and (17), V2 is extractable to a Spec, C position. That is, these structures are not subject to the Coordinate Structure Constraint, and therefore are not covert coordination structures.

(16) *citalem-ay* *tu* *ku* *mi-lenac-an* *n-i* *aki* *k-u*
sharp- FAC ASP KU UV- grind-UV GEN-PPN Aki NOM-CM
pu'ut
Knife
'The knife that Aki grinded was sharp.'

(17) *teli-an* *i* *palikuan* *ku* *mi-sanga-an* *n-i*
put-UV PREP wardrobe KU UV-make clothes- UV GEN-PPN
Aki *k-u* *liku'*
Aki NOM-CM clothes
'What Aki did after making the clothes is to put it in the wardrobe.'

Thus, we may conclude that these structures are syntactically different from those with the linker. In addition, the structures are also distinguishable from SVCs of the other Formosan languages, in that there is no AV-only restriction, nor any atemporal restriction, on the lexical verb. Attention now turns to what the syntactic structures of this construction are, and how they are derived.

3 Assumptions and proposed analysis

This study suggests that these two types of SVCs can be analysed under Baker & Stewart's (2002) treatment for consequential SVCs (CSVC) and purposive SVCs (PSVC), which relies upon adjunction. In addition, I suggest that these two types of SVCs differ from each other in terms of adjunction sites.

According to Baker and Stewart (2002), in a CSVC, there are two transitive verbs in which the shared arguments are both subject and object, and in which the adjoined verb assigns a theme to a *pro*. The evidence for *pro* comes from the distribution of a reflexive-like constituent, the availability of an E-type reading, and extraction possibilities

in Edo. In Edo, *tòbórè* is a reflexive-like element, but it behaves similarly to the English quantifier *all*. It can appear after the second verb in a CSVC but not an RSVC. The implication of this phenomenon is that there is an empty category residing the second vP of a CSVC, but there is no such empty category in a RSVC. Also, in Edo CSVCs, the interpretation of the shared object corresponds to an E-type reading (Evans 1980, cited in Baker and Stewart 2002), supporting the argument that a null pronoun is situated in the second vP, which is not c-commanded by its antecedent. Another piece of evidence supporting a *pro* analysis is observed from the double object construction (DOC). In a CSVC involving DOC, there is no island effect when the goal or locative object is clefted out. In other words, the gap is not a *wh*- trace and the gap is not bound by an operator. Thus, the null constituent of the second vP of a CSVC is an empty category, or a *pro*, precisely. Summarizing observations above, Baker and Stewart (2002) propose that a CSVC is constructed with a second vP adjoining to the first vP and "the object of the second vP is a null pronominal linked to the object of the first vP" (p. 1). This renders the second vP of a CSVC "parallel to infinitival and participial relatives" (p. 40). The reanalysed configuration of a CSVC is presented in Figure 3.1.

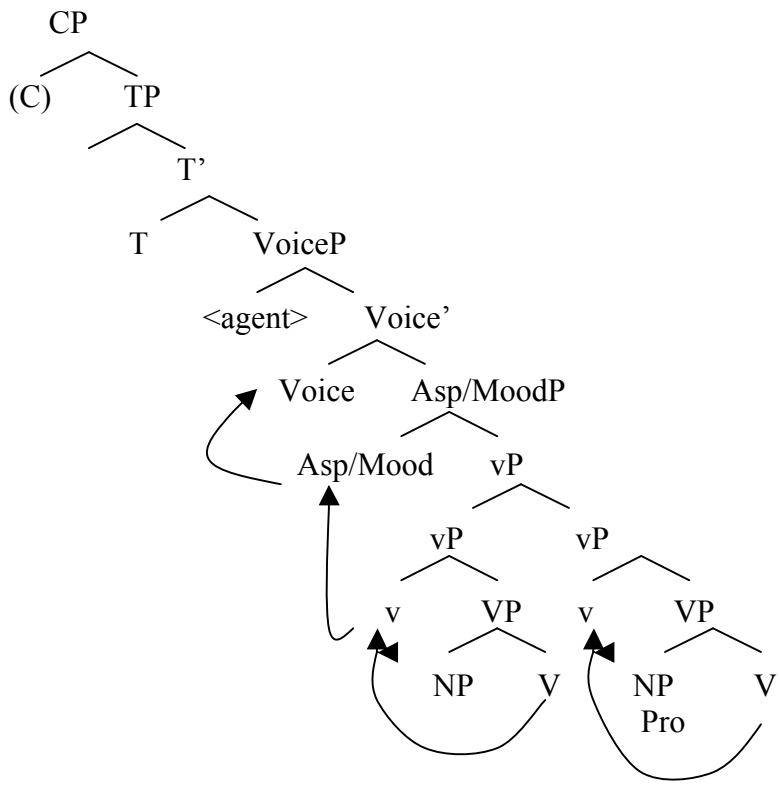


Figure 3.1: CSVC (Baker and Stewart, 2002, p. 9)

In addition to the CSVC, Baker and Stewart (2002) propose another type of SVC, the *purposive serial verb construction*, PSVC. Syntactically, the PSVC behaves quite differently from the CSVC; for instance, in a PSVC, a theme object cannot be clefted out, nor can a theme object of V2 be extracted out of a DOC. Baker and Stewart (2002) attribute observed island effects to an internal operator movement which blocks extraction of the theme object. This leads Baker and Stewart (2002) to postulate that a PSVC is configured as in Figure 3.2.

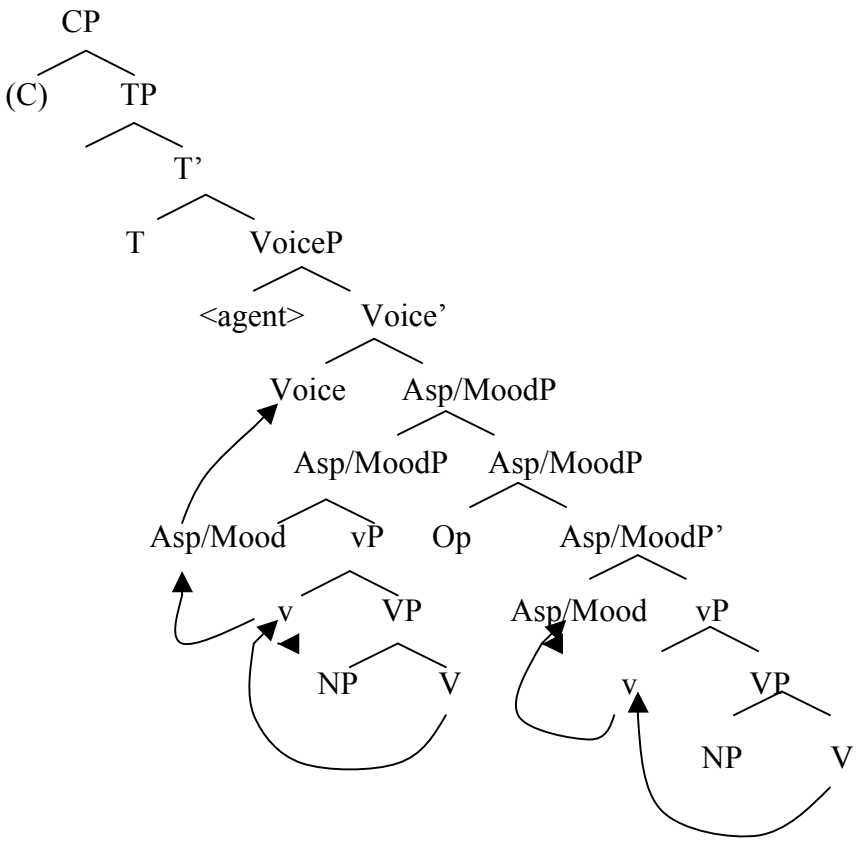


Figure 3.2: PSVC (Baker and Stewart2002, p.10)

Following Baker and Stewart's (2002) adjunction analysis for SVCs, this study proposes that the second verb/predicate of an Amis RSVC is adjoined to the VP of the first verb/predicate, and that of an Amis SSVC is adjoined to an Event Phrase (hereafter, EP).

SVCs are monoclausal. The claim of monoclausality can be bolstered by the negation test. As mentioned in Section 2, in Amis, it is unacceptable for negator to precede the second verb/predicate in an SVC, while it is acceptable in the construction with the linker *a*.[6] The implication is that the clause containing the second verb/predicate is not the complement of the preceding verb, and SVCs hence are not bi-clausal.

One piece of evidence to rule out the complement analysis is that the second verb/predicate is omissible. If the second verb/predicate is a complement of V1, its presence should be obligatory.

(6) *mi-kucakuc-ay* *ø -ci* *aki* *t-u* *kilang*
AV-climb-FAC NOM-PPN Aki DAT-CM tree
'Aki (has) climbed to the top of the tree.'

(9) *mi-sanga-an* *n-i* *aki* *k-u* *liku'*
UV-make clothes-UV GEN-PPN Aki NOM-CM clothes
'Aki made clothes.'

The second piece of evidence stems from the location of locative PPs. If the resultative expression is the complement of V1, then at the level of Logical Form, this expression

[6] According to Tsai (2007), *a* is a non-finite complementizer in Amis. Thus, it also marks the embedded status of a clause.

should be merged after V1, but before the locative PP, which adjoins to V1 later. However, the data contradicts this possibility. Thus, the resultative expression of RMVC is not a complement of V1.

(18) *mi-lenac-an* *n-i* *aki* *k-u* *pu'ut* *i* *kufaw*
UV-grind-UV GEN-PPN Aki NOM-CM knife PREP kitchen
citalem-ay *tu*
sharp- FAC ASP
'Aki sharpened the knife in the kitchen.'

(19) *mi-lenac-an* *n-i* *aki* *i* *kufaw* *k-u* *pu'ut*
UV-grind-UV GEN-PPN Aki PREP kitchen NOM-CM knife
citalem-ay *tu*
sharp- FAC ASP
'Aki sharpened the knife in the kitchen.'

(20) **mi-lenac-an* *n-i* *aki* *k-u* *pu'ut* *citalem-ay* *tu*
UV-grind-UV GEN-PPN Aki NOM-CM knife sharp- FAC ASP
i *kufaw*
PREP kitchen

(21) *mi-kucakuc-ay* *ø -ci* *aki* *t-u* *kilang* *i* *paputal*
AV-climb-FAC NOM-PPN Aki DAT-CM tree PREP yard
tahatuluc-ay *tu*
top-FAC ASP
'Aki climbed to the top of the tree in the yard.'

(22) *mi-kucakuc-ay* *ø -ci* *aki* *i* *paputal* *t-u* *kilang*
AV-climb-FAC NOM-PPN Aki PREP yard DAT-CM tree
tahatuluc-ay *tu*
top-FAC ASP
'Aki climbed to the top of the tree in the yard.'

(23) **mi-kucakuc-ay* *ø -ci* *aki* *t-u* *kilang* *tahatuluc-ay*
UV-climb-UV NOM-PPN Aki DAT-CM tree top-FAC
tu *i* *paputal*
ASP PREP yard

As shown above, in the Amis RSVC, the locative PP should be merged prior to the resultative expression. When the locative PP is situated in the sentence-final position, (20) and (23) are judged to be ungrammatical. On the other hand, the locative PP is free to be situated either after the agent, as in (19) and (22), or after the theme of the first verb, as in (18) and (21). The possible implication of these phenomena is that the second Mod-AspP is an adjunct to rightmost position under VP. Thus, the second Mod-AspP is located on the right-hand side of the locative PP.

Traditionally, the shared argument is assumed to be either a *pro* or an operator. This study assumes that, in Amis SVCs, the shared argument is a null operator, not a *pro*. According to Rizzi (1986), a *pro* must be governed. In (26), based on the Minimal Distance Principle, the closest antecedent is the instrument *fekeroh* 'stone'. However, if the

second verb/predicate is a complement of V1 and the empty category is a *pro*, then the *pro* is not c-commanded by the argument from which the *pro* is interpreted.[7]

(26) *sa-pi-lenac* *n-i* *aki* *t-u* *pu'ut* *k-u* *fekeroh*
SA-PI-grind GEN-PPN Aki DAT-CM knife NOM-CM stone
citalem-ay
sharp- FAC
'Aki sharpened the knife with a sharpening stone.'

In Amis SSVCs, the locative PP cannot be clefted. There are two possible explanations for this. First, Liu (2003) and Wei (2009) all assume that the clefted constituent is base-generated in its surface position, and is predicated of the *ku*- clause by operator variable binding. In such cases, if the locative PP is base-generated in the cleft position, as in (27) and (28), it is not co-indexed with any argument inside the V1 clause. Inside the V1 clause, one argument is covertly shared with another in the V2 clause, in order to saturate both predicates. Therefore, to cleft the locative PP in (29) will leave both predicates unsaturated, resulting in ungrammaticality.

(27) V1 Agent Theme_i $[\text{op}_i$ [V2 $\_\_\_\__i$ PP_k]]

(28) *mi-sanga-an* *n-i* *aki* *k-u* *liku'*$_i$ *(*a)*
UV-make clothes-UV GEN-PPN Aki NOM-CM clothes (COMP)
teli-an $\_\_\_\_\__i$ *i* *palikuan*
put-UV PREP wardrobe
'Aki made the clothes and put them in the wardrobe.'

(29) **i* *palikuan* *ku* *mi-sanga-an* *n-i* *aki*
PREP wardrobe KU UV- make clothes–UV GEN-PPN Aki
k-u *liku'* *teli-an*
NOM-CM clothes put-UV

Another possible explanation is that the configuration above, where V2 is ditransitive, would predict that the prepositional object, a locative PP, is unable to move out, because the extraction is blocked by the gap formed by null operator movement. This prediction is borne out. As shown in (29), the extraction of the locative PP *i palikuan* results in ungrammaticality. This phenomenon is in fact also a case of subject sensitivity, as Liu (1999) and Wei (2009) demonstrate that only a nominative NP is extractable in cleft and pseudo-cleft structures in Amis.

The adjoined constituent is assumed to be a Mod-AspP. In either the RSVC or SSVC, the second verb/predicate is fully fledged, as it allows the presence of the factuality marker *–ay* and aspectual marker *tu*. In fact, the two types of predicative clauses in Williams (1980) are either a PRO in a non-finite clause, or an operator in a finite clause (CP). In Amis SVCs, we also see this combination: an operator in a finite clause.

With respect to the adjunction site, this study assumes that the RSVC and SSVC each adjoin to different sites, thereby explaining the different voice markings in the two kinds of SVCs. In an RSVC, because the second verb/predicate specifies a resultative state, often via an intransitive verb/predicate, the default voice marking is either neutral or actor. On the other hand, in an SSVC, the second verb/predicate is transitive, and the voice marking

[7] Typical examinations of *pro* are not testable in Amis. For example, the universal quantifier is often transformed to a verb. Also, it seems there is no specific reflexive element in Amis. The few choices here make it difficult to offer a clear argument against *pro* here.

corresponds to that of the first verb/predicate. To account for the harmony of voice marking in SSVCs, this study assumes that the second Mod-AspP adjoins to EP.

For SSVCs, this study assumes the adjunction site should be EP. Following Chen (2008), E1 determines which argument undergoes LF movement to Spec, EP1. Following Williams' (1980) analysis of predicative clauses and Baker and Stewart's (2002) analysis of SVCs, Mod-AspP2 becomes predicated of EP1 because the empty category of Mod-AspP2 is coindexed with an NP in some other phrase: precisely Spec, EP1, the one c-commanding the empty category. Thus, the features of E2 and E1 must be the same in order for the empty category to be co-indexed with its correspondent.

As for RSVCs, this study assumes Mod-AspP2 adjoins to the VP of the first verb/predicate, based on the distribution of circumstantial adverbials and the epistemic adverb *alatek* 'probably'. Recall that, in Amis RSVCs, the locative PP is restricted to positions preceding the second verb/predicate. On the other hand, this phenomenon is offset in the case of epistemic adverbs. *Alatek* may be one of the few real adverbs in Amis;[8] however, it can occur in the sentence-final position, as it does in regular sentence, but it cannot occur between an argument and the second verb/predicate in an RSVC. A possible explanation is that *alatek* is an adverb of the speaker-oriented class, or that "higher" adverbs and circumstantial adverbials are located within the VP domain (see the adverbial classification in e.g., Jackendoff 1972, Cinque 1998). If Mod-AspP2 is right-adjoined to VP, it blocks the rightmost position where circumstantial adverbials of the first verb/predicate can adjoin to, making them only occur to the left of the second verb/predicate. On the other hand, speaker-oriented adverbs do not originate at the bottom of VP; even though VP is occupied by Mod-AspP2, it can still adjoin to the pre-VP domain, for instance, to CP and IP.

Given the assumptions above, here are illustrations of how Amis SVCs are derived, using examples (4) and (18).

To begin with (18), the V head *lenac* 'grind' is first merged with the undergoer NP *pu'ut* 'knife', forming a VP.A *v* head is then merged with the VP, projecting a *v*P, and *lenac* undergoes head movement to *v*. The actor *Aki* 'Aki' is then merged in the Specifier position of vP. An E head, merged next in the derivation, carries with it a [+undergoer] feature, which needs to be checked at Logical Form. This feature attracts the undergoer argument *pu'ut* to Spec of EP, covertly. When the verb *lenac* moves again to the head of EP, the [+undergoer] feature spells out as the undergoer voice morphology *mi-...-an*. Next merged in the derivation is the Mod-AspP head, to whch the verb then raises.

Mod-AspP2 adjoins to VP1.To make Mod-AspP2 a possible predicate, Mod-AspP2 and VP1 c-command each other, and there is an operator movement in which the null operator is co-indexed with *pu'ut* in this case.

---

[8] In Amis, most "adverbs/adverbials" are transformed into predicate or verb-like constituents. The epistemic *alatek* is one of very few "adverbs/adverbials" that can occur freely in Amis. Also, it doesn't take any voice marking or TMA markers, nor does it influence the marking of the other lexical verb.

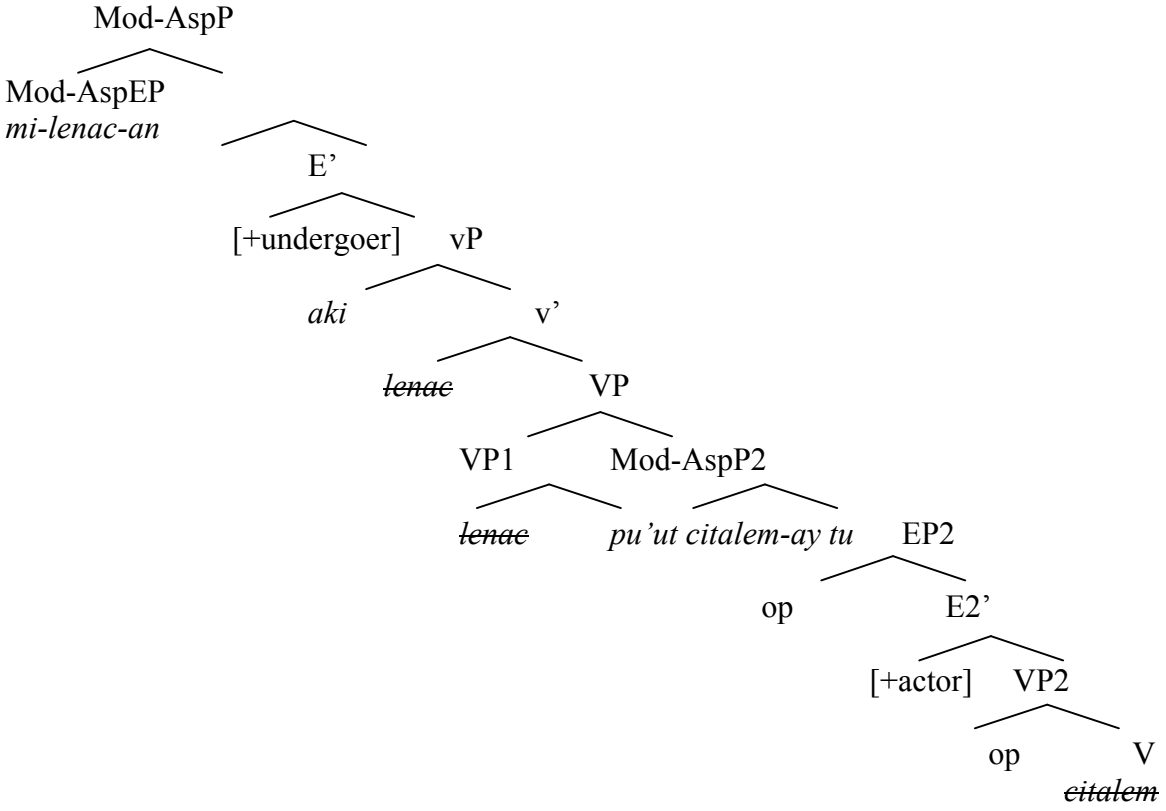


Figure 3.3: RSVC in Amis

Now let's turn to the SSVC in (4). Firstly, the actor *Aki* 'Aki' and the verb *sulikul* 'turn' are merged in an eventual *v*P structure. The [+actor] feature on the E head, merged next in the derivation, must be checked, so the actor undergoes covert movement to Spec of EP1. To make Mod-AspP2 being a predicate possible, the operator inside of Mod-AspP2 (here assumed to be Spec of EP2) is co-indexed with an NP in some other phrase, here the Spec of EP1. Thus, features of E1 and E2 must be the same for the empty category to get its correspondent co-indexed, which, in turn, ensures that the voice marking on two verbs/predicates is in agreement.

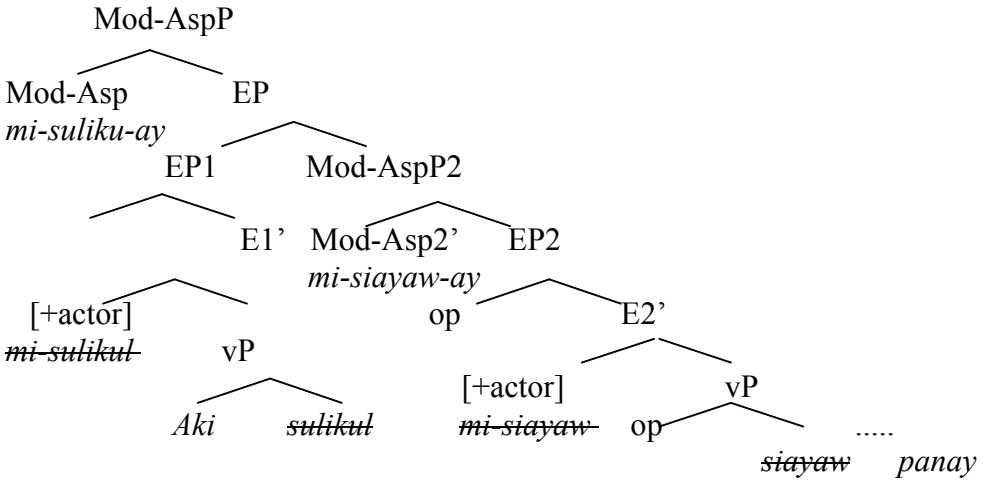


Figure 3.4 SSVC in Amis

4 Typological discussion

Overall speaking, the two types of no-linker SVCs in Amis possess the following characteristics:

a. they are possibly monoclausal;
b. there exist at least two verbs/predicates;
c. there is no linker showing clausal dependency between two verbs/predicates;
d. they share at least one argument, but it is not necessarily an internal one;
e. they share only one TMA and polarity specification, though the second verb/predicate can take TMA marker(s).

They differ from many SVCs in Formosan languages in some aspects: notably, the possibility of TMA marking on the second verb/predicate, and the lack of a linker.

The linkerless condition is often a controversial criterion in defining SVCs in Formosan languages. Huang (1997) and Shibatani (2009) argue that the linkerless condition should be removed, since this criterion poses a dilemma for lexical autonomy, especially of the second verb. Thus, it is often proposed that there is a linker in SVCs in Formosan languages (e.g. Huang 1997, Hsieh 2012).

In studies of Formosan SVCs, an atemporal condition is very often adopted. Researchers such as Chang (2006), Shibatani (2009), and Shibatani and Huang (2006), impose certain restrictions on the second verb. Specifically, it is argued that the second verb cannot host a pronominal clitic, cannot be negated, and cannot be marked for temporal, aspectual, and modal categories. As a result of these restrictions, TMA can be marked *once*, generally on the first verb/predicate, while the second verb/predicate of the SVC exhibits default values, usually carrying actor voice. However, in the structures presented here, the second verb/predicate is eligible for TMA marking, and not subject to an actor-sensitive condition.

Though the structures presented here do not show well-adopted patterns in Formosan SVCs, they are still qualified to be called SVCs.

First, they are monoclausal. Although SVCs have been studied extensively and cross-linguistically, there is no consensus as to how they might be defined or analysed. However, the criterion of monoclausality seems to be widely accepted. Aikhenvald (2006) defines an SVC as "a sequence of verbs which act together as a single predicate, without any overt marker of coordination, subordination, or syntactic dependency of any other sort"
(p. 1). Crowley (2002) treats "serial verbs [as] syntactic constructions involving what can be analysed at the surface level as single clauses, but which are nevertheless expressed by means of multiple predicates" (p. 9). What is implied from quotes above is that an SVC should be a mono-clause with multiple predicates.

Secondly, there is no linker between the two verbs/predicates. Aikenvald (2006) emphasizes that there should be no syntactic dependency marker occurring between multiple predicates, but a marker which bears no relation to syntactic dependency is fine.

Third, there is only one temporal/aspectual specification for both verbs/predicates. Across a significant number of previous studies, whether the second verb/predicate can be TMA marked is not a consistent criterion. For instance, Collins (1997) proposes that, in Fongbe, a TMA marker can only occur once with the first verb/predicate. On the other hand, Noonan (1992) presents SVCs in Lango in which the two verbs/predicates must exhibit the same inflectional marking. In Aikhenvald's (2006) typological survey, the only limitation about TMA marking is that two verbs/predicates must bear the same tense and

aspect setting and share the same polarity value. That is, the second verb/predicate is allowed to be marked, as long as two verbs/predicates share the same TMA and polarity value.

Fourth, argument sharing is observed, although scholars also have different views on the criterion of argument-sharing. If the criterion is just 'argument sharing', then there is no requirement specifying which arguments can be shared in an SVC. Aikhenvald (2006) indicates that subject-sharing is the most common configuration (p. 14) and it is "considered a feature of prototypical SVCs" (p. 14). On the other hand, scholars such as Foley and Olson (1985), Baker (1989), Collins (1997), and Stewart (2001), inter alia, argue that the shared argument must be an internal argument. The subject of a unergative or transitive verb is thereby precluded from being shared in SVCs, since this is assumed to be an external argument. Collins (1997) provides examples, reproduced below, to motivate the limitation that the shared argument be an internal one.

(32) *kofi* *a* *ts˙∴* *ati-ε* *fo* *yao*
Kofi FUT take Stick-DEF hit Yao
'Kofi took the stick and hit Yao with it.'

(33) *Kofi* *a* *liε* *ati* **(a)* *gbe* *ne*
Kofi FUT climb stick *(FUT) pick coconut
(Collins, 1997, p. 466)

According to Collins (1997), (32) is a typical SVC in Fongbe, but (33) is an instance of covert coordination. In (32) the shared argument is an instrument, which is considered to be an internal argument. However, although in (33) the external argument is shared by two verbs, the occurrence of two future markers shows (33) to be a case of covert coordination (CC).

Overall, the composition of these Amis SVCs is similar to what Aikhenvald (2006) classifies as symmetrical SVCs. According to Aikhenvald (2006), SVCs can be grouped into symmetrical and asymmetrical classes, in terms of their composition, Symmetrical SVCs refer to those whose verbs/predicates come from a "relatively large, open, or otherwise unrestricted class" (Aikhenvald, 2006, p.21). By contrast, in an asymmetrical SVC, one verb, the "major verb", comes from an unrestricted class, and the other, the "minor verb", from a closed class. These two types of SVCs, according to Aikhenvald (2006), differ from each other not only in their composition, but also in semantics, iconic constituent order, and grammaticalisation path. Table 4.1 shows their differences.

Table 4.1: Asymmetrical and symmetrical serial verbs (Aikhenvald, 2006, p. 35)

Properties of serial constructions	Asymmetrical	Symmetrical
Semantics	Aspectual, directional, modal, associative, causative	Sequence of events, cause-effect, manner SVCs with synonymous verbs
Iconic constituent order	No: depends on construction type	Yes: for sequential and cause-effect SVC No: for manner and synonymous SVC
Grammaticalisation or lexicalization	Grammaticalisation	Lexicalization

In Amis, these SVCs often present a sequence of events and cause-effect in terms of their semantics, and both types are composed of two fully fledged verbs/predicates in an iconic order.

SVCs in Amis are insensitive to the universal hierarchy of serializing verbs proposed by Foley and Olson (1985). According to Foley and Olson (1985), motion intransitive verbs, such as *go* and *come*, are the most common type of verb to be serialized. For Aikhenvald (2006), typologically speaking, the motion verbs *go*/*come* are among those that occur most frequently in the minor verb slot of an asymmetrical SVC, and tend to be grammaticalised into a particle-like constituent for direction marking. When a motion verb in an SVC loses its lexical status to certain degree, it tends to be situated in the second verb position. In Amis, where *go*/*come* is located depends on the nature of the other verb in the sequence. When put together with another motion intransitive verb, such as *walk* or *run,* and a process verb, such as *go home*, *go*/*come* is preferred to be the second verb position. However, some examples of such cases are not instances of SVCs. When combining a motion intransitive verb, the linker *a* is often allowed (35). The one exception is to combine with a verb whose ending sound is [t] and to produce the sequence in a rapid manner, as in (34).

(34) *r<um>akat=ayra* *ø–ci* *aki*
walk<AV>=go NOM-PPN Aki
'Aki walked there.'

(35) *r<um>akat-ay* *(a)* *tayra* *ø–ci* *aki*
walk<AV>-FAC LNK AV .go NOM-PPN Aki
'Aki walked there.'

Overall, these no-linker structures differ syntactically from the other complement constructions with a linker in Amis. They do possess characteristics of asymmetrical and symmetrical SVCs as suggested by Aikhenvald (2006), and they can be analysed similarly to SVCs in Baker and Stewart (2002). Although some properties of these SVCs are not in concord with characteristics of SVCs in other Formosan languages, these properties in fact are controversial criteria for defining SVCs in general. However, cases of SVCs are not common in Amis: for instance, a resultative expression is more often expressed via result-cause order. To call Amis a serializing language seems a paradox. On one hand, SVCs do exist; on the other hand, they are not prevalent.

5 Conclusions

This paper describes two types of SVCs in Amis. Key properties are that (a) no linker is allowed; (b) the second verb can be marked with temporal, aspectual or modal markers; (c) there is a shared argument; (d) the construction often expresses a resultative state or a sequence of actions. This study also shows that these two types of SVCs are not instances of VP/IP coordination, since they are not subject to the Coordinate Structure Constraint. They also differ syntactically from other complement structures with a linker: a negator can precede a complement clause with a linker, whereas it cannot do so in an SVC.

Following Baker and Stewart (2002), this paper tentatively proposes that V2 is contained in a full Mod-AspP that adjoins to either VP or EP for resultative or sequential SVCs, respectively. This argument is also in line with one characterization of SVCs, where an SVC denotes one single event or one event of sub-events. In an RSVC, cause-effect is constructed under one single event, and in an SSVC, one big event dominates two

sub-events. Evidence for this claim comes from (a) the placement of locative PPs and speaker-oriented adverbs; (b) co-indexed voice markers; (c) non-binding effects; and (d) island effects.

References

Aikhenvald, Alexandra, 2006, Serial verb constructions in typological perspective. In Alexandra Aikhenvald and Robert Malcolm Ward Dixon, eds. Serial verb constructions: a cross linguistic typology, 1-68. Oxford: Oxford University Press.

Baker, Mark, 1989. Object sharing and projection in serial verb constructions. Linguistic Inquiry 20: 513-553.

Baker, Mark, and Osamuyimen T. Stewart. 2002, A serial verb construction without constructions. M.s. Rutgers University.

Chang, Henry, 2006, The guest playing host: Adverbial modifiers as matrix verbs in Kavalan. In Hans-MartinGärtner, Paul Law, and Joachim Sabel, eds. Clause structure and adjuncts in Austronesian languages, 43-82. Berlin: Mouton de Gruyter.

— 2009, Adverbial verbs and adverbial compounds in Tsou: A syntactic analysis. Oceanic Linguistics 48: 439-476.

— 2010, On the syntax of adverbial verbs in Formosan languages. In Raphael Mercado, Eric Potsdam, and Lisa de Mena Travis, eds. Austronesian and theoretical linguistics, 183-212. Amsterdam: John Benjamins.

Chen, Yi-Ting, 2008, A minimalist approach to Amis structure and complementation. PhD dissertation. Tempe: Arizona State University.

Collins, Chris, 1997, Argument sharing in serial verb constructions. Linguistic Inquiry 28: 461-197

Crowley, Terry, 2002, Serial verbs in Oceanic: a descriptive typology. Cambridge: Cambridge University Press.

Foley, William, and Mike Olson. 1985, Clausehood and verb serialisation. In Johanna Nichols and Anthony C. Woodbury, eds. Grammar inside and outside the clause: some approaches to theory from the field, 17-60. Cambridge: Cambridge University Press.

Huang, Lilian, 1997, Serial verb constructions in some Formosan languages: a typological view. Paper presented at the Eighth International Conference on Austronesian Linguistics, Taipei, Taiwan.

Liu, En-Hsing, 2003, Conjunction and modification in Amis. Master thesis. Hsin-Chu, Taiwan: National Tsing-Hua University.

Liu, Tsai-Hsiu, 1999, Cleft constructions in Amis. Master thesis, Taipei, Taiwan: National Taiwan University.

Newmeyer, Frederick J., Some thoughts on the serial verb construction. Paper presented at La notion de « construction verbale en série » est-elle opératoire? Paris, France.

Noonan, Michael, 1992, A grammar of Lango. New York: Mouton de Gruyter.

Rizzi, Luigi, 1986, Null object in Italian and the theory of pro. Linguistic Inquiry 17: 501-557.

Sebba, Mark, 1987, The syntax of serial verbs. Amsterdam: John Benjamins.

Shibatani, Masayoshi, 2009, On the form of complex predicates: toward demystifying serial verbs. In Johannes Helmbrecht, et al., eds. Form and function in language research, 309-336. Berlin: Mouton de Gruyter.

Shibatani, Masayoshi, and Lilian Huang, 2006, Serial verb constructions in Formosan languages: their form and grammaticalization pattern. Paper presented at the 3rd Oxford-Kobe Linguistics Seminar: The Linguistics of Endangered Languages, Kobe, Japan.

Stahlke, Herbert, 1970, Serial verbs. Studies in African linguistics 1: 60-99.

Stewart, Osamuyimen T., 2001, The serial verb construction parameter. New York: Garland Publishing.

Tsai, Wei-Tien, 2007, Conjunction reduction and its origin: a comparative study of Tsou, Amis and Squliq Atayal. Oceanic Linguistics 46: 585-602.

Wei, Ting-Chi, 2009, Interrogatives in Amis. Language and Linguistics 10: 315-374.

Williams, Edwin, 1980, Predication. Linguistic Inquiry 11: 203-238.

Wu, Chun-Ming, 2006, Verb serialization in Kanakanavu. USTWPL 2: 109-139.

Wu, Joy, 1995, Amis complex constructions. Master thesis. Taipei, Taiwan : National Taiwan Normal University.

— 2006, Verb classification, case marking, and grammatical relations in Amis. PhD dissertation. Buffalo, New York: the State University of New York at Buffalo.

Yeh, Yuting and Shuanfan Huang. 2009, A study of triple verb serialization in four Formosan languages. Oceanic Linguistics 48: 78-110.

22 *Conjunction in Thao*

PAUL JEN-KUEI LI

1 Introduction

Thao is a Formosan language spoken at Sun-Moon Lake in central Taiwan, with a small population of 728. It is a highly endangered language, with fewer than ten older speakers left. Most Thao people speak Taiwanese (Southern Min) in their daily life. Intermarriage between Thao and Bunun is common. Thao is not very well studied, as compared with many other Formosan languages. There are a few valuable linguistic studies on Thao, but none of them deal with discourse.

In this paper I shall discuss the conjunctives in Thao discourse and grammar.[1] All examples are based on the 42 Thao texts I collected from two main informants, one female and one male, over two different periods: 1975-1976 with the female informant, Puni, and 1996-2003 with the male informant, Kilaʃ (Li 2011). These texts cover a wide range of topics, including their personal experience, work and life, festivals, myth and traditions. The female speaker's narratives are in simple style and straightforward, while the male speaker's narratives are much more elaborate with great details. The length of each text varies from less than ten sentences to more than two hundred.

A few other Thao texts have been collected by F. K. Li (1956), Blust (2003:244-74), Wang (2004:351-72) and Kuo (2008). Since F. K. Li collected his two short texts from an older Thao informant, while Blust and Wang collected theirs from the same informant, Kilaʃ, I shall touch on F. K. Li's texts for a comparison.

Thao conjunctives include *numa* 'then', its derivative *numawan* 'therefore', *maqa* 'because, so that' and *ya* 'if, when, complementizer'. There is some overlap between the conjunctive *numa* and coordinator *masa* 'and', so I shall also touch on coordination.

Thao conjunctives are often optional. As stated in Blust (2003:217), "Although conjunctions do occur in some environments, clauses typically are conjoined with no overt connecting element."

2 Conjunctives in Thao

2.1 The coordinator *masa* 'and'

Thao uses the coordinator *masa* 'and' to conjoin smaller units, including two nouns (1), two verbs (2), and two verb phrases (3), as illustrated below:

[1] I'd like to thank Wayan Arka and the anonymous reviewer for their helpful comments, which lead to great improvement of the paper.

(1) *ma-faðaq* *m-apa* *buna* *masa* *kawi.*
AF[2]-know AF-carry.on.back sweet.potatoes and firewood
'I know how to carry sweet potatoes and firewoodon my back.' (5-7)[3]

(2) *ma-faðaq* *pa-tun-ruða* *masa* *pitʔia.*
AF-know CAUS-ride-boat and cook
'I know how to steer a boat, and to cook.' (5-8)

(3) *ma-faðaq* *ma-didir* *paðay* *masa* *q<m>aʃiʃi* *ðaʃuq.*
AF-know AF-husk grain and sift<AF> rice
'I know how to husk rice and sift rice.' (5-2)

The coordinator is not always required. In fact, it is often optional, as in:

(4) *uka-yða* *sa*[4] *ina* *ama,* *numa* *ma-kasim* *a* *t<un>maða.*
not.have-ALR DET mother father then STA-sad LIG hear<AF>
'Having no mother or father, then (one is) sad listening (to the songs)' (2-3)

(5) *numa* *maniun* *ya* *miku-ki-náy,* *u-tusi-wan* *maniun* *mita* *a* *taun,*
then you.pl if want-stay-here go-there-please you.pl our LIG house

m-ara *sa* *fa-finʃiq* *a* *paðay,* *buna,* *ɬari,* *suksuk,* *makamun*
AF-bring DET RED-sow LIG rice sweet.potato taro ginger pepper

pu-náy *maɬuyða.*
CAUS-here AF.plant
'If you (pl) like to live here, go to our houses to bring the seeds of rice, sweet potatoes, taros, ginger (and) pepper to plant here.' (4-21)

Notice that there is no coordinator between *ina* 'mother' and *ama* 'father' in (4), nor is there any coordinator to conjoin a series of nouns in (5) above. The coordinator *masa* does not appear in F. K. Li's two texts.

2.2 The conjunctives *numa* 'then' and *numawan* 'therefore'

The conjunctive *numa* 'then' is used to conjoin two larger units, such as two clauses, as in (4) above, and two sentences, as in (5) above. It appears very frequently in the free running texts, mostly in the sentence- or clause-initial position, conjoining the sentences in each text. It appears in almost all the Thao texts I have collected, even in the shortest text with only three sentences, namely Text 2. It appears very frequently in some texts, e.g., 96 times in 202 sentences in Text 34. In other words, it appears almost in every other sentence in that text. It appears even more frequently in a few other texts (Texts 4, 11 and 28). It is rare that it does not appear in the whole text at all, as in Text 13 with 14 sentences and

[2] Nonstandard abbreviations (those not included in the Leipzig Glossing Rules) used: AF, Agent-focus; ALR, already; INCH, inchoative; IRR, irrealis; LIG, ligature; COLL, collective; PF, Patient-focus; RED, reduplication; STA, stative.

[3] All examples are based on the Thao texts I have collected: 5-7 indicates that the example is taken from Text 5 and Sentence No.7 (Li 2011). Minor changes were made in a few examples, as suggested by my informant, Kilaʃ (2013/8/1).

[4] The particle *sa* is the most commonly used determiner in Thao (Wang 2004:314). Its shortened form is *s*.

Text 19 with 11 sentences. It appears 655 times out of a total of some 1,900 sentences, roughly once in every 3 sentences (2.9, to be precise).

Similarly, the conjunctive *numa* 'then' appears 10 times in 30 sentences in F. K. Li's (1956) first text and 4 times in 12 sentences in his second text; that is, it appears precisely once in every three sentences.

There is some overlap between the conjunctive *numa* and coordinator *masa* in Thao. The form *numa* can also function as a coordinator between two nouns (as in (6)), two verbs (as in (7)) or two verb phrases (as in (8)) although there is already a coordinator *masa* 'and', as illustrated in (1)-(3) above.

(6) *paʃtay* *mu-náy* *sa* *ayuði* *numa* *aðaðak* *a-tiʃ-an* *maqa*
all come-here DET man and child IRR-rub-LF so.that

a-ma-kaliŋkin *sa* *but.*
IRR-AF-healthy DET body
'All men and boys must come to have (their sleeves) rubbed to stay healthy.' (8-10)

(7) *numa* *i-náy-ða* *yaku* *tu*[5] *min-faðaq* *maʃ-babiar* *numa* *min-paraw*
then LOC-here-ALR I TU INCH-know strike-pestle and MIN-dance

numa *ma-qa-quyaʃ.*
and AF-RED-song
'And here I learned how to strike (the ground) with a pestle, dance and sing.' (1-5)

(8) *numa* *sa* *ma-puði-puði* *a* *qnuan* *mu-saðum,* *ma-kuna* *maŋ-kaθi,*[6]
then DET STA-RED-white LIG deer go-water AF-go.to MAN-there

numa *uka-yða* *ʃaktu-n.*
and/then not-ALR visible-PF
'Then the white deer went into the water, going over to the other side, and then disappeared.' (4-7)

Notice that the conjunctive *numa* appears between every two verbs, as in (7) above. Also notice that the function of the second *numa* in (8) above is ambiguous as to whether it is a conjunctive or a coordinator. The ambiguity makes it possible for the conjunctive and coordinator to overlap. Tsai (2007) postulated that coordinate structures evolve into subordinate structures in both Chinese historical syntax and Formosan languages, such as Tsou, Amis and Atayal. If that is so, *numa* in a subordinate clause in Thao, as in (8), may have functioned as a coordinator 'and' first and then as a subordinate marker 'then' later.

However, *numa* functioning as a coordinator appears much less frequently, altogether only 68 times in all my Thao texts with a total of some 1,900 sentences; it nearly appears once in every 30 sentences. Its normal function is still a conjunctive, conjoining clauses or

---

[5] The particle *tu* is another determiner with different but more restricted distributions than *sa*. It precedes an NP or nominalized clause. It does not co-occur with personal or demonstrative pronouns (Wang 2004:322-26). But it may even precede *sa* to indicate an emphasis, e.g., *tu sa suma wa anyamin* 'It's the other people's stuff'.

[6] Thao has numerous prefixes, including *malan-, maŋ-, mataŋ-*, *min-, mun-* and *tiŋ-*, as appeared in the examples. It would take up a lot of space to give precise glosses for each of these in the interlinear glosses. Moreover, the functions of some of the prefixes remain unclear. Readers may refer to Blust (2003:91-203) for a detailed discussion of Thao affixes.

sentences at this stage. On the other hand, the coordinator *masa* never functions as a conjunctive, conjoining two larger units as does *numa*. The coordinator *masa* 'and' appears 84 times in all the texts.

There may be a competition between these two forms, *numa* and *masa*, with the same function of coordination. Does this indicate that the coordinator *masa* is phasing out and that the conjunctive is taking over the function of coordination and will eventually replace *masa*, or the other way around? Can we predict what will take place? These questions are intriguing.

The conjunctive *numa* may also show a contrast, as in:

(9) *yutu s suma wa θaw, numa suma wa θaw ʃ<uŋ>kaʃ.*
there DET some LIG erson and/then some LIG person fear<AF>
'Some people stayed there, while the others were afraid.' (37-25)

(10) *sa suma ʃ<uŋ>kaʃ mapa-ʔuʃa, numa sa suma mapa-ʔu-ʃuqiʃ.*
DET some fear<AF> COLL-go and/then DET some COLL-go-return
'Some were frightened and ran away, while the others returned.' (37-26)

It is not clear whether to gloss *numa* as 'and' or 'then' in the examples (9) and (10) above. But it is clear it indicates a contrast, functioning as 'while' in English.

A derivative of *numa* 'then' is *numawan* 'therefore'. Like *numa*, *numawan* also appears in the sentence- or clause-initial position and functions as a conjunctive. But the derivative appears much less frequently than the root form in the texts: It appears only once in all Puni's texts, while it appears 58 times in all Kilaʃ's texts, as illustrated in (11)-(12) and (13). The total occurrence of *numawan* is only 59 times in all my Thao texts, as compared with that of *numa*, which is 655 times.

(11) *m<in>aɬi-ʔaðaðak-iða makθin a qali, m-uʃa-yða mu-buhat.*
AF<PFV>give.birth-child-ALR ten LIG day AF-go-ALR go-work
'Ten days after having a baby, (the mother) goes to work.' (10-16)

(12) *numawan antu ma-kaliŋkin sa but.*
therefore not STA-strong DET body
"(She is), therefore, not healthy." (10-17)

(13) *numa puʃa-n yamin, numawan manʃa sa funfun ɬay itan.*
then release-PF we therefore AF.give.gift DET seeds give we/OBL
'Then we released him, so they brought seeds as gifts to us." (33-34)

The derived form, *numawan*, occasionally functions like the root form, *numa*, as in:

(14) *ya ma-ʔania, numawan ðay-n sa θaw, a-malan-ta-tnur.*
if STA-capable therefore ask-PF DET people IRR-MALAN-RED-practice
'If capable (of shooting, they) were then asked to compete by the people.' (39-51)

In fact, it is not always clear whether to interpret the conjunctive as 'then' or 'therefore', as in (14) above. As suggested by Wayan Arka (p. c.), *numawan*, like *numa*, has a rather broad meaning, allowing for both sequential and causal relations.

2.3 The conjunctive *maqa* 'because, so that'

Another conjunctive in Thao is *maqa* 'because, so that'. As implied in the meanings of the form, it has restricted distribution, as it much depends on the context. Like the other conjunctives *numa* and *numawan*, *maqa* occurs in the sentence- or clause-initial position. It appears 274 times in all the texts.

(15) *maqa ya s<m>apuk sa łkaribuʃ, ma-θuaw undu-an.*
because when catch<AF> DET wild.animal STA-very capable-LF
'Because (he) was very capable when (he) was catching wildanimals' (32-9)

(16) *maqa a-saran-an ya mu-ʃuqiʃ, pa-tqa-tqal-an.*
so.that IRR-road-LF when AF-return CAUS-RED-mark-LF
'(They) kept making marks so that (they) could find their way back.' (4-5)

(17) *numa iðáy ła-ławʃin a fafaw quθquθ-an minabút, maqa antu*
then that RED-swing LIG above tie-LF straw so.that not

a-m-utun, maqa a-ma-daqri ya iðáy a fa ławʃin-an.
IRR-AF-break so.that IRR-STA-smooth if that LIG creeper swing-LF
'Then the top of this swing is tied with rice straw, so that (it) won't break, so that this creeper swing will be smooth when swung.' (6-9)

2.4 The conjunctive *ya* 'if, when; complementizer'

A conditional sentence is indicated by the form *ya* 'if, when', as in (14)-(17) above. It conjoins two clauses in a sentence. A tricky problem about the form is that it is not always clear whether we should translate it as 'if' or 'when' in some context, as illustrated in (18)-(19).

(18) *maqa ma-tiŋ-qaran mun-ławʃin, maqa antu ma-qirið ya*
so.that STA-TIN-happy MUN-swing so.that not STA-tired if/when

mu-buhat.
go-field
'To go swinging so that one is happy, so that one does not get tired if/when working in the field.' (6-12)

(19) *numa ya m<in>u-lalu-yða, k<m>an afu.*
then if/when AF<PFV>hold-ceremony-ALR eat<AF> rice
'If/When the ceremony is over, (we) eat cooked rice.' (8-14)

Both conjunctives 'if' and 'when' head a subordinate clause in English and they have closely related meanings. Thao speakers may not distinguish between the two meanings.

In some context, it seems clear that *ya* functions as 'if', as in:

(20) *m-ayaw sa θaw ya antu mu-tantu.*
STA-shy DET person if not go-there
'(I would have been) sorry to people if (I had) not gone there.' (13-5)

In addition to being a conditional marker, the form *ya* seems to function as a complementizer preceding a nominalized clause,[7] as in the following examples:

(21) *myaðáy a malan-ta-tnur t<m>aðam ya tima sa*
in.that.way LIG MALAN-RED-practice test<AF> COMP who DET

ma-ʔania panaq sa iðáy ɬpaðiSan. (39-50)
STA-capable shoot DET that arrow
'For such a practice, it was to test to see who was capable of shooting with an arrow.'

(22) *maqa tu kahiwan ma-θuaw antu pidara ya s<m>iwsiw,*
because DET formerly STA-very not pardon COMP do.magic<AF>

antu ma-rarabaða.
not STA-joke
'People in the past took a magic (a ritual ceremony) very seriously; they did not allow it to be slighted.' (39-57)

(23) *iðáy binanawʔað, ya itiða sa rariɬið, ʃi-na-haðiʃ-uan, kanun-iða*
that woman when arrive DET plate go-to-far-please in.case-ALR

suma wa binanawʔað ya ɬquðan.
some LIG woman COMP pregnant
'When the wooden plate of ancestors arrives, women should stay away, in case someone should be pregnant.' (39-211)

(24) *numa mataŋ-kaktun-iða, parʃian ya ma-qa-quyaʃ ʃa-ʃayla*
then MATAN-finish-ALR taboo COMP AF-RED-sing RED-perform

wa quyaʃ. (39-236)
LIG song
'When all is over, to keep singing and dancing for the ritual ceremony is forbidden.'

It requires further study to determine the various grammatical functions of *ya*.

3 Summary and discussion

I have discussed the coordinating conjunction in Thao discourse and grammar. All examples are based on the 42 Thao texts with a total of some 1,900 sentences I collected. The Thao texts cover a wide range of topics, with stylistic differences by different speakers.

---

[7] Blust (2003:1057) also treats *ya* as two separate lexical entries, one as "conditional: if" and the other as "grammatical particle of uncertain function." Wang (2004:326-28) treats a form of *ya* as a determiner, which co-occurs with a time phrase to designate a future (1) or habitual (2) interpretation, or a determiner with an oblique/locative function (3), as below:

(1) *a muʃa iDa yaku ya saqaDi.* 'I'll leave this noon' (Blust 2003:1043)
(2) *ya ʃaʃanu maθuaw waDaqan maharbuk.* 'In the morning the lake is very foggy' (Blust 2003:398)
(3) *hadana ita iDa ya aDaDak.* 'Let's adopt a child' (Blust 2003:393)

Thao conjunctives include *numa* 'then', its derivative *numawan* 'therefore', *maqa* 'because, so that', and *ya* 'if, when', which seems to function as a complementizer also.

The coordinating conjunction *numa* 'then' appears frequently in the free running texts, mostly in the sentence-initial or clause-initial position, linking the sentences in each text. It appears in all the texts, even in the shortest one with only three sentences. The same form can also function as a coordinator between two nouns, two verbs or two verb phrases, although there is already a coordinator *masa* 'and' in this language. These two functions, glossed as 'then' and 'and', seem to be closely related syntactically and semantically. It will be interesting to find out how similar or dissimilar the functions of the Thao form *numa* are to other Formosan languages. For instance, the conjunction *la* 'then' links two clauses or sentences in all Rukai dialects except Mantauran, but not between two nouns, in which case *si* 'and' is used (Li 1975:1), while *mani* 'then' and *la* 'and' are used respectively in Mantauran (Zeitoun 2007:452-56).

Another conjunction in Thao is *ya* 'if, when', which conjoins two clauses: I have discussed some problems of analyzing the conjunction. One problem is that it is not always clear whether to gloss *ya* as 'if' or 'when' in some context. The same is found in other Formosan languages also.

It will be interesting to compare forms with similar functions in the other Formosan languages.

Appendix. Number of sentences and occurrences of conjunctives in Thao

In the following table, information is given for the total number of sentences and occurrences of *numa* 'then' or 'and', *numawan* 'therefore', *maqa* 'because', and *masa* 'and' in each text. Texts 1-30 except 4 in Part I are based on Puni, while Texts 31-42 in Part II are based on Kilaʃ. In fact, the total number of sentences in the texts should be larger than given in Table 1. A quote, for instance, may contain more than a sentence in the texts.

Table 1: Number of occurrences of the Thao conjunctives

Part I Text No.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
No. of sentences	5	3	11	24	15	14	10	18	12	18
occur. of *numa* 'then'	1	1	4	15	1	4	3	5	3	8
occur. of *numa* 'and'	2			1				3	3	
occur. of *numawan* 'therefore'										1
occur. of *maqa* 'because'				1		5	1	2	2	3
occur. of *masa* 'and'					3					

Part I Text No.	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
No. of sentences	7	8	13	28	14	11	24	18	11	17
occur. of *numa* 'then'	3	1		6	5	7	6	3		2
occur. of *numa* 'and'			1	3	1					
occur. of *numawan* 'therefore'										
occur. of *maqa* 'because'	1	2	2		1	3	6		1	
occur. of *masa* 'and'										1

Part I Text No.	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**
No. of sentences	16	27	22	16	17	18	30	11	9	18
occur. of *numa* 'then'	8	9	6	4	1	6	1	7	3	7
occur. of *numa* 'and'		1					1			
occur. of *numawan* 'therefore'										
occur. of *maqa* 'because'	3	6	4	1		1	6	1	2	2
occur. of *masa* 'and'						1				1

Part II Text No.	**31**	**32**	**33**	**34**	**35**	**36**
No. of sentences	50	99	36	202	61	98
occur. of *numa* 'then'	21	46	24	96	12	31
occur. of *numa* 'and'	1			1		9
occur. of *numawan* 'therefore'	3		5	1	5	8
occur. of *maqa* 'because'	4	6	4	11	2	21
occur. of *masa* 'and'	1	2	1	27	2	6

Part II Text No.	**37**	**38**	**39**	**40**	**41**	**42**	**Total**
No. of sentences	154	178	236	117	146	36	1,878
occur. of *numa* 'then'	41	70	83	38	59	4	655
occur. of *numa* 'and'	5	12	8	11	5		68
occur. of *numawan* 'therefore'	13	4	6	9	4		59
occur. of *maqa* 'because'	32	25	43	28	34	8	274
occur. of *masa* 'and'	6	7	15	16	4		93

References

Blust, Robert. 2003. *Thao Dictionary.* Language and linguistics monograph series No.5. Taipei: Academia Sinica.

Gärtner, Hans-Martin, Paul Law, and Joachim Sabel, eds. 2006. Clause structure and adjuncts in Austronesian languages: A critical introductory survey. In Gärtner, Hans-Martin, Paul Law and Joachim Sabel (eds.), *Clause structure and adjuncts in Austronesian languages*, 1-42. Berlin: Mouton de Gruyter.

Kuo, Ching-hua. 2008. Thao texts. In Seiya Abe, Yoshio Nagashima, and Sumino Niida, eds., *Thao Vocabulary with English and Japanese Indices: Thao Language Studies II*, 261-320. Research Institute for Oriental Cultures, Gakushuin University, Occasional Papers No.54.

Li, Fang Kuei *et al*. 1956. Notes on the Thao language (in Chinese). *Bulletin of the Department of Archaeology and Anthropology* 7:23-51.Taipei: National Taiwan University.

Li, Paul Jen-kuei. 1975. *Rukai Texts*. Taipei: Institute of History and Philology, Academia Sinica Special Publications, No.64-2.

Li, Paul Jen-kuei. 2011.*Thao Texts and Songs*. Language and linguistics monograph series No.44. Taipei: Academia Sinica.

Tsai, Dylan Wei-tian. 2007. Conjunctive reduction and its origin: A comparative study of Tsou, Atayal, and Amis. *Oceanic Linguistics* 46.2:585-602.

Wang, Shan-shan. 2004. *An Ergative View of Thao Syntax.* Unpublished PhD dissertation, University of Hawaii at Manoa.

Zeitoun, Elizabeth. 2007. *A Grammar of Mantauran (Rukai)*. Language and Linguistics Monograph Series A4-2, Institute of Linguistics, Academia Sinica.