

*The many faces of Austronesian
voice systems: some new
empirical studies*

Pacific Linguistics 571

Pacific Linguistics is a publisher specialising in grammars and linguistic descriptions, dictionaries and other materials on languages of the Pacific, Taiwan, the Philippines, Indonesia, East Timor, southeast and south Asia, and Australia.

Pacific Linguistics, established in 1963 through an initial grant from the Hunter Douglas Fund, is associated with the Research School of Pacific and Asian Studies at The Australian National University. The authors and editors of Pacific Linguistics publications are drawn from a wide range of institutions around the world. Publications are refereed by scholars with relevant expertise, who are usually not members of the editorial board.

FOUNDING EDITOR: Stephen A. Wurm

EDITORIAL BOARD: John Bowden, Malcolm Ross and Darrell Tryon (Managing Editors),
I Wayan Arka, David Nash, Andrew Pawley, Paul Sidwell, Jane Simpson

EDITORIAL ADVISORY BOARD:

Karen Adams, *Arizona State University*
Alexander Adelaar, *University of Melbourne*
Peter Austin, *School of Oriental and African Studies*

Byron Bender, *University of Hawai'i*
Walter Bisang, *Johannes Gutenberg-Universität Mainz*

Robert Blust, *University of Hawai'i*
Lyle Campbell, *University of Utah*
James Collins, *Universiti Kebangsaan Malaysia*

Bernard Comrie, *Max Planck Institute for Evolutionary Anthropology*
Soenjono Dardjowidjojo, *Universitas Atma Jaya*

Matthew Dryer, *State University of New York at Buffalo*

Jerold A. Edmondson, *University of Texas at Arlington*

Nicholas Evans, *University of Melbourne*

Margaret Florey, *Monash University*

William Foley, *University of Sydney*

Karl Franklin, *Summer Institute of Linguistics*

Charles Grimes, *Universitas Kristen Artha Wacana Kupang*

Nikolaus Himmelmann, *Ruhr-Universität Bochum*

Lillian Huang, *National Taiwan Normal University*

Bambang Kaswanti Purwo, *Universitas Atma Jaya*

Marian Klamer, *Universiteit Leiden*

Harold Koch, *Australian National University*

Frantisek Lichtenberk, *University of Auckland*

Patrick McConvell, *Australian Institute of Aboriginal and Torres Strait Islander Studies*

William McGregor, *Aarhus Universitet*

Ulrike Mosel, *Christian-Albrechts-Universität zu Kiel*

Claire Moyse-Faurie, *Centre National de la Recherche Scientifique*

Bernd Nothofer, *Johann Wolfgang Goethe-Universität Frankfurt am Main*

Ger Reesink, *Universiteit Leiden*

Lawrence Reid, *University of Hawai'i*

Jean-Claude Rivierre, *Centre National de la Recherche Scientifique*

Melenaite Taumoefolau, *University of Auckland*

Tasaku Tsunoda, *University of Tokyo*

John Wolff, *Cornell University*

Elizabeth Zeitoun, *Academica Sinica*

The many faces of Austronesian voice systems: some new empirical studies

Edited by

I Wayan Arka and Malcolm Ross



Pacific Linguistics
Research School of Pacific and Asian Studies
The Australian National University

Published by Pacific Linguistics
Research School of Pacific and Asian Studies
The Australian National University
Canberra ACT 0200
Australia

Copyright in this edition is vested with Pacific Linguistics

First published 2005

National Library of Australia Cataloguing-in-Publication entry:

I Wayan Arka and Malcolm Ross, eds
The many faces of Austronesian voice systems: some new empirical studies

Bibliography
ISBN 0 85883 556 8.

1. Austronesian languages – Voice. I. Arka, I Wayan.
II. Ross, Malcolm (Malcolm D.). III. The Australian National
University. Research School of Pacific and Asian Studies.
Pacific Linguistics. IV. International Conference on
Austronesian Linguistics (9th :, 2002 : Canberra, ACT).
V. International Conference on Oceanic Linguistics
(5th :, 2002 : Canberra, ACT).

499.2

Copyedited by Basil Wilson
Typeset by Jeanette Coombes
Cover design by Cirils Printers
Printed and bound by Cirils Printers, Fyshwick, Canberra

Table of contents

1	Introduction <i>I Wayan Arka and Malcolm Ross</i>	1
2	Kelabitic languages and the fate of ‘focus’: evidence from the Kerayan <i>Beatrice Clayre</i>	17
3	The Palu’e passive: from pragmatic construction to grammatical device <i>Mark Donohue</i>	59
4	Passive without passive morphology? Evidence from Manggarai <i>I Wayan Arka and Jeladu Kosmas</i>	87
5	Two passive-like constructions in Tongan <i>Yuko Otsuka</i>	119
6	Grammatical relations in Puyuma <i>Stacy Fang-ching Teng</i>	137
7	Voice, ergativity and transitivity in Tagalog and other Philippine languages: a typological perspective <i>Masumi Katagiri</i>	153
8	Aspects of pragmatic focus in Tagalog <i>Daniel Kaufman</i>	175
9	The richness of Madurese voice <i>William D. Davies</i>	197
10	Topic continuity, voice and word order in Pendau <i>Phil Quick</i>	221
11	Semantic analysis of the Moronene verbal prefix <i>moN-</i> <i>Suree Andersen and T. David Andersen</i>	243

1 *Introduction*

I WAYAN ARKA AND MALCOLM ROSS

1 Preliminaries

The Ninth International Conference on Austronesian Linguistics and the Fifth International Conference on Oceanic Linguistics were both held at The Australian National University in Canberra during January 2002. Rather than publish a single very diverse collection of conference papers, the organisers favoured a series of smaller compilations on specific topics. One such volume, on Austronesian historical phonology, has already been published as Lynch (2003). The present volume represents another such compilation, although three of the papers (those by Davies, Donohue and Teng) were not presented at the conferences.

The Austronesian voice (or ‘focus’) system has been widely acknowledged as a significant challenge to historical, descriptive and typological linguistics, as well as to theoretical syntax. Since the publication of Schachter (1976), who questioned the universality of the grammatical relation ‘subject’ from a Philippine perspective, there has been much research on Austronesian voice, with data from both Philippine and non-Philippine languages, for example the collection on papers in Wouk and Ross (2002). This volume, in a sense a follow-up to that volume,¹ contains ten papers in the area of Austronesian voice which provide fresh data and some new perspectives on old problems. To our knowledge, no data from two of the ten languages, Palu’e and Moronene, have been published previously, whilst data from Puyuma, Madurese, Pendau and Manggarai are not easily accessible. In this introductory chapter, we present a brief overview of each paper (§2), and then comment on the significance of these papers in relation to research on Austronesian voice systems (§3).

A note on terminology: our authors vary in their chosen terminologies, and we have made no attempt to bring them into line with each other. Indeed, there are good reasons for many of their differences. For the purposes of this introduction, however, we will talk

¹ A ‘follow-up’ both in the thematic sense and in the sense that Wouk and Ross represents a thematic compilation of papers from a workshop at the Eighth International Conference on Austronesian Linguistics held in Taipei in 1997.

about subjects (rather than ‘topic’ or ‘pivot’), about voice (rather than the Philippinists’ ‘focus’), about the semantic roles of actor and undergoer, and about actor and undergoer voices (preferring ‘undergoer’ to ‘patient’, as Philippine-type voice systems can often be said to make a primary contrast between actor and undergoer voices, with more than one undergoer voice, one of which is the patient voice; see, for example, Himmelmann 2005).

2 The papers

Most of the papers in the volume are descriptive, though often with theoretical insights, and each mainly concerns a single language. One paper is an exception to this generalisation. **Beatrice Clayre** offers an areal survey with historical implications. She outlines the voice system in Kelabitic languages (Lun Dayeh, Kerayan dialects and Sa’ban), a subgroup of the North Sarawak group, describing and comparing their verbal voice and aspect affixes, the marking of noun phrases, the number and function of pronoun sets, and word order. She shows that a relatively elaborate Philippine-type voice (‘focus’) system is encountered in Lun Dayeh in the north of the area, but has become a more simplified system in all of the southern dialects of Kerayan and Sa’ban. Lun Dayeh has three voices (actor, patient—Clayre’s ‘undergoer’—and instrument) with a remnant of the locative/benefactive voice, appropriate verb affixation and three separate pronoun sets showing distinct encoding of subject vs non-subject on one hand and actor vs non-actor roles on the other. Sa’ban and all other Kerayan dialects have only two voice types: actor and undergoer. Verb affixation and pronoun sets are also reduced. Pronominal forms showing certain distinctions of grammatical relations have changed or been lost. For example, set I pronouns, which mark subject in Lun Dayeh, mark core relations in Sa’ban. The Lun Dayeh set II pronouns are almost gone in Sa’ban. With set II pronouns gone, similar proclitics corresponding to Lun Dayeh set III pronouns are found only in certain Kerayan dialects. The verbal morphology has been much simplified, e.g. the reduction of the actor-voice prefix to *N-*, which is already gone in certain dialects (i.e. the actor-voice verb has no actor-voice prefix), yielding a more isolating structural type. The demise of voice morphology correlates with the increasing significance of periphrastic constructions and relatively fixed word order (e.g. the emergence of a verb phrase where a non-subject core argument must follow the verb). However, fossilised remnants of morphologically marked voice are still found in a small group of conservative verbs suggesting earlier versions of the dialects.

The remaining papers in the volume are largely descriptive. The next two concern Palu’e and Manggarai respectively, languages of the Indonesian island of Flores which both reflect an isolating tendency like that of Sa’ban and the Kerayan dialects but carried significantly further.

Mark Donohue discusses voice in Palu’e and argues that Palu’e is a language that has just begun grammaticalising a topic construction into a passive voice, thereby providing a piece of evidence for an origin of voice systems in pragmatic structuring devices. While reflexive binding data is inconclusive, a number of other tests (adverbial placement, quantifier float, conjunction reduction, purposive clause) clearly show that AVP (Agent-Verb-Patient) and PAV constructions are two grammatically different constructions in Palu’e. The P in the PAV structure is grammatically subject and is not therefore simply a preposed or topicalised P of the corresponding active AVP structure. Donohue concludes that the PAV construction is an instance of passive, despite lack of passive morphology. He

asks whether this unusual passive is a start towards a more elaborate system or a relic of an earlier more elaborate system.

I Wayan Arka and **Jeladu Kosmas** discuss a passive construction in Manggarai. They argue that, despite the absence of voice morphology, syntactically Manggarai has a passive, in which the agent appears with the preposition *l(e)* ‘by’. After outlining basic grammatical relations in Manggarai (isolating, canonical SVO order with a possible VOS structure and possibly with enclitic subject), they give evidence to support the idea that the *l(e)* construction is passive. Firstly, the backgrounded Agent marked by *l(e)* is indeed syntactically oblique, even though it is generally obligatorily present. Secondly, there is a word-order change, accompanied by a change in pronominal subject enclitic that is no longer in agreement with the backgrounded agent. This indicates that the *le* agent is not the subject in the *le* construction. Thirdly, reflexive binding further supports the claim that the *le* Agent is Oblique. Arka and Kosmas show that reflexive binding in Manggarai is sensitive to a surface grammatical relations hierarchy where the reflexive must be syntactically outranked by the binder. Thus, the *le* agent cannot bind a reflexive subject but a reflexive *le* Agent can be bound by a non-agent subject. Arka and Kosmas also discuss issues associated with the analysis in a wider typological context.

Although Tongan is geographically and genealogically far from the other languages discussed in this volume, it also has a morphologically isolating tendency which has led to analytical controversy. **Yuko Otsuka** shows that Tongan is ergatively aligned, despite morphosyntactic similarities with the accusative languages of eastern Polynesia. She also shows that the Tongan reflexes of the famous *-Cia* verbal form are not passive in Tongan, unlike their cognates in various other Polynesian languages. The focus of her discussion, however, is two constructions that appear passive-like and might each appear to be a passive: the agentless passive construction and the VOS construction. She shows on the basis of syntactic criteria, however, that despite appearances neither of these constructions is passive. Her tests include the use of possessor pronouns with nominalisations, the distribution of clitic pronouns, and behaviour under coordination with the conjunction *mo*. Instead, she suggests that the agentless transitive construction is precisely that—a construction with an unspecified agent — and the VOS construction is a permutation of the default VSO construction which expresses a particular kind of focus.

At the opposite extreme of complexity among Austronesian languages is the voice system of Puyuma, a language of southern Taiwan, described by **Stacy Fang-Ching Teng**. While there are four different voices — actor, patient, locative and instrumental/beneficiary — not all verbs exhibit all four voice alternations. Teng shows that there are asymmetries between the actor voice and non-actor (i.e. undergoer) voice sentences in terms of verbal morphology and pronominal clitic marking. There is a unique affix for each undergoer voice type, e.g. *-aw* for patient, *-ay* for locative, *-anay* for instrumental/beneficiary, and there are at least five affixes for the AV actor voice (*ma-*, **, *mu-*, *mi-* and zero). There are two clitics on an undergoer voice verb, a genitive actor proclitic and a nominative (subject) enclitic but there is only a single clitic, namely the nominative (subject) enclitic, on an actor voice verb. The third person genitive and nominative clitics can often be cross-referenced by free argument NPs, which receive different case markings depending on their grammatical functions (subject or oblique), noun classes (common, proper, or location nouns), definiteness, and number. Puyuma is also shown to have complex free pronominal forms, distinguished by number and the associated grammatical functions (neutral or oblique, possessor of subject or non-subject, and definiteness). Teng

discusses the issues of transitivity and grammatical system alignment in Puyuma and concludes that Puyuma is morphosyntactically ergative as far as independent clauses are concerned but accusative in regard to relative clause and serial verb constructions.

Moving southwards into the Philippines, the next two papers concern Tagalog. **Masumi Katagiri** re-examines data from Tagalog and other Philippine languages and discusses some of the issues surrounding the ergative hypothesis from a typological perspective. Based on Dixon's (1994:146) antipassive criteria, she argues that the actor voice construction is not an antipassive as there is no good evidence that it is derived from the patient voice. Actor voice verbal morphology, e.g. Tagalog *mag-*, is arguably not a unique formal marking of antipassive-intransitive derivation because it is morphologically indistinguishable from the intransitive *mag-* verb which differs from it only in the absence of an undergoer. Crosslinguistically, as in Dyirbal (Australia), the two kinds of intransitive are marked differently. Moreover, the claim that the patient voice verb is more basic than the actor voice verb does not always get support because the morphological structure of *mag-* verbs in Tagalog is not necessarily more complex than the patient voice verbs. In fact, certain patient voice verbs are marked for their voice while the corresponding actor voice verbs are unmarked. Furthermore, there is conflicting evidence as to the syntactic status of the undergoer of an actor voice clause: it ought to be oblique under the antipassive analysis. However, adjunct fronting and participial adjunct tests show that the genitive-marked undergoer argument of the actor voice construction is a core (or term) argument. However, Katagiri shows that a dative-marked undergoer associated with certain actor voice verbs appears to be an oblique as two tests for core status in Tagalog suggest that it is demoted to Oblique. Crucially, a serious problem arises because a verb may allow a genitive-marked undergoer or a dative-marked undergoer, the choice of which is governed by the specificity of the undergoer: a specific undergoer is dative-marked whereas a non-specific undergoer is genitive-marked. Katagiri concludes that there is no plausible morphological or syntactic evidence to support an ergative analysis for the Tagalog voice system. She proposes instead that Tagalog shows signs of a split alignment system, with the actor voice construction representing an accusative construction and the undergoer voice construction an accusative construction. The conditions for the split are the definiteness and affectedness of the undergoer and tense/aspect/mood. For example, the undergoer voice construction is typically used, sometime obligatorily, when the undergoer is definite or highly affected, or when the construction is in realis or perfective aspect. However, there is variation among the Philippine languages as to how these conditions apply.

Daniel Kaufman discusses the *pragmatic* relations of focus and topic in Tagalog (as opposed to 'focus' and 'topic' in their Philippinist sense). He shows that Tagalog possesses regular syntactic expressions of focus and topic. The two are highly grammaticalised in Tagalog. He further argues that prosody functions as a back-up alternative in marking the focus when the syntax is unable to do so. Illustrations are given of constructions with double focus, e.g. where the effects of the syntactic focus are pragmatically inappropriate. Based on observation of the interaction between the syntax and the prosody of topic and focus, he explains previously unaccounted cases of 'scrambling' whereby one variant constituent order is preferred over another in particular pragmatic contexts.

William Davies presents data from Madurese, a language of the west Java region, arguing that its ostensible two-voice (actor/undergoer) system is actually as rich as those of the Philippines. He proposes an ‘extended voice system’ analysis for Madurese whereby the derivational affixes *-agi* and *-e* (usually each identified along with their Javanese, Balinese and Indonesian counterparts as both causative and applicative) are considered part of the voice system. He demonstrates how argument mappings (semantic roles and their surface grammatical relations) in Madurese and Tagalog are basically the same. The key difference is in the morphology involved. For example, when the beneficiary is linked to subject in Tagalog, a single morphological rule marked by *ipag-* is required. In contrast, the same process requires two morphological rules in Madurese, marked by the two affixes *e-* and *-agi*. The analysis is also supported by evidence from connected discourse, which shows a high percentage (80%) of verbs with *-agi* (or *-e*) in undergoer voice. Comparative data from Madurese and Cebuano texts shows a striking distributional similarity between the undergoer voice in Madurese (including *-agi* and *-e* forms) and the undergoer voices in Cebuano.

Phil Quick deals with two major voice forms, *nong-* verbs and *ni-* verbs, in Pendau. He argues that *nong-* verb and *ni-* verb clauses are both transitive, and he analyses them as active and inverse voice respectively. Quantitative evidence and criteria are presented to support the analysis. He shows that relative topicality or continuity of actor and undergoer is crucial for the selection of a *nong-* verb or a *ni-* verb construction. When the actor argument is more topical (manifests greater topic continuity) than the undergoer argument, then the *nong-* verb construction tends to be chosen, otherwise the *ni-* verb construction serves as the default. Quick opposes the analysis of *ni-* verbs as passive because they are syntactically transitive and the actor argument is highly topical and rarely omitted: these are not characteristics of passives. His investigation also reveals that, whatever its causes, constituent order variation is not determined by topic continuity in Pendau.

Suree and David Andersen discuss voice alternation in Moronene, a Bungku-Tolaki language of south-east Sulawesi. The discussion focuses on the semantic contrast between clauses in which the undergoer/object is indexed on a plain verb stem by an absolutive clitic (the Verb-ABS construction) and those in which the verb takes the prefix *moN-* (the *moN-*verb construction). The *moN-*verb tends to have an ordinary, general or non-specific undergoer/object. It also appears to follow the antipassive pattern found in other languages with a highly topical agent and a low topicality patient. However, it is possible for the free undergoer noun phrase cross-referencing the absolutive clitic to be indefinite and the undergoer noun phrase of the *moN-*verb clause to be definite. If the undergoer is definite, the choice of *moN-*verb or verb-ABS construction reflects certain semantic differences: the *moN-*verb construction is associated with whole processes, durative/non-punctual, non-volitional, or irrealis events whereas the verb-ABS construction is associated with a particular act/action and is punctual/aspectually bounded, volitional, and realis. Factors relating to individuation such as animacy and plurality are also shown to play a role; e.g. a definite animate undergoer typically co-occurs with a verb-ABS construction whereas a definite inanimate undergoer tends to co-occurs with a *moN-*verb. In Hopper and Thompson’s (1980) terms the verb-ABS construction expresses higher transitivity whereas the *moN-*verb construction serves for lower transitivity.

3 Discussion

3.1 Historical themes

As we noted earlier, the paper with the strongest historical implications is Clayre's. The Kelabitic languages offer us in microcosm a picture of a process which must have occurred independently in a number of different places and at different times. This is the breakdown of a Philippine-type voice system like those of Puyuma (Teng) and Tagalog (Katagiri) into an actor/undergoer voice opposition. The beginnings of the breakdown process are already visible in Lun Dayeh, where the locative voice has fallen out of use, but simplification has proceeded much further in Sa'ban and the Kerayan dialects. The instrumental voice is lost, leaving a simple actor/undergoer contrast and aspect morphology is replaced by an auxiliary. Proto Kelabitic evidently distinguished the actor voice affixes **m-* (from PMP **-um-*), also serving as dynamic intransitive, **N-* (from PMP **maN-*) and zero (from PMP zero; for PMP reconstructions see Ross 2002a:48–51), and had a separate stative intransitive **me-*. The first two have been partially merged in Lun Dayeh (where Clayre's *N-* is a diaform encapsulating lexically determined *m-* and *ng-*). In Sa'ban and the Kerayan dialects these morphological distinctions (which were probably never complete anyway) are being or have already been lost; and the functional load of pronominal case marking is reduced as genitive forms (set II) are displaced by nominative forms (set I) so that the latter come simply to mark core (term) grammatical function.

It is instructive to note that this is not the same breakdown process as has occurred in other places (indeed, it is more extreme). Throughout much of western Indonesia and Malaysia a system like the one described by Davies for Madurese has emerged (cf. for example, Poedjosoedarmo 2002 on Javanese, Quick 2002 on Pendau and Arka 2003 on Balinese). The system has on the surface also been reduced to an actor/undergoer contrast, but applicative suffixes have been recruited to allow the selection of a variety of semantic roles as subject, resulting in a system with the same flexibility as the Philippine systems, as Davies points out. In such systems the genitive/nominative case marking distinction is usually maintained in pronouns (but often not in noun phrases). The emergence of 'Indonesian-type' systems like Madurese has been reconstructed by Wolff (1996; see also Ross 2002a:52–56, 2002b:451–453), whilst Mead (2002) presents Sulawesi variants of this emergence and Ross (2002b:459–464) suggests an overarching morphosyntactic history accounting for most of the languages of Sulawesi.

A historical point of a different kind emerges from Clayre's data. She shows that all the languages she describes have a clausal constituent order in which a non-subject core noun phrase (a full noun phrase) must directly follow the verb. The non-subject core noun phrase is the actor in an undergoer voice construction and the undergoer in an actor voice construction. This symmetry has also been recorded for Balinese (Arka 2003), for Batak (Sumatra; Emmorey 1984; Schachter, ed. 1984; Norwood 2002) and for Pendau (Sulawesi; Quick, this volume). This wide distribution suggests that it was present at an early stage in the history of Indonesian-type languages. It is not apparently a general Philippine feature, however, and this raises the question: Did this innovation occur once in a common ancestor of the Kelabitic languages, Balinese, Batak and Pendau? Or did it occur independently, perhaps with the loss of noun-phrase case marking, in different places? If it occurred only once, then when did it occur? Holmer (2002) and Huang (2002) agree that a similar constituent order is the default in Seediq (northern Taiwan), but this is not true of Puyuma, where constituent order is freer (Stacy Teng, pers. comm.). Did this innovation occur as a rigidification of the order in a language like Seediq? This matter is significant in

the reconstruction of the history of the languages of western Indonesia and Malaysia and deserves further attention.

3.2 Typological themes

3.2.1 Types of Austronesian voice systems

The voice systems of Austronesian languages are typologically quite diverse. There are three major types with regard to voice, namely Types 1, 2 and 4 below. We have added Type 3, whose membership is much smaller but which is represented by two contributions to this book.

1. Languages with multiple voice types, marked by verbal morphology and often accompanied by case marking of free nominal arguments. There is always one actor voice, which is either intransitive or lower in transitivity than the remaining voices, which are conveniently grouped together as undergoer voices. The undergoer voices allow noun phrases with a variety of semantic roles to become subject: patient, theme, location, instrument, beneficiary etc. Such systems have been labelled by Himmelmann (2002) '**Philippine-type**' systems, although they occur not only in the Philippines but also in Taiwan, northern Borneo, northern Sulawesi and Madagascar. Philippine-type languages described in the contributions to this volume are Puyuma, Tagalog and Lun Dayeh. They are assumed to be the most conservative of Austronesian voice systems.
2. Languages conventionally analysed as having two voices, actor and undergoer, supplemented by applicative suffixes which allow locations, instruments, beneficiaries and noun phrases of other semantic roles to become the undergoer. These may be labelled '**Indonesian-type**' languages, although they are located in western Indonesia and Malaysia. Their development from Philippine-type systems is described by Ross (2002a, 2002b). In the more conservative Indonesian-type languages, the undergoer voice continues to be the default transitive, but in some languages the actor voice is competing for this role (see, for example, Wouk 2002 on Sasak). (Conservative) Indonesian-type languages described in the contributions to this volume are Pendau, Moronene and Madurese. Other well-known members of this type are Classical Malay, Balinese (Arka 2003) and Toba Batak (Schachter 1984).
3. Languages with two voice types, active and passive, but not morphologically marked on the verb. This is a characteristic of certain isolating languages of Flores, of which Manggarai and Palu'e are described in this volume.
4. Languages with no voice alternation. These languages could be labelled the 'eastern Austronesian type', as they are located in eastern Nusa Tenggara, New Guinea, Island Melanesia, Polynesia and Micronesia. These languages typically have pronominal subject proclitics or prefixes and object enclitics or suffixes on the verb, and have an applicative derivational morpheme (Pawley 1973; Ross 2004). A free NP may scramble to reflect certain pragmatic information, but scrambling does cause a change in the grammatical relation of the argument(s). No language of this type is discussed

in his volume, as the contributions are concerned with voice systems. Descriptions of such languages are relatively plentiful.²

These four types of course by no means exhaust the voice-related types of Austronesian language. Sa'ban and the Kerayan dialects described by Clayre in this volume resemble Type 1 but with only one undergoer voice. Some isolating languages of Flores appear to lack voice, but also lack the morphology attributed to Type 3, e.g. Nage-Keo (Baird 2002). A few languages which otherwise belong to Type 4 have innovated a passive voice. One is Roviana (northwest Solomons; Corston 1996). Others are found in Micronesia. Polynesian languages are aberrant members of Type 4. Some of these have a passive voice, but the one described in this volume, Tongan, has passive-like constructions which are evidently not actually passives.

3.2.2 *Syntactic issues*

Two major groups of syntactic issues arise in the papers in this volume. The first centres on the question: How should Austronesian voice systems be characterised in terms of syntactic typology? The second concerns the use of these voice systems in discourse: How does a speaker decide which voice to use? We do not attempt to keep these issues apart in the discussion below, as they are loosely intertwined.

The first set of issues has been a concern at least since Schachter (1976). Davies' observation that the Madurese verbal suffixes *-e* and *-agi* should be regarded, functionally at least, as part of the voice system, effectively takes an earlier suggestion by Starosta (1986) and stands it on its head. Verhaar (1984) also treats Indonesian *-i* and *-kan* as being part of a single system, as they form circumfixes *me- -i* and *me- -kan*. Starosta recalls that applicative suffixes like *-e* and *-agi* have the function of allowing a noun phrase with a semantic role other than patient or theme (e.g. instrument, beneficiary, location) to be 'promoted' to the grammatical function of undergoer.³ He then points out that this is precisely the function of locative and instrument/beneficiary voice morphology in Philippine-type languages. The difference between applicatives as conventionally understood and Philippine-type voice morphology is that an applicative suffix typically promotes an oblique noun phrase to object (the undergoer grammatical function in an accusative language) whereas, at least in the default (undergoer voice) construction of a Philippine-type language, this promotion is to subject. The functional parallel is clear, however. What is more, as Davies observes, the oblique noun phrase in Madurese is also promoted to subject if the verb is in the undergoer voice. Thus an alternative analysis, not canvassed by any of the contributors to this book (but see Himmelmann 2005), would be to

² For example, Tetun (Timor; Van Klinken 1999), Taba (Maluku, Indonesia; Bowden 2001), Manam (north New Guinea; Lichtenberk 1983), Tawala (southeast Papua; Ezard 1997), Tolai (New Britain; Mosel 1984), Hoava (northwest Solomon Is; Davis 2003), Kwaio (southeast Solomon Is; Keesing 1985), North-East Ambae (northern Vanuatu; Hyslop 2001), Sye (southern Vanuatu; Crowley 1998), Fijian (Dixon 1988), Samoan (Mosel & Hovdhaugen 1992). A collection of sketches of languages mostly of this type is also found in Lynch, Ross and Crowley (2002).

³ It should be noted that *-e/-agi* and similar suffixes in other Indonesian-type languages (e.g. *-i/-kan* in Indonesian and *-in/-ang* in Balinese) may have a range of functions with the applicative function being only one of them. These suffixes commonly also have a causative function, where an actor, rather than an undergoer, argument is introduced. It is unclear and debatable whether they should be regarded as being 'polysemous' or 'homonymous' forms and whether or not they are part of (extended) voice systems.

treat both Indonesian- and Philippine-type languages as having a single undergoer/actor voice alternation (with undergoer as the default voice in many languages), enriched by applicative morphology which allows a location, instrument or beneficiary noun phrase to be promoted to undergoer. The applicative morphemes would be the Madurese suffixes *-e* and *-agi* and their equivalents in other Indonesian-type languages, and the locative and instrument/beneficiary voice morphology of Philippine-type languages. This equivalence between the two language types is functional. Systemic and morphological differences remain. Thus Madurese (or any other Indonesian-type languages such as Indonesian, Javanese and Balinese) has three actor voice forms, one unsuffixed and two suffixed respectively with *-e* and *-agi*, whereas Puyuma and Tagalog have only one. The applicatives in Indonesian-type languages take the form of suffixes. In Philippine-type languages the locative and instrument/beneficiary voice forms are constructed with a variety of affixes and, importantly, are morphologically sometimes no more complex than patient voice forms.

In Philippine-type systems in particular there has been some debate as to whether the actor voice form is intransitive, and sometimes categorical statements have been made asserting that it is (Starosta 1999). Teng shows that in Puyuma the actor voice is indeed intransitive: the undergoer can only be expressed as an oblique. Clayre indicates that the same is true of Lun Dayeh. In Tagalog, on the other hand, as Katagiri shows, actor voice clauses are not categorically intransitive. In Pendau and Moronene, both of which hover somewhere near the boundary between Philippine-type and Indonesian-type systems, Quick and the Andersens show that the actor voice is not only not intransitive, but is not limited to an indefinite undergoer argument.

This leads to a further question: if the undergoer voice is the default voice in most systems of both the Philippine and Indonesian types, then what is the function of the actor voice, and how should it and the voice system as a whole best be classified typologically? Our authors have varying views, and rightly so, as the voice systems they describe differ from each other, often in quite subtle ways. Occasionally they disagree simply because their terminological usage differs.

According to Teng, Puyuma independent clauses are syntactically ergative, as the actor of a transitive clause — always undergoer voice — is encoded differently (with the genitive) from the subject of an intransitive (with the nominative). Puyuma is the clearest case among the languages described here of such a system, and many linguists would in such circumstances call the actor voice ‘antipassive’. Teng does not choose to do so, perhaps with good reason, as the terms ‘ergative’ and ‘antipassive’ are used by linguists to denote a variety of features, and not all of them are present in Puyuma. Katagiri argues that one set of definitions involves the notion that ‘passive’ and ‘antipassive’ are morphologically derived from the form of the default voice. No Tagalog voice form is clearly derived (and the same would be true of Puyuma), so none should be labelled ‘passive’ or ‘antipassive’. The criterion that Teng uses for applying the label ‘ergative’ to Puyuma depends on the fact that only undergoer voice clauses are transitive, but this criterion does not apply to Tagalog, where Katagiri shows that some actor voice clauses are arguably transitive. Katagiri instead suggests that the Tagalog voice system tends towards the features of a split system as it is described by Dixon (1994). The undergoer voice is favoured in the realis mood and when the undergoer is highly affected. The actor voice is at least less disfavoured in the irrealis. Katagiri also draws attention to the fact that neighbouring Philippine languages behave in different ways. Cebuano has less

constraints on the use of the actor voice and may perhaps be said to be further from an ergative system than Tagalog is. Kapampangan, on the other hand, has a system of pronominal crossreferencing which is ergative in its alignment.

Teng's characterisation of the Puyuma actor voice as intransitive does not mean that an actor voice clause never has a patient argument, but rather that the actor voice is intransitive in its morphosyntax and any patient is not a core argument but is encoded with the oblique case and interpreted as indefinite. In Tagalog such a patient/undergoer is often encoded with the genitive, and Katagiri finds its core/peripheral status is ambiguous. Nonetheless, there is a strong tendency to interpret it as indefinite. Moronene takes us a step further, according to the Andersens: syntactically both the verb-ABS (undergoer voice) and the *moN*-verb (actor voice) constructions are transitive, but voice selection is based on criteria which resemble those in Tagalog. If the actor voice is chosen, the undergoer will in most cases be interpreted as indefinite. But sometimes the undergoer is definite, and the actor voice encodes the fact that the clause is low in transitivity by the criteria of Hopper and Thompson (1980), e.g. the undergoer is not well individuated or the event is irrealis or non-punctual or non-volitional.

There is no indication in Davies' account of Madurese voice that the definiteness of the undergoer affects voice selection, but Madurese voice selection seems otherwise to be determined by factors similar to those at work in Tagalog and in Moronene. We write 'seems' because Davies describes voice selection in a rather different framework from Katagiri or the Andersens. He says that Madurese follows Classical Malay, as described by Hopper (1979), in using the undergoer voice for foreground information (the story line) and the actor voice for background information (descriptive support for narrated events). It is reasonably clear that background information clauses are likely to have the low-transitivity features listed by Hopper and Thompson, i.e. that the Moronene and Madurese criteria for voice selection are similar.

Whether Moronene and Madurese apply these criteria to the same degree or in the same way is of course another matter. The differences among Philippine languages should make us wary, but it is worth noting that Davies finds roughly the same distribution of actor and undergoer voices in narrative text as Bell (1988) does in Cebuano.

Interestingly, the Andersens choose to use the label 'antipassive' for the Moronene actor voice where Katagiri rejects it in rather similar circumstances.⁴ The reason for this is simple. As discussed above, Katagiri uses the terminology of Dixon (1994) and his predecessors, whereas the Andersens use Givón's (1994) rather different definition, which is based on his measures of discourse 'topicality' (Referential Distance and Topic Persistence), i.e. the continuity of references to a particular discourse participant. In Givón's usage an antipassive voice is one in which the actor is not only more topical than the undergoer (this is also true of his 'active' voice) but the undergoer is of very low topicality. The Andersens' application of topicality measures to the actor and undergoer of Moronene actor voice clauses reveals a pattern similar to that recorded by Forrest (1994) for the Bella Coola antipassive.

The difference between Dixon's and Givón's definitions of 'antipassive' is one of morphosyntax vs discourse function, and also one of categoriality. If one applies Dixon's criteria, one arrives at a categorial morphosyntax-based definition of 'antipassive'

⁴ In fact Katagiri's morphological criterion for an antipassive is fulfilled in Moronene, as the actor voice form in *moN*- can be said to be derived from the uninflected undergoer voice form. The syntactic criterion for an antipassive is not fulfilled, however, as the actor voice form is transitive.

(although Katagiri shows that this isn't as easy as it seems). If one applies Givón's criteria, one arrives at a topicality-based definition where the measure is a matter of degree (although Givón recognises an 'active'/'antipassive' contrast only where there are two constructions in contrast).

A large part of Quick's paper is devoted to an application of Givón's topicality measures to Pendau. He finds that where the topicality of the undergoer is greater than or equal to that of the actor, the undergoer voice is chosen. Where the topicality of the actor is greater than the topicality of the undergoer, then the actor voice is selected. This suggests that, at least with regard to topicality, the determinants of voice selection in Moronene and Pendau are quite similar. Although the Andersens and Quick both apply Givón's measures, comparable figures are limited to the following, showing the percentages of actors and undergoers in actor voice clauses with a Referential Distance of 1–3 (high topicality) and > 3 (low topicality) in the two languages:⁵

	Moronene		Pendau	
	A	U	A	U
1–3	97	45	88	57
> 3	2	55	12	43

The two sets display similar tendencies but indicate that the choice of actor voice in Moronene is, if anything, more categorical than in Pendau. Its A is highly topical in 97% of cases in the Moronene corpus, 88% in the Pendau corpus.

However, undergoer topicality as measured by a Referential distance of 1–3 is perhaps not the only determinant of voice selection. Pastika (1999), for example, reports the following statistics of topicality for actor voice (AV) and undergoer voice (UV) clauses in Balinese:

	AV		UV	
	A	U	A	U
1–3	89	14	80	77
>3	11	76	20	23

As far as actor voice clauses are concerned, Balinese shows a similar tendency to Moronene and Pendau. However, it should be noted that the topicality of the actor is still higher than that of the undergoer even in undergoer voice clauses in Balinese. The significant factor for voice selection in Balinese appears to be the topicality of the undergoer rather than the topicality of the actor.

At first sight the relationship between these topicality figures and the determinants of voice selection mentioned above in regard to Tagalog, Moronene, Madurese and Balinese is perhaps not immediately obvious. However, there is such a relationship. Definiteness reflects topicality: topical items are definite or generic. Topicality entails discourse continuity, which is more likely to occur in the story-line (foreground) of a narrative than in the supporting background. And, as we observed above, the transitivity of two-argument clauses telling the story is likely to be higher than of those expressing the background. So topicality, reflected in definiteness, is associated with the higher-

⁵ Like Givón, the Andersens categorise Referential Distance into 1 (the referent is mentioned in the previous clause), 2–3 (mentioned 2–3 clauses back) and > 3 (mentioned more than 3 clauses back). We have combined their figures for 1 and 2–3 to render them comparable with Quick's.

transitivity clauses which tell the story. Pastika (1999) reports that a combination of topicality and foregrounding is a strong predictor of voice selection in Balinese.

In Puyuma and apparently Lun Dayeh, independent clauses with an actor and an undergoer are in undergoer voice unless the undergoer is indefinite, in which case (intransitive) actor voice is selected.⁶ In the other languages discussed above, this categoriality is weakened, but the undergoer voice remains the default choice for independent clauses with an actor and an undergoer. The actor voice is the marked choice, selected when (i) the undergoer is indefinite, or (ii) the actor significantly outweighs the undergoer in topicality, or (iii) the clause has low transitivity in Hooper and Thompson's terms (these three conditions are partially interdependent). These statements clearly apply to Tagalog and to Moronene (although not necessarily to the same degree). They also appear to apply to Madurese, since Davies says that the undergoer voice also reflects topic continuity of the undergoer.⁷ We do not know whether (i) and (iii) apply to Pendau, but (ii) certainly does.

The generalisations about voice selection above apply only to independent clauses. There is widespread evidence that voice selection in subordinate clauses is syntactically constrained in many Philippine- and Indonesian-type languages. Only the subject of a relative clause may be relativised, and so the speaker selects the voice which places the to-be-relativised noun phrase in the subject slot. Similarly, in complement clauses with a deleted controlee: the controlee must be the subject of the complement clause. Contributions to this volume indicate that these statements are true of Puyuma, Tagalog and Madurese. Contributions to Wouk and Ross, eds (2002) indicate that they are true of Seediq (northern Taiwan) (Holmer 2002), Bonggi (Sabah) (Boutin 2002) and Sasak (Wouk 2002).

From Kaufman's contribution it is clear that the syntactic constraint on voice selection plays a major role in conversational Tagalog. Kaufman is not concerned with syntax *per se* but with the Tagalog constructions that are used to express pragmatic focus and topicalisation. His allosentences entail a predicate nominal and a subject consisting of an embedded clause with a deleted subject corresponding to the predicate nominal. Again, the deleted noun phrase must be the subject of the embedded clause. The fourth left-hand allosentence illustrates topicalisation. Here the topic marker *ay* is preceded by a topicalised noun phrase, and yet again the corresponding noun phrase is the deleted subject of the clause following *ay*.

Kaufman's concern, incidentally, with applying Lambrecht's (1994) information structure theory to Tagalog represents an important step in understanding the mapping of grammar onto discourse structure, a step which strikes us as long overdue in the study of Austronesian languages. Associated with this, as Kaufman shows, is the study of intonational phonology and its relationship to information structure (see e.g. Ladd 1996; Beckman et al. in press), which to our knowledge has yet to be carried out in relation to an Austronesian language.

Three papers in the volume concern languages which are of neither the Philippine nor the 'Indonesian' type. We write 'Indonesian' in inverted commas because two of these, Palu'e and Manggarai, are languages of Flores, in central Indonesia. They have, however, lost the voice morphology which characterises Indonesian-type languages. Both Donohue,

⁶ The same generalisation is made by Gault (2002:371) about Sama Bangingi', a language of the Sulu archipelago in the south-western Philippines.

⁷ On Holmer's (2002) account these statements apparently apply to Seediq (northern Taiwan), too.

for Palu'e, and Arka and Kosmas, for Manggarai, describe a passive which is not marked morphologically. Instead, in Palu'e it is marked by undergoer–actor–verb constituent order, whilst in Manggarai the passive construction as a whole is usually identifiable through the presence of an oblique actor marked with the preposition *l(e)*.

These papers bring us back yet again to the question: How should Austronesian voice systems be characterised in terms of syntactic typology? Whereas in Philippine-type and Indonesian-type languages it is the actor voice about which this question is most often asked (although Quick does ask if the Pendau undergoer voice is a passive, and shows that it isn't), in Palu'e and Manggarai, the question is asked about a putative undergoer voice construction: Is it a passive? Now when analysts ask whether the actor voice in a Philippine- or Indonesian-type language is an antipassive (or when Quick asks whether the Pendau undergoer voice is a passive), the question is not, 'Is this a voice?' but only, 'What voice is this?' In the cases of Palu'e and Manggarai, however, the more fundamental question is asked: If this is a voice, it is passive, but is it in fact a voice? The authors of both papers adduce strong syntactic arguments to show that, despite the lack of inflection, the constructions they describe are both passive voices, and that there are typological precedents for them.

The third paper in this trio is Otsuka's on Tongan. She describes three constructions in Tongan which could be or have been claimed to be passive. The *-Cia* construction she dismisses as a fossil, focusing on the agentless transitive and VOS constructions. She shows that, despite appearances, they are not passives.

References

- Arka, I Wayan, 2003, *Balinese morphosyntax: a lexical-functional approach*. Canberra: Pacific Linguistics.
- Baird, Louise, 2002, A grammar of Keo: an Austronesian language of East Nusantara. PhD thesis, The Australian National University.
- Beckman, Mary E., Julia Hirschberg and Stefanie Shattuck-Hufnagel, in press, The original ToBI system and the evolution of the ToBI framework. In Sun-an Jun, ed. *Prosodic models and transcription: towards prosodic typology*. New York: Oxford University Press.
- Bell, Sarah J., 1988, Voice and foreground/background in Cebuano. In Richard McGinn, ed. *Studies in Austronesian linguistics*, 427–440. Athens, Ohio: Center for Southeast Asian Studies, Ohio University Center for International Studies.
- Boutin, Michael E., 2002, Nominative and genitive case alternations in Bonggi. In Wouk and Ross, eds 2003:209–239.
- Bowden, John, 2001, *Taba: description of a South Halmahera language*. Canberra: Pacific Linguistics.
- Corston, Simon H., 1996, *Ergativity in Roviana*. Canberra: Pacific Linguistics.
- Crowley, Terry, 1998, *An Erromangan (Sye) grammar*. Honolulu: University of Hawai'i Press. Oceanic Linguistics Special Publication 27.
- Davis, Karen, 2003, *A grammar of the Hoava language, Western Solomons*. Canberra: Pacific Linguistics.
- Dixon, R.M.W., 1988, *A grammar of Boumaa Fijian*. Chicago: University of Chicago Press.

- 1994, *Ergativity*. Cambridge: Cambridge University Press.
- Emmorey, Karen, 1984, The intonation system of Toba Batak. In Schachter, ed. 1984:37–58.
- Ezard, Bryan, 1997, *A grammar of Tawala: an Austronesian language of the Milne Bay area, Papua New Guinea*. Canberra: Pacific Linguistics.
- Forrest, Linda B., 1994, The de-transitive clauses in Bella Coola: Passive versus inverse. In Givón, ed. 1994:147–168.
- Gault, JoAnn, 2002, Some aspects of focus in Sama Bangingi. In Wouk and Ross, eds 2002:367–378.
- Givón, Talmy, ed., 1994, *Voice and inversion*. Amsterdam: John Benjamins.
- Himmelman, Nikolaus P., 2002, Voice in two northern Sulawesi languages. In Wouk and Ross, eds 2002:7–16.
- 2005, The Austronesian languages of Asia and Madagascar: typological characteristics. In N.P. Himmelman and Alexander Adelaar, eds *The Austronesian languages of Asia and Madagascar*. London: Routledge Curzon.
- Holmer, Arthur, 2002, The morphology and syntax of Seediq focus. In Wouk and Ross, eds 2002:333–354.
- Hopper, Paul J., 1979, Aspect and foregrounding in discourse. In Talmy Givón, ed. *Discourse and syntax*, 213–241. New York: Academic Press.
- Hopper, Paul J. and Sandra Thompson, 1980, Transitivity in grammar and discourse. *Language* 56:251–299.
- Huang, Shuanfang, 2002, The pragmatics of focus in Tsou and Seediq. *Language and Linguistics* 3:665–694.
- Hyslop, Catriona, 2001, *The Lolovoli dialect of the North-East Ambae language, Vanuatu*. Canberra: Pacific Linguistics.
- Keesing, Roger M., 1985, *Kwaio grammar*. Canberra: Pacific Linguistics.
- Ladd, D. Robert, 1996, *Intonational phonology*. Cambridge: Cambridge University Press.
- Lambrecht, Knud, 1994, *Information structure and sentence form: topic, focus and the mental representations of discourse participants*. Cambridge: Cambridge University Press.
- Lichtenberk, Frantisek, 1983, *A grammar of Manam*. Honolulu: University of Hawai'i Press. Oceanic Linguistics special publication No. 18.
- Lynch, John, 2003, *Issues in Austronesian historical phonology*. Canberra: Pacific Linguistics.
- Lynch, John, Malcolm Ross and Terry Crowley, 2002, *The Oceanic languages*. Richmond: Curzon Press.
- Mead, David, 2002, Proto Celebic focus revisited. In Wouk and Ross, eds 2002:143–177.
- Mosel, Ulrike, 1984, *Tolai syntax and its historical development*. Canberra: Pacific Linguistics.
- Mosel, Ulrike and Even Hovdhaugen, 1992, *Samoan reference grammar*. Oslo: Scandinavian University Press.
- Norwood, Clodagh, 2002, Voice and valency alternations in Karo Batak. In Wouk and Ross eds, 2002:181–207. Canberra: Pacific Linguistics.

- Pastika, I Wayan, 1999, Voice selection in Balinese narrative discourse. PhD thesis, The Australian National University
- Pawley, Andrew K., 1973, Some problems in Proto Oceanic grammar. *Oceanic Linguistics* 12:103–188.
- Poedjosoedarmo, Gloria, 2002, Changes in word order and noun phrase marking from Old to modern Javanese: implications for understanding developments in western Austronesian ‘focus’ systems. In Wouk and Ross, eds 2002:311–330.
- Quick, Phil, 2002, A sketch of the primary transitive verbs in Pendau. In Wouk and Ross, eds 2002:101–122.
- Ross, Malcolm, 2002a, The history and transitivity of western Austronesian voice and voice-marking. In Wouk and Ross, eds 2002:17–62.
- 2002b, Final words: research themes in the history and typology of western Austronesian languages. In Wouk and Ross, eds 2002:451–474.
- 2004, The morphosyntactic typology of Oceanic languages. *Language and Linguistics* 5:491–541.
- Schachter, Paul, 1976, The subject in Philippine languages: topic, actor, actor-topic or none of the above. In Charles Li, ed. *Subject and topic*, 491–518. New York: Academic Press.
- Schachter, Paul, ed., 1984, *Studies in the structure of Toba Batak*. UCLA Occasional Papers in Linguistics 5.
- Starosta, Stanley, 1986, Focus as recentralisation. In Paul Geraghty, Lois Carrington and S.A. Wurm, eds *FOCAL I: Papers from the Fourth International Conference on Austronesian Linguistics*, 73–95. Canberra: The Australian National University.
- Starosta, Stanley, 1999, Transitivity, ergativity and the best analysis of Atayal case marking. In Elizabeth Zeitoun and Paul Jen-kuei Li, eds *Selected papers from the Eighth International Conference on Austronesian Linguistics*, 371–392. Taipei: Institute of Linguistics (Preparatory Office), Academia Sinica.
- Van Klinken, Catherina, 1999, *A grammar of the Fehan dialect of Tetun, an Austronesian language of West Timor*. Canberra: Pacific Linguistics.
- Verhaar, John M.W., 1984, Affixation in contemporary Indonesian. In Bambang Kaswanti Purwo, ed. *Towards a description of contemporary Indonesian: preliminary studies*, Part I, 1–26. Jakarta: NUSA.
- Wolff, John U., 1996, The development of the passive verb with pronominal prefix in western Austronesian languages. In Bernd Nothofer, ed., *Reconstruction, classification, description: festschrift in honor of Isidore Dyen*, 15–40. Hamburg: Abera.
- Wouk, Fay, 2002, Voice in the languages of Nusa Tenggara Barat. In Wouk and Ross, eds 2002:285–309.
- Wouk, Fay and Malcolm Ross, eds, 2002, *The history and typology of western Austronesian voice systems*. Canberra: Pacific Linguistics.

2 *Kelabitic languages and the fate of ‘focus’: evidence from the Kerayan*

BEATRICE CLAYRE

1 Introduction

Kelabitic dialects or languages¹ are a subgroup of the North Sarawak group of languages (Blust 1997). They are spoken across the borders of Sabah, Sarawak and Kalimantan Timur in the interior of Borneo (see Map 1). Members of this language group include Lun Dayeh, Kelabit, and Sa’ban.² Linguistically, they occupy an area of transition between the languages of Sabah to the north which have a developed focus (or voice) system, and the languages of Sarawak and Kalimantan to the south which have a much reduced, or even lost, focus (or voice) system.³

The term ‘voice’ will be used here in preference to ‘focus’. The aim of this paper is to outline the voice system as it occurs in Lun Dayeh and Sa’ban and to compare this with the results of a brief survey of Kelabitic languages in Kalimantan Timur undertaken in 1995.⁴

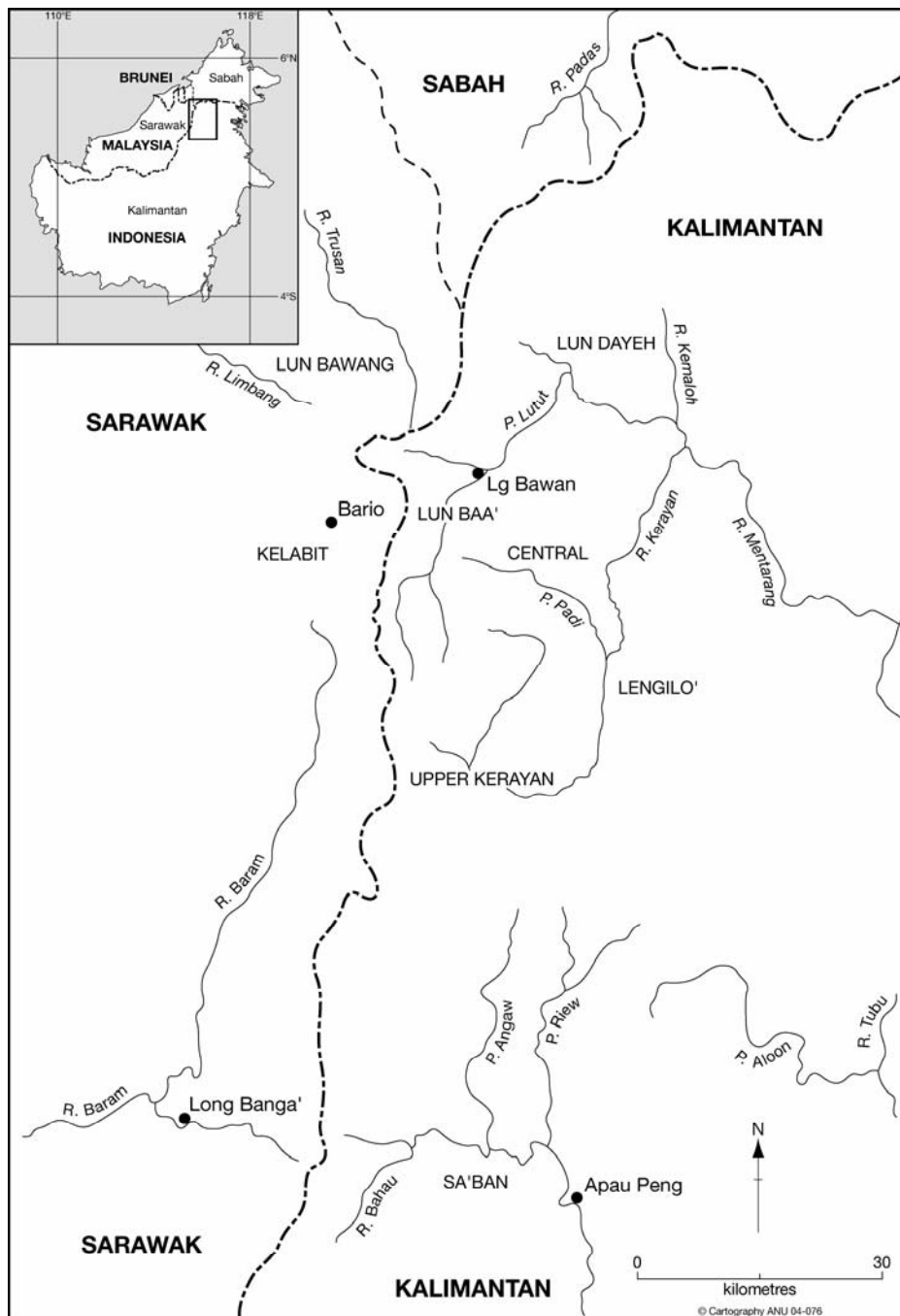
¹ This language group was formerly called the ‘Apo Duat’ group (Hudson 1978). The terms dialect and language are used loosely here. No dialect testing has been undertaken, but there appears to be considerable mutual intelligibility between many of the dialects. On the other hand, there are examples of lack of mutual intelligibility, for example between Lun Dayeh and Sa’ban despite many similarities between some dialects, for example Kelabit and Lun Dayeh, some speakers maintain that they are separate languages.

² In placenames and language names, an apostrophe indicates a glottal stop. Lg and K. are abbreviations for *long* and *kuala* meaning ‘river mouth’, P. and B. are abbreviations for *paa*’ or *baa*’ ‘river’, and T. is an abbreviation for *tanjung* ‘headland’.

³ The term ‘focus’ is used here in the sense of Philippine-type focus.

⁴ I am grateful to the Culture and Conservation research programme of the Kayan Mentarang National Park project and the Ford Foundation (Jakarta) who made this survey possible, the staff of the World Wide Fund for Nature in Samarinda who took care of my travel arrangements in Kalimantan, and Samuel S.T. Padan who accompanied me on the survey. I am indebted to Dr Bernard Sellato who first interested me in this survey, and who so generously shared his knowledge of the area and its languages with me.

Issues particularly addressed here are: voice and aspect affixes with verbs, the marking of nouns, the number and function of pronoun sets, and word order.⁵



Map 1: Kelabitic languages in Sabah, Serawak and Kalimantan Timur

⁵ These were some of the issues which, in 1995, Dr René van den Berg and I agreed to look at in our respective areas of interest, Sulawesi and central northern Borneo, with a view to preparing a preliminary typological picture of ‘focus’ in these regions (Clayre 1995; Wouk 2002:1).

2 The voice system in Lun Dayeh and Sa'ban

2.1 Lun Dayeh

Historically, the Lun Dayeh ('people of the interior') inhabited the mountainous areas of north-west Kalimantan Timur. There they are also known as Lun Tana' Luun (highland people who cultivate dry rice), Putok or Kemaloh Murut. Around the eighteenth century some Lun Dayeh groups spread down the Trusan, Limbang and Lawas rivers into north Sarawak where they were, until recently, known as Murut.⁶ They now call themselves Lun Bawang ('people of the country'). Over the last hundred years, other Lun Dayeh have migrated from Kalimantan, via the Padas river to Sabah, where they have recently legally registered their name as Lundayeh.⁷

The dialect of Lun Dayeh described here is that of the Kemaloh area of the Kerayan in Kalimantan.⁸

2.1.1 Lun Dayeh verbal affixation

2.1.1.1 Voice and aspect

Lun Dayeh has three voices: actor, undergoer and instrument, but there are indications that at an earlier stage in the language there was at least one other voice which may have been locative or benefactive.⁹ Table 1 shows the affixes that signal the voice of a Lun Dayeh verb. The only other affixes that may occur with a voice affix are those marking aspect. In Lun Dayeh imperfective aspect is unmarked, only perfective aspect is marked, and in perfective aspect undergoer voice is unmarked. The voice and aspect affixes of Lun Dayeh transitive clauses are given in Table 1, together with the affixes that signal imperative mood.

⁶ In the literature they were also called Sarawak Murut to distinguish them from the ethnically and linguistically distinct Sabah Murut (also known as Timugon, Tagal etc.).

⁷ For further information on Lun Dayeh and related groups see Bolang and Harrisson (1949), Cense and Uhlenbeck (1958:21–23), Crain (1978), Langub (1987), Lees (1959), Schneeberger, (1979), Southwell (1949), Tuie (1995).

⁸ I am most grateful to Semion Lalung of Sipitang, Sabah, for working with me on the preliminary analysis of his language in 1972 when he was a student in the UK. His parents came from Lg Nuat in the Kemaloh region of Kalimantan. I am also grateful to Daniel Teg^hang and Hendrik Gug^hang Tebari, originally from Lg Nuat, now living in Kampong Baru, Kerayan region, Kalimantan Timur, for language help in 1995.

⁹ The evidence for this voice is firstly the occurrence of two imperative suffixes *-a* and *-i* (see §2.1.1.2), and secondly one example in my data of the affix combination *piN-√an*, as in the following perfective aspect sentence with the verb root *beré* 'give':

Idé nepimeréan mu lawid di.
 who PF.?V:give GEN.2S fish AP
 'To whom did you give the fish?'

It was explained to me that this construction would be used if, for example, you were checking that a child had delivered the fish to the right person.

Table 1: Lun Dayeh voice and aspect affixes

	Imperfective	Perfective	Imperative
Actor	<i>N-</i>	<i>ne-N-</i>	<i>N-</i>
Undergoer	<i>-en</i>	<i>-in-/-i-</i>	<i>-u</i>
Instrument	<i>piN-</i>	<i>ne-piN-</i>	Ø
?Locative/benefactive*	Ø	Ø	<i>-aʔ, -i</i>

* See footnote 9.

The prefix *N-* symbolises homorganic nasal substitution and it signals actor voice. Before roots with initial *l* or *r*, it occurs as *nge-* (for example *ngelamud* ‘to mix’, and *ngeramit* ‘to scratch’) and before vowel-initial roots as *ng-* (for example *ngalap* ‘to take’). Examples (1–2) below illustrate verbs marked for actor voice with the verb roots *kekeb* ‘cover’ and *put* ‘blowpipe’.¹⁰ This prefix occurs with verbs in imperfective aspect in contrast to verbs in perfective aspect which add *ne-* before the *N-* prefix.

- (1) *Ngekeb lacang nih uih atun ...*
 AV:cover pot this PIV.1S first
 ‘I am covering the pot first/before ...’
- (2) *Ngukab serei? di uih em nebila? nah ieh.*
 AV:open bottle AP PIV.1S and ST:shatter PT 3S
 ‘I was opening the bottle and it shattered.’

Perfective aspect in actor voice is signalled by adding the prefix *ne-* before the *N-* prefix as illustrated in (3) with the root *beruʔ* ‘wash something’.

- (3) *Idé nemeru? lacang inih.*
 PIV.who PF.AV:wash cooking pot this
 ‘Who washed this cooking pot?’

Undergoer voice is marked by the suffix *-en* added to the root, as illustrated in example (4), with the root *beli* ‘buy’. This suffix occurs only with verbs in imperfective aspect.

- (4) *Belien kuh lal ineh ku usin inih.*
 buy:UV 1S.GEN PIV.hen that with money this
 ‘I’ll buy that hen with this money.’

The infix *-in-* signals perfective aspect, and in Lun Dayeh this infix occurs only with verbs in undergoer voice, which are otherwise unmarked. This affix is infixed following the initial consonant of the root (for example *karem* ‘capsize’, *kinarem* ‘capsized’). It occurs as *in-* before a vowel-initial root (for example *ukab* ‘open’, *inukab* ‘opened’); however, roots with a schwa as the penultimate vowel show a pattern of verbal ablaut in

¹⁰ Abbreviations used in interlinear glosses: 1, 2, 3 first, second, third person; 1EP first person plural exclusive; 1IP first person plural inclusive; AP anaphoric particle; AV actor voice; D dual pronoun; FP friendly particle; GEN genitive; IMP imperative; IMPF imperfective; IV instrument voice; LP locative particle; NEG negative; NPNA non-pivot non-actor; P plural; PIV pivot; PF perfective; PM personal noun marker; PT particle; QP question particle; REL relative pronoun; S singular; ST stative; UV undergoer voice; () brackets indicate voice markings which are not present; { } brackets indicate voice markings which are cancelled; the symbols /c/ and /j/ represent voiceless and voiced affricates, /ng/ the velar nasal, /e/ the schwa, /é/ the front open-mid vowel, and /:/ lengthened vowel.

which the schwa is replaced by *-i-* (for example *beré* 'give', *biré* 'given').¹¹ Examples (5–6) illustrate verbs in perfective aspect (and undergoer voice) with the infixes *-in-* (5) and *-i-* (6). The verb root in (5) is *pipi?* 'form or shape' and in (6) *kekeb* 'cover'.

- (5) *Kareb tana? pangeh p-in-ipi? neh ...*
 when land finish <UV.PF>-form GEN:3S
 'When she had created the land ...'
- (6) *K-i-keb kuh lacang di peh.*
 <UV.PF>-cover 1S.GEN pot AP PT
 'I covered the pot already.'

Examples (7–9) illustrate verbs in instrument voice and imperfective and perfective aspect. The verb root in (8) is *taban* 'to take away', in (9) it is *até* 'death'. Examples (4) and (7) highlight the difference between undergoer and instrument voice.

- (7) *Pimeli kuh lal usin inih.*
 IV:buy 1S.GEN hen PIV.money this
 'I will use this money to buy the hen.'
- (8) *Bekang inih pinaban kuh barang inih.*
 PIV.basket this IV:take away GEN.1S things this/these
 'I will use this basket to take these things away.'
- (9) *Anun nepingaté neh neneh.*
 PIV.what PF.IV:die 3S.GEN NPNA:3S
 'What did he use to kill him?'

2.1.1.2 Imperatives

Examples (10–12) illustrate the use of the imperative suffixes. These suffixes are attached to an unaffixed verb root and are not in common use, probably because it is regarded as more polite to use a verb in indicative mood as illustrated in examples (13–14). All three imperative suffixes may occur with the same root, for example *libud-u*, *libud-a?* and *libud-i* 'roll!'. The suffix *-u* (10) signals undergoer voice, while the *-a?* suffix (11) seems to be used when the object is at a distance or out of sight¹² and was translated by informants as 'go and do x'. The function of the *-i* suffix (12) is uncertain. The verb root in example (10) is *kekeb* 'cover', and in (11–12) it is *beru?* 'wash something'.¹³

- (10) *Kekebu lacang ineh mikat.*
 cover:UV.IMP PIV.pot that quickly!
 'Cover the pot quickly!'
- (11) *Berua? sawan ineh kareb m-ecang.*
 wash:?V.IMP cup that while ST-light
 'Go and wash that/those cup/s while it is still (day)light!'

¹¹ See Blust (1997) for a discussion of ablaut in western Borneo.

¹² This is the way it was explained to me by a Lun Dayeh speaker.

¹³ The final glottal stop of the root disappears when a suffix is added.

- (12) *Berui bigan ineh.*
wash:V.IMP plate that
'Wash that plate!'

More polite imperatives are obtained by using indicative mood in actor or undergoer voice, as illustrated in examples (13–14). The verb root in both examples is *beré* 'give'.

- (13) *Meré buku ineh ni=Dawat.*
AV:give book that PM.NPNA=Dawat
'Give that book to Dawat!'
- (14) *Beré:n mu kuyuh inih ni=Balang!*
give:UV 2S.GEN PIV.shirt this PM.NPNA=Balang
'Give this shirt to Balang!'

2.1.1.3 Intransitive verbs in Lun Dayeh

Intransitive verbs occur either as simple, unaffixed roots or with the infix *-em-* (*m-* before vowel-initial roots), which is the reflex of PAN **-um-* in Lun Dayeh.¹⁴ Perfective aspect is signalled by adding the prefix *ne-*. Examples (15–16) illustrate two intransitive verbs with the *-em-* infix, *tulud* 'fly' and *languy* 'swim', the former in imperfective aspect (unmarked), and the latter in perfective aspect. Example (17) illustrates an intransitive verb marked for perfective aspect (*ne-*) and actor voice (*m-*), and example (18) illustrates two unaffixed roots (*tudo* 'stay/sit', *rudap* 'sleep') with the perfective prefix *ne-*.

- (15) *Ieh temulud rat pe-tinueh guta dalam amé pe-kabing,*
PIV.3S <AV>:fly from LP-right across path go LP-left
inih bian luk doo? tutu.
this omen REL good very
'It is a very good omen, (when) it flies from right to left across the path.'
- (16) *Idi uko nelemanguy nengubit puet neh.*
then PIV.dog PF.<AV>:swim PF.AV:bite buttock GEN:3S
'Then the dog swam (to him) and bit his buttocks.'
- (17) *Idé nemenad bua? kuh di.*
PIV.who PF.AV:climb fruit 1S.GEN AP
'Who climbed up my fruit (tree)?'
- (18) *Idi bang se-decem kai netudo nerudap*
then in one-night PIV.1EP PF.(AV):stay PF.(AV):sleep
bang pulung.
in forest
'Then for one night we stopped and slept in the forest.'

¹⁴ One verb in Lun Dayeh retains the *-um-* form of the infix: *kuman* 'to eat'. It occurs with other voice and aspect affixes as follows: *nekuman* (AV.PF), *kenen* (UV.IMPF), *kinan* (UV.PF). Notice that the *-um-* and *-in-* infixes do not co-occur (as they do in some Dusunic languages), but *-in-* replaces the *-um-* infix. The stative forms of this verb (see §2.1.1.4) are *mekan* and *nekan*.

In Lun Dayeh *-em-* occurs only with intransitive verbs.¹⁵ There is one known example of its occurrence with a meteorological verb: *m-udan* ‘raining’ (Himmelman 2002:9, fn.5). In Tagalog and Kimaragang Dusun the infix *-um-* occurs in change of state predicates (Kroeger 1990:§1.2.6), but this usage does not occur in Lun Dayeh which signals change of state with *amé* ‘go’ as illustrated in (19).

- (19) *Don ineh amé me-sia?*
leaf that go ST-red
‘The leaf becomes red.’

2.1.1.4 Other verb affixes in Lun Dayeh

The stative verb affixes in Lun Dayeh are *me-* ‘present state’ and *ne-* ‘completed state’ (*m-* and *n-* before vowel-initial roots), and the two prefixes can also respectively signal the potential for an action to be performed, or involuntary action. The stative prefix is added to the verb root, but does not give rise to homorganic nasal substitution, which occurs only with dynamic actor voice verbs. Notice too that the stative perfective prefix *ne-* replaces the imperfective *me-* prefix; it is not added to it as occurs with the dynamic actor voice perfective prefix *ne-* (3, 16–17). Stative verbs with a *me-* prefix are illustrated in examples (20, 22), and with *ne-* in (21).¹⁶

- (20) *Me-beré mu kuyuh inih ni=Balang.*
ST-give GEN.2S shirt this PM.NPNA=Balang
‘Are you able to give this shirt to Balang?’ OR: ‘Can you give ...’
- (21) *Ne-kekeb kuh lapung nah.*
ST-cover GEN.1S lamp PT
‘I accidentally covered the lamp.’
- (22) a. *Na meraut ku apuy, meseb iko napeh.*
NEG. AV:play with fire ST:burn 2S later
‘Don’t play with fire, you are liable to get burnt.’
(OR: ‘... you may burn yourself.’)
- b. *Na meraut ku apuy mekeseb mu ruma? napeh.*
NEG. AV:play with fire ST:ke:burn GEN.2S house later
‘Don’t play with fire, you are liable to burn the house down.’

In addition, the stative prefixes can combine with other affixes, such as *-ke-* and *-te-*, whose function is not yet fully understood (22b).¹⁷ Lun Dayeh stative constructions are always in undergoer voice. Stative affixation is not discussed further here, but it may be noted that there are some similarities between the Lun Dayeh stative system and that described by Kroeger (1990) for Kimaragang Dusun in Sabah.

¹⁵ With the possible exception of *k-um-an* ‘eat’. This verb, along with *beré* ‘give’ and *(t)aru?* ‘do’ or ‘make’, is remarkable in all Kelabitic languages for retaining affixes which appear to have been lost in other verbs (see §2.2.1.3, and §3.1.2.1).

¹⁶ Compare the stative form of these verbs with the actor voice forms *ngekeb* and *meré* in (1) and (13). The actor voice form of *meseb* (22) is *ngeseb* ‘burn something’

¹⁷ The affix *-ke-* (*-k-* before vowel initial roots) may signal the presence of a non-pivot actor.

Other verb affixes in Lun Dayeh are: *pe-* ‘reciprocal’, *peri-* ‘multiple’ (action or actors), and *si-* pretence action, all occur in actor voice type of constructions, and all signal perfective aspect by adding the prefix *ne-*. These affixes are not described further here.

2.1.2 *Lun Dayeh pronoun sets*

Lun Dayeh has three sets of pronouns distinguished by their function in a clause. Pronouns of set I are free pronouns, and mark the pivot of the clause whether it is the actor, the undergoer, or the instrument. Set II pronouns always occur immediately following the verb or noun they qualify. They mark the non-pivot actor when they follow a verb, and the possessor when they follow a noun. Set III pronouns indicate non-pivot and non-actor arguments, such as the undergoer in actor voice, or an oblique in any voice.

The first and second person plural pronouns of sets I and II are identical in form, but different in function. Their function is made clear by word order in the clause. Only the occurrence of the proclitic *ne=* distinguishes the pronouns of set III from those of set II but the two sets function in different ways. The proclitic *ne=/n=*¹⁸ seems to function in a similar way to the personal noun marker *ni=*.¹⁹

Table 2: Lun Dayeh pronoun sets

	Pivot	Non-pivot	
	I (pivot)	II (actor/genitive)	III (non-actor)
1S	<i>uuh</i>	<i>kuh</i>	<i>nekuh</i>
2S	<i>iko</i>	<i>mu</i>	<i>nemu</i>
3S	<i>ieh</i>	<i>neh</i>	<i>neneh</i>
1IP	<i>tau</i>	<i>tau</i>	<i>netau</i>
1EP	<i>kai</i>	<i>kai</i>	<i>nekai</i>
2P	<i>muyuh</i>	<i>muyuh</i>	<i>nemuyuh</i>
3P	<i>ideh</i>	<i>deh</i>	<i>ndeh</i>

The use of set II pronouns as genitives (GEN) is illustrated in (23 and 25, 29), and as actors in undergoer voice clauses in (26–28). Examples (24, 25, 27, 28) illustrate the use of pronouns of set I as the pivot (PIV) of the clause. In (24–25) the pivot is the actor in an actor voice construction and in (27, 28) the pivot is the undergoer in an undergoer voice construction. Pronouns of set III (NPNA) are illustrated in (24–26, 28, 29); in (24–25, 29) in the role of undergoer in actor voice clauses, and in (24, 26, 28) as obliques in both actor and undergoer voice clauses. In (24) the oblique is preceded by the verb, *amé* ‘go’, which is commonly used to indicate direction towards, and in (28) by the preposition *rat* ‘from’.

- (23) *rumah kuh*
house GEN:1S
‘my house’

¹⁸ In some dialects of Lun Bawang in Sarawak, the proclitic occurs as *ke=/k=*.

¹⁹ See §2.1.3. *Ne=* cannot be interpreted as a preposition, as it does not occur elsewhere in the language, and pronouns with this proclitic can co-occur with a preposition (see (28)).

- (24) *Iko nguit neneh amé nekuh.*
 PIV.2S AV:bring NPNA:3S go NPNA:1S
 'You bring it/him to me.'
- (25) *Kareb uih isuut idi tinan kuh meb^hang nekuh*
 while PIV.1S small then mother GEN.1S AV:forbid NPNA:1S
melalid kudeng uih na maya? bala neh
 ST:naughty if PIV.1S not AV:follow word GEN.3S
idi ieh ngeteb²⁰ nekuh.
 then PIV.3S AV:strike NPNA:1S
 'When I was small, my mother forbade me to be naughty, if I did not obey her then she would strike me.'
- (26) *Beked ini beré:n kuh nemu.*
 PIV.shirt this give:UV GEN.1S NPNA:2S
 'I'll give this shirt to you.'
- (27) *Uih initun neh.*
 PIV.1S (UV).PF:ask GEN.3S
 'He asked me.' ('I was asked by him.')
- (28) *Inapung kuh ieh rat neneh.*
 (UV).PF:hide GEN.3S PIV.3S from NPNA:3S
 'I hid it from him.'
- (29) *Na doo? tu irin kai ngaceku na lemluen*
 NEG. good very vegetable GEN.1EP because NEG. person
ngasa? neneh
 AV:look after NPNA:3S
 'Our vegetable(s) is/are not very good because no one looks after it.'

2.1.3 Lun Dayeh noun phrase markers

In Philippine type 'focus' languages, noun phrase markers indicate the semantic role of an argument. It will be apparent already from the examples given above that there are no noun phrase markers for common nouns in Lun Dayeh. In a clause with two core arguments the role of each is indicated by its constituent order (see §2.1.4).

There are two personal noun phrase markers in Lun Dayeh: *i=* and *ni=*,²¹ whose function seems to have been similar to that of pronouns of sets I and III, but which today are used less consistently, if at all. The marker *i=* occurs with personal nouns that are syntactic pivots, but may also be found before a personal noun that is an undergoer. On the other hand, the marker *ni=* never occurs with pivots, but does occur before personal

²⁰ *Ngeteb* normally means 'to cut' or 'to sting'.

²¹ In some dialects of Lun Dayeh the noun markers are *i=* and *ki=*.

nouns and some kinship terms²² that are either undergoers or non-pivots and non-actors.²³ Examples (13–14, 20, 31–32) illustrate the use of *ni=* before personal nouns or kindred terms which are not pivots, and example (30) the use of *i=* before a personal noun which is the pivot of the clause.

- (30) *I=Agong ngenecuk nekuh ngarem namu, leh!*
 PM.PIV=Agong PF.AV:order NPNA:1S AV:capsize NPNA:2S FP
 ‘Man! Agong told me to capsize you!’
- (31) *Uih nenecat ni=Pengiran.*
 PIV.1S PF.AV:hit PM.NPNA=Pengiran
 ‘I hit Pengiran.’
- (32) *Kareb uih anak megai? uih ngitun*
 While 1S child often PIV.1S AV:ask
ni=apu? ratnan ...
 PM.NPNA=grandfather about
 ‘When I was a child I often asked grandfather about ...’

The interrogative pronoun *idé* ‘who’ occurs with the proclitic *n=* when it is not the pivot of the clause. In example (33–34) *idé* is the pivot in an actor voice and an undergoer voice clause respectively, but in (35) *nidé* is the non-pivot-non-actor pronoun in an actor voice clause.

- (33) *Idé nenawar nekuh ina dei?*
 PIV.who PF.AV:call NPNA.1S just now
 ‘Who called me just now?’
- (34) *Idé daren mu.*
 PIV.who call:UV GEN.2S
 ‘Who are you calling?’
- (35) *Nawar nidé iko nah?*
 AV:call NPNA:who PIV.2S PT
 ‘Who are you calling?’

2.1.4 Word order in Lun Dayeh

In Lun Dayeh only the pronoun sets, and less consistently the personal noun markers, are marked to indicate the semantic role of an argument in a clause. Common nouns are never marked by any kind of phrase marker. This lack of morphological information means that constituent order is important for the interpretation of the clause. In any clause with one or two core arguments, the prescribed position of the non-pivot argument is immediately following the verb. This applies to the undergoer in actor voice constructions

²² Boutin (2002:215) describes similar usage in Bonggi. In Lun Dayeh the noun phrase marker also occurs in folk stories with animal names, e.g. *i=/ni=pelanok* ‘Mr Mousedeer’, and examples of *ni=* before the word *lun* ‘person(s)’ have been noted.

²³ The set III pronoun proclitic *ne=* and the noun phrase marker *ni=* may be allomorphs of a marker for oblique case or non-pivot arguments.

(marked by set III pronouns, for example (24)) and to the actor in undergoer (and instrument) voice constructions (marked by set II pronouns, for example (7, 26–28)). The pivot (marked by set I pronouns) whose semantic role, as actor, undergoer, or instrument, is signalled by the voice affix on the verb, is more free. It may occur before or after the verb-plus-non-pivot construction: compare examples (1, 24) in actor voice and (27–28) in undergoer voice. Instruments in instrument voice constructions occur either at the beginning or end of the clause (7–9). Obliques follow the verb and its core arguments (4, 24, 26).

2.2 Sa'ban

Originally, the Sa'ban lived along the upper Bahau river and its tributaries in Kalimantan Timur, where they were also known as Berau or Merau (Schneeberger 1979).²⁴ The Bahau is a tributary of the Kayan river, and the area settled by the Sa'ban is located south of the mountains which separate the Bahau-Kayan river system from the Kerayan-Mentarang-Sesayap river system. The area was also inhabited by other ethnic groups, such as the Murik and the Kenyah.²⁵ There were once several Sa'ban villages, each with its own slightly different dialect, but the region is now deserted, and the Sa'ban have dispersed, some to the upper Baram river in Sarawak,²⁶ others to the upper Kerayan, and others again downriver to Apau Peng, Tanjong Selor, and other towns in Kalimantan Timur.

The dialect of Sa'ban described here is that spoken today in Lg Banga' in the upper Baram area of Sarawak.²⁷ Although Sa'ban is a member of the Kelabitic language group, it looks and sounds very different from Lun Dayeh because of the many sound changes which have taken place in the language (Blust 1999; Clayre 1972; Clayre in preparation).²⁸

²⁴ The Merau Kalun mentioned by Schneeberger (1979:31, 33) can be identified as Sa'ban from the Kalun river (known to the Sa'ban as Pei' Aloon) whose dialect is somewhat different to that of Sa'ban from the Bahau area.

²⁵ The Murik (also known as Ngurik) moved away from the Bahau area in the nineteenth century, some migrating to Sarawak. There were several Kenyah groups in the area, one of which, the Lepo' Ké Kenyah, migrated with the Sa'ban to Sarawak after the war, and settled at Lg Banga'.

²⁶ Sa'ban first started moving to the Baram area about the beginning of the twentieth century, but the greatest migration took place following World War II (Bolang & Harisson 1949:117).

²⁷ I would like to record my thanks to all members of the Sa'ban communities of Lg Banga' and Lg Puak who over the years have helped me learn their language. In 1967–68 I had lived in Lg Banga' for eighteen months, and I am grateful to the Committee for South-east Asian studies of the British Academy for grants which allowed me to return to Lg Banga' in 1999 and 2000 for two periods of research, each of three months duration. I also wish to express my gratitude to the Sarawak State Planning Unit for permission to carry out research on the Sa'ban language, and to the Sarawak Museum and the Majilis Adat Istiadat in Kuching for sponsoring my research.

²⁸ In Sa'ban orthography /h/ in word-final position represents the voiceless counterpart preceding a voiced sonorant; /a/ represents an open-mid fronted back unrounded vowel [ʌ], contrasting with the close-mid central unrounded vowel [ə] represented in the orthography by /e/; /ú/ represents a near-close central lightly rounded vowel [u̯]; all vowels may be lengthened and are written as two like vowels, /aa/ etc.; all consonants except glottal stop, /h/, /w/ and /y/ may occur as geminates in word-initial position (Clayre, in preparation).

2.2.1 Sa'ban verb affixation

2.2.1.1 Voice and aspect

Compared to Lun Dayeh the Sa'ban voice system is much reduced. There are only two voices in Sa'ban transitive clauses: actor and undergoer. In addition, actor voice is marked morphologically only in imperfective aspect and undergoer voice only in perfective aspect. Perfective aspect in actor voice is signalled by an auxiliary verb, and imperfective aspect in undergoer voice by a periphrastic construction. Sa'ban voice and aspect affixes are set out in Table 3.

Table 3: Sa'ban voice and aspect affixes

	Imperfective	Perfective
Actor	<i>N-</i>	∅
Undergoer	∅	<i>i-</i>

The actor voice affix is, as in Lun Dayeh, a nasal prefix which assimilates to the position of the initial consonant of the root. In Sa'ban, however, there has been extensive erosion of word-initial consonants, vowels and even syllables of Sa'ban words and roots. This means that when the *N-* prefix is added to an eroded root it assimilates to the position of what in Proto Kelabit-Lun Dayeh (PKLD) was a medial consonant, and produces word-initial consonant clusters. Table 4 compares a selection of verb roots from PKLD, Lun Dayeh and Sa'ban to illustrate the effect of the erosion on Sa'ban verbs.

Table 4: A comparison of verbs marked for actor voice in PKLD, Lun Dayeh and Sa'ban

	PKLD ²⁹	Lun Dayeh	Sa'ban
to arrange	* <i>nepipin</i>	<i>mipin</i>	<i>mpén</i>
to give	* <i>meray</i>	<i>meré</i>	<i>mraay</i>
to smell	* <i>muen</i>	<i>muen</i>	<i>mwin</i>
to ask	* <i>mutuh</i>	<i>mutuh</i>	<i>ntew</i>
to burn (tr)	* <i>ngeseb</i>	<i>ngeseb</i>	<i>nsep</i>
to shave	* <i>ngiki</i>	<i>ngiki</i>	<i>ncey</i>
be ashamed	* <i>miguq</i>	<i>migu?</i>	<i>njewə</i>
cover	* <i>kekeb</i>	<i>kekeb</i>	<i>kap (n)</i> <i>ngkap (v)</i>

The basic form of the actor voice, nasal prefix (*N-*) in Sa'ban is *m-*, which contrasts with Lun Dayeh where the basic nasal prefix is *ng-/nge-*. A Sa'ban verb can be marked in other ways, and these are: gemination of the initial consonant of the root (for example, *lluen* 'to roll up', *rrúet* 'to sew', and *nnal* 'to see'³⁰), or use of the prefix *IV-* (*le-*, *la-*, *lu-*,

²⁹ PKLD examples here and elsewhere in this paper are taken from Blust (1999).

³⁰ In Lun Dayeh these would be *ngelulun*, *ngerut* and *nier* respectively.

for example *luwat* in example (52)),³¹ which may be a borrowing from Kenyah or Kayan (Galvin 1967, *passim*; Southwell 1990:509). There are also several common verbs which occur without any affixation in actor voice, for example *alaak* 'to take', *aro?* 'to do/make', *abeh* 'to carry on the back', *kaap* 'to look for'. All such verbs appear to function in the same way as a verb marked with a *N-* actor voice affix.

Examples of verbs marked for actor voice with a *N-* prefix are given in (51 (*mawaal*), 58–59 (*mpal*)), with initial geminate consonants in (38 (*nnet*), 49 (*nnal*)) and with a *IV-* prefix in (52 (*luwat*)). Examples of the unaffixed verb roots, *aroə* and *alaak* occur in (37 and 39).

Perfective aspect in actor voice is not signalled by affixation as it is in Lun Dayeh. Instead, the verbal auxiliary *pi* 'finished' or 'already' is placed before the verb, which retains its actor voice marking, as illustrated in example (36). A very common construction with *pi* is: 'after doing x he did y' as illustrated in example (37).

- (36) *Dei ieh lé? ieh pi mataay.*
 then 3S (AV):know 3S finish AV:die
 'Then he knows that he has died.'

- (37) *Pi deh aro? nguen ieh iwét deh nguen*
 finish 3P (AV):make coffin 3S (UV).PF:take 3P coffin
ay may lem maaə ay.
 SP go in house SP
 'After they have made his coffin they take it into the house.'
 (it is taken into the house).

Undergoer voice is not morphologically marked in Sa'ban, but the affix *i-* that signals perfective aspect occurs only in undergoer voice. This affix has no allomorphs in Sa'ban, but there is a conditioned variant *y-* which occurs before roots beginning with the vowel *a*, and this has an optional variant *j-* [d₃], for example, *yawaal* or *jawaal* 'called' (*mawaal* 'to call'). Undergoer voice verbs in perfective aspect (*iwét* 'taken', *itap* 'bitten', *ibút* 'poured') are illustrated in examples (37–39).

- (38) *Pi nnet nah Ra' Kueng itap Ra' ay hroel*
 finish AV:climb PT Ant Giant PF.(UV):bite Ant AP egg
ssúek ay.
 macaque AP
 'After Giant Ant had climbed up, he bit the macaque on his testicles.'

- (39) *Deh alaak hloeng ieh, aro? ieh lem ajueng, ibút*
 3P (AV):take bone 3S (AV):do 3S in gong PF.(UV):pour
deh pey? réek may wan hloeng ay.
 3P water ricewine go on bone AP
 'They take his bones and put them (him) in a gong, then they pour rice wine over the bones.'

³¹ In some cases *IV-* appears to make a root transitive or causative, e.g. *manoot* 'float' (by itself) and *lanoot* 'to make something float', or *wak* 'fence' and *luwak* 'make a fence'. On the other hand, there are verbs such as *kku?* and *leku?* which both mean 'to sit' (intransitive) and are used interchangeably in Lg Banga' today.

2.2.1.2 *Undergoer voice in imperfective aspect: a periphrastic construction*

Undergoer voice is not marked morphologically on the Sa'ban verb. In imperfective aspect, it occurs as a periphrastic construction formed with the verb *aroʔ* 'do/make'. In fact, depending on the form of *aroʔ*, both imperfective and perfective periphrastic constructions can be produced.³² The main verb in these constructions retains its actor voice markings but these are cancelled by the signals on the preceding auxiliary verb *aroʔ*. This occurs as *an*³³ in imperfective constructions and *aroʔ* in perfective constructions and is followed by the non-pivot (NP) actor. The following diagram, in which curly brackets indicate the affixes that are cancelled, shows this periphrastic construction:

Imperfective	<i>an</i>	+	NP actor	+	{AV}:verb
Perfective	<i>aroʔ</i>	+	NP actor	+	{AV}:verb

The imperfective periphrastic undergoer construction is illustrated in example (40) and the perfective in (41–42). The verbs, *puet* 'jump' and *kkuʔ* 'sit', in (40) are unaffixed roots which are typical of Sa'ban actor voice intransitive verbs.

- (40) *Puet deh kkuʔ deh wan ddueng lemluen,*
 (AV):jump 3P (AV):sit 3P on shoulder person(s)
an deh mluʔ oeng lún.
 do:UV 3P {AV}:cut throat person(s)
 'They jump up and sit on peoples' shoulders and cut their throats.'

- (41) *Aroʔ lún mpat ieh koʔduet ún ddéeng.*
 (UV.PF):do person(s) {AV}:back 3S like leaf dried.in.sun
 'They left her behind like a withered leaf.'

- (42) *Aroʔ ieh mayong luwak ley raah, Yomban ay.*
 (UV.PF):do 3S {AV}:spear chest man big Yomban SP
 He speared the respected elder, Yomban, in the chest.'

2.2.1.3 *Undergoer voice in imperfective aspect: fossilised remains of a former suffix*

A very few Sa'ban verbs have retained fossilised remnants of the undergoer voice morpheme *-en*. These are *rien* (*mraay* 'to give'), *an* (*aroʔ* 'to do/make'), *nan* (*maan* 'to eat'), *aben* (*abeh* 'carry on back'), and *apan* 'take'.³⁴ Examples (43–45) illustrate the use of *apan* (43), *rien* (44) and *aben* (45).

³² This type of construction also occurs in Lun Dayeh, where it can occur in imperfective and perfective aspect and imperative mood (Clayre 2002).

³³ *Aroʔ* is one of the few Sa'ban verbs which retains a fossilised remnant of the PAn undergoer voice suffix *-en*. In some dialects of Sa'ban, such as P. Riew, *an* [ʌn] is realised as [ɔn]. The form *an* is a reduced form of *ruen*, *tuʔen*, and *uen*, which are the equivalent forms in Lun Dayeh, Kelabit, and Sembudud dialects respectively, see examples (73–76) and Table 12. Although the Sa'ban form of *uen* is *an*, *uem* occurs in Sa'ban imperative constructions (see (55)), the final consonant of the word being the fossilised remains of the set II second person pronoun *mu*.

³⁴ The verb *apan* 'take' (and *apuem* 'take.IMP-2S') seems to exist only as undergoer voice forms. For example:

- (43) *Apan deh kat ay an deh mrang ieh wan ataw.*
 take:UV 3P fish AP do:UV 3P {AV}:place 3S on stone.
 'They take the 'kat' fish and they place it on a stone.'
- (44) *Si' pah naan ataw nok la? rien namay ay.*
 PT where place stone REL want give:UV NPNA:1EP AP
 'Where is the stone that (you) want to give to us?'
- (45) *Ken aben ceh tah kaang nai ngaan anak ceh nah.*
 QP carry:UV 2S PT backpack this with child 2S PT
 'Are you going to carry this backpack as well as your child?'³⁵

2.2.1.4 Sa'ban intransitive verbs

The intransitive actor voice infix *-um-* (*-em-* in Lun Dayeh) occurs as *m-* before vowel-initial roots in Sa'ban, for example, *malaan* 'to walk', but the form *-em-* is rare and occurs only as a fossilised form in verbs such as *temdiew* 'to wail' and *lemdiek* 'to slash undergrowth'.³⁶ Some intransitive verbs have geminate initial consonants, for example, *rruet* 'to descend', *nnet* 'to climb' (example (38)), but most occur as roots, for example, *puet* 'to jump', *kku?* 'to sit' (example (40)), *sieng* 'to arrive', *diep* 'to sleep' (example 46), *languey* 'to swim', or *hlut* 'to fly'.

- (46) *Malaan Ataw Apuey iraat lem hnong yaay dey ieh*
 AV:walk Stone Fire from in headwaters yonder then 3S
sieng lem Pey? Angaw nah, dey ieh diep si? nah.
 (AV):arrive in river Angaw PT then 3S (AV):sleep PT that
 'Fire Stone walked/journeyed from yonder headwaters, then he arrived at the river Angaw, and slept there.'

2.2.1.5 Imperatives in Sa'ban

There are no imperative suffixes in Sa'ban. Periphrastic undergoer constructions (47) or verbs marked for actor voice (48) are used as imperatives, although the former is regarded as more polite.

- (47) *An ciem ngkoet tana? nok-nay!*
 do:UV 2P {AV}:dig ground REL-this
 'Dig this ground!'
- (48) *May ceh alaak ayew ntaah!*
 AV:go 2S (AV):take wood NPNA:1ID
 'Go and collect firewood for the two of us!'

Apan ciem al, luwa taam.
 take:UV 2P fowl smear with blood 1P
 '(You plural) Get a chicken and let us smear ourselves with (its) blood.'

³⁵ The verb *abeh* means to carry on the back, and a *kaang* is a basket designed to be carried on the back.

³⁶ In Lun Dayeh these verbs are *temido* and *lemidik* (PKLD **l-em-idik*) respectively.

2.2.1.6 Other verb affixes in Sa'ban

Unlike Lun Dayeh, where stative verbs are marked by the prefixes *me-* or *ne-*, in Sa'ban stative affixation varies according to the verb root, and generally refers only to completed state. In the following examples both the stative and the actor voice forms of the verb are given: *hmat* 'held' (*mmat* 'to hold'), *hlén* 'bolted' (*llén* 'to lock/bolt'), *hlaa?* 'thrown down' (*mlla?* 'to throw/scatter'); *hrap* 'immersed' (*rrap* 'to immerse'), *belséeng* 'bright' (*leséeng* 'to make light'), *peldok* 'muddy' (*ledok* 'to muddy', *dok* 'mud'), *plaap* 'be inside something' (*llaap* 'put something inside another'), *palaay* 'tame' (*malaay* 'to tame'); *ssin* 'left behind' (*nsin* 'to leave behind'), *tto?* 'fall/drop' (*nto?* 'drop' (intentionally)), *tapét* 'stuck' (*mapét* 'stick something'), *telpew* 'freed' (*lepew* 'free something'), *ta?oot* 'be afraid' (*ma?oot* 'to frighten'), *telnyet* 'obscured' (*lenyet* 'to obscure'). A stative form is occasionally used to indicate involuntary action, for example *tto?* 'fall/drop', but ability to perform an action is indicated by the auxiliary verb *parap* 'able' and not by verb morphology. Stative verbs are not discussed further here.

One other affix found in Sa'ban is the 'reflexive' prefix *se-/s-*, as in *sapueng* 'hide oneself'. Continuous or multiple action or actors is indicated by reduplication of the full verb root.

2.2.2 Sa'ban pronouns and noun markers

Table 5: Sa'ban pronoun sets

	I	(II)	III
1S	<i>éek</i>	<i>(-?)</i>	<i>ne?éek</i>
2S	<i>ceh</i>	<i>(-m)</i>	<i>nceh</i>
3S	<i>ieh</i>	<i>(-n)</i>	<i>nyeh</i>
1IP	<i>taam</i>	–	<i>ntaam</i>
1EP	<i>amay</i>	–	<i>namay</i>
2P	<i>ciem</i>	–	<i>nciem</i>
3P	<i>deh</i>	–	<i>ndeh</i>

There are two sets of pronouns in Sa'ban, I and III, and fossilised remains of pronouns of set II.³⁷ Pronouns of set I fulfil the functions of sets I and II in Lundayeh; that is, they signal the pivot of the sentence (the actor in actor voice clauses and the undergoer in undergoer voice clauses), the actor in undergoer voice clauses, and following a noun, the possessor. In addition, in Sa'ban these pronouns also signal the undergoer in actor voice clauses, a function fulfilled by set III pronouns in Lun Dayeh. Table 6 compares the functions of the pronoun sets in Lun Dayeh and Sa'ban. Brackets are used to indicate the functions of the fossilised pronouns.

³⁷ In addition, Sa'ban has complete sets of dual, and paucal (3–10 people) pronouns. These show dialectal variations; for example, 'they two' is *daweh* in the P. Riew dialect and *déeh* in the P. Angaw dialect.

Table 6: A comparison of pronoun functions in Lun Dayeh and Sa'ban

Language	Actor voice			Undergoer voice			Possession
	Actor	Undergoer	Oblique	Actor	Undergoer	Oblique	
Lun Dayeh	I	III	III	II	I	III	II
Sa'ban	I	I	III	I(II)	I	III	I(II)

The function of a pronoun in a clause is shown by its constituent order,³⁸ and when present, by the affix on the verb, as shown in examples (49) and (50). Examples (49, 51–53) are actor voice clauses. The verb is marked by an actor voice prefix, and the non-pivot undergoer, a set I pronoun, immediately follows the verb. Example (50) is an undergoer voice clause, in which the verb is marked for perfective aspect (which occurs only in undergoer voice) and the non-pivot actor, a set I pronoun, follows the verb.

- (49) *Éek nnal ieh.*
 PIV.1S AV:see 3S
 'I see him.'
- (50) *Éek inal ieh.*
 PIV.1S (UV).PF:see 3S
 'He saw me.'
- (51) *Wey? ceh mawaal éek, dey éek aray apu? ceh.*
 good PIV.2S AV:call 1S then PIV.1S (AV):come meet 2S
 'Call me, then I will come and meet you.'
- (52) *Oh, dey nah éek pi lataat ceh ina in*
 oh, then PT PIV.1S already advise 2S just now don't
luwat ieh.
 (AV):empty out 3S
 'Oh, I told you just now not to empty it out.'
- (53) *Deh temdiew ieh iraat wan sataat yaay.*
 PIV.3P <AV>wail 3S from on ground³⁹ yonder
 'They wail (for) her from down there on the ground.'

Sa'ban set III pronouns, like Lun Dayeh set III pronouns, occur with a proclitic *ne=* or *n=*, but whereas Lun Dayeh pronouns are basically set II pronouns with the proclitic, the Sa'ban set III pronouns are basically set I pronouns with the proclitic. Furthermore, Sa'ban set III pronouns only signal an oblique argument as illustrated in examples (48) and (54). They do not signal the undergoer in actor voice as they do in Lun Dayeh.

- (54) *Ndeh nok-nay ciek eu-eu iraray ieh ne?éek*
 why REL-this small very (UV).PF:give 3S NPNA:1S
 'Why did he give me this very small (portion)?'

³⁸ See §2.2.3

³⁹ The *sataat* is an area of ground around a house which is kept clear of weeds and rubbish.

Table 5 lists three singular pronoun forms which are similar to the pronouns of set II in Lun Dayeh. These are: *-ʔ* (1S), *-m* (2S) and *-n* (3S). They occur as fossilised remnants at the end of certain nouns which are kindred terms; for example: *hnaa-ʔ* ‘my mother’, *hnaa-m* ‘your mother’, and *hnaa-n* ‘his/her mother’, *tama-ʔ* ‘my father’, *tama-m* ‘your father’ and *tama-n* ‘his/her father’. The third person form is also the general word for ‘mother’ and ‘father’. The same Sa’ban verbs that retained fossilised remains of the undergoer suffix *-en* have also retained fossilised remains of these pronouns. They occur in undergoer voice imperative constructions, for example: *rie-m* ‘(you) give!’, *noe-m* ‘(you) eat!’, *apue-m* ‘(you) take!’, and *ue-m* ‘(you) do!’ as illustrated in example (55).

- (55) *Uem ntoʔ siʔ úeng ay nan éek, ih!*
 do:UV.2S(GEN) {AV}:drop one banana AP eat:UV 1S FP
 ‘Friend, throw down a banana for me to eat!’

There is no noun phrase marker for common nouns or proper nouns in Sa’ban, but *n-* or *nah* occurs before oblique nouns that refer to humans, as for example in (56–57).

- (56) *Dey ssúel ieh mraay léʔ n=lun awéeng*
 then wife GEN.3S AV:give know PM.NPNA=people village
nok-nah ...
 REL-that
 ‘Then his wife informs the people of that village ...’
- (57) *Hmew ssúel ieh nah=hley ieh nok pi*
 word wife GEN.3S PM.NPNA=husband 3S REL finish
mataay nah ...
 die PT
 ‘(His) wife says to her husband who has died ...’

2.2.3 Word order in Sa’ban

In Sa’ban, as in Lun Dayeh, the non-pivot core argument must follow immediately after the verb, for example (49) and (50) in actor voice and undergoer voice respectively. In actor voice transitive clauses the pivot core argument usually occurs before the verb, as for example (58), but it can occur after the verb-plus-non-pivot core argument, as in (59); it may not occur immediately following the verb (60).

- (58) *Éek mpal kuʔ nok-nah.*
 PIV.1S AV:beat dog REL-that
 ‘I beat that dog.’
- (59) *Mpal kuʔ nok-nay éek.*
 AV:beat dog REL-this PIV.1S
 ‘I beat this dog.’
- (60) * *Mpal éek kuʔ nok-nay.*
 AV:beat PIV.1S dog REL-this
 *‘This dog beat me.’

In actor voice intransitive clauses, the actor may occur before or after the verb, for example (38, 40, 46). In undergoer voice clauses the undergoer pivot is relatively free and may occur before the verb-plus-non-pivot core argument, as illustrated in (50, 54) or after it, as illustrated in (37, 38). Oblique arguments (marked by set III pronouns) occur after the verb and its core arguments (48, 54).⁴⁰

3 Kelabitic dialects or languages in Kalimantan Timur

In Kalimantan Timur on the other side of the border from Sarawak and Sabah is a district called Kerayan, which takes its name from the Kerayan river, one of the headwaters of the Mentarang-Sesayap river system. The people living in the Kerayan were known to speak dialects related to Lun Dayeh, but references in the literature were brief. Indeed, Jay Crain, writing in 1978, commented on 'the lacunae of knowledge about the heartland of the [Lun Dayeh] population in the interior of East Kalimantan' (Crain 1978:123). These Kerayan dialects are of particular interest linguistically, since they are located geographically between Lun Dayeh spoken in the north, and Sa'ban in the south of a region that is only fifty-seven miles long (Schneeberger 1945). The first serious attempt to study the languages of the Kerayan was made in 1993 by Bernard Sellato who, in addition to many word lists, also collected information on cultural features and history (Sellato 1997).⁴¹ The research for this paper was carried out in July 1995. A month had been set aside for the project, but bad weather seriously delayed flights to and from the area,⁴² and prevented trips to the more remote villages. Twenty-three language informants were interviewed in the Kerayan, five in Samarinda and one in Tarakan.⁴³ They were, with few exceptions, mature adults and their details are given in Appendix 1, and the location of their home villages shown on Map 2.

The linguistic picture in the Kerayan district has been made more complex by the policy carried out in the early 1970s, of regrouping or relocating villages, whereby eighty-nine villages (each probably with its own minor dialect) were reduced to twenty-seven. Today, as a result, two or more distinct dialects may be spoken in one village, as for example at Lg Layu, where Sa'ban, Lengilo' and other dialects are spoken, and Sellato observed in 1993 that regional *linguae francae* were developing (Sellato 1997).⁴⁴

⁴⁰ In stative verb clauses the verb usually occurs in clause-initial position. For example:

Ttuet nah kuet ieh pi? abieng.
 smashed PT leg 3S LP left
 'His left leg was completely smashed.'

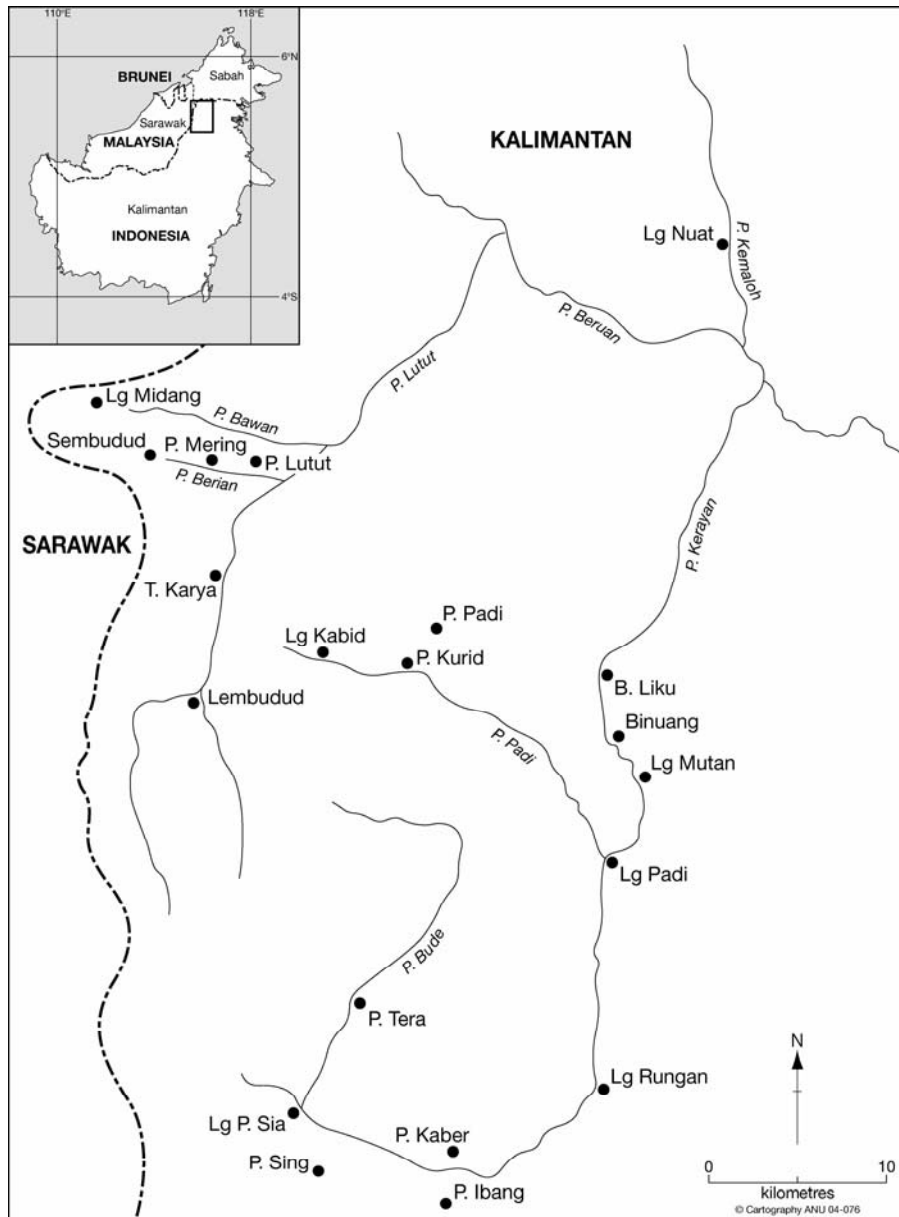
⁴¹ Dr Sellato generously made his word lists available to me.

⁴² The Kerayan region is a plateau, 1000–1100 m high, and access from the east coast of Kalimantan is by small, usually single-engine, aircraft.

⁴³ Language work in Samarinda concentrated on a study of the pronouns, and of the reflexes of Blust's voiced aspirates (**b^h*, **d^h*, **g^h*). Following the Kerayan survey I interviewed a Kelabit, John Terawé, from Bario, Sarawak.

⁴⁴ Sellato has recorded in detail, all the known movements of people (and therefore of dialects) in the Kerayan area within living memory (Sellato n.d.).

Sellato tentatively distinguished four major dialect groups in the Kerayan region: Lun Dayeh in the northwest, Kelabit, including Lun Baa' in the west, Lengilo' in the mid and upper Kerayan, and Sa'ban further south. These labels, with minor modifications, are followed here.



Map 2: Villages in the Kerayan district, Kalimantan Timur

The term Kelabit will be restricted to the language spoken by the Kelabit people who live across the border in Sarawak in an area known as the Kelabit highlands. The relationship between the Kelabit language and Lun Dayeh has long been recognised. Kelabit word lists and vocabularies have appeared in various publications since 1896, but the most authoritative is Blust's 'Kelabit-English vocabulary' published in 1993 (Blust 1993; Bolang and Harrison 1949:119-122; Douglas 1911; Ray 1913:155-196; Roth 1896:2, appendix:103; Southwell 1949:104-105). The ethnonym Lun Baa' will be used to

refer to the 'Kelabits' in the Kerayan region. These are people who cultivate wet rice and inhabit the Bawan and Belawit valleys in the west of the Kerayan area. (Langub 1987:33; Schneeberger 1979:30; Sellato 1997; Southwell 1949:105–107).

The Lengilo' are probably the people recorded in 1939 by Schneeberger as Ulun Milau and Ulun Kerayan. They were located in the east of the region along the middle reaches of the Kerayan river, but Sellato records that some Lengilo' have since moved into the upper reaches of that river and are living alongside Sa'ban who have migrated there from the upper Bahau (Schneeberger 1979:22, 29–31; Sellato 1997 and n.d.).

Two general terms will be used to refer to the remaining dialects in the survey: central dialects and upper Kerayan dialects. This is because the dialects spoken in these areas are diverse and cannot yet be included under a single umbrella linguistic or ethnic term. Central dialects consist of a small group of dialects clustered around the Kurid river in the centre of the region. Sellato identified them as a subgroup of Lun Baa' called Kurid, but central dialects, such as that of P. Padi, exhibit interesting differences from the Lun Baa' dialects and for this reason have been kept separate here.⁴⁵ The upper Kerayan area is located in the south of the Kerayan region and, as a result of the relocation policy, the dialects spoken there are more diverse and include major groups such as Sa'ban and Lengilo'.

3.1 Verbal affixation in Kerayan dialects

Kerayan dialects have only two voices: actor and undergoer. No examples of instrument voice were found outside Lun Dayeh. Compared to Lun Dayeh there is a noticeable reduction in the number of affixes occurring with verbs. There are for example no suffixes, and the number of pronoun sets is also reduced.

3.1.1 Actor voice

3.1.1.1 Imperfective aspect

In transitive verbs, most Kerayan dialects signal actor voice by a nasal prefix (*N-*), and imperfective aspect is unmarked. In Lun Dayeh the basic actor voice prefix is *ng-/nge-*, but many of the Kerayan dialects mark actor voice more commonly with the prefix *m-/me-*, as in Sa'ban. These are the Lun Baa' dialects, and two dialects of the upper Kerayan, P. Sing, and P. Kaber, while the dialects that use *ng-/nge-* are the Lengilo' dialects, and P. Tera. Kelabit appears to belong to the latter group. The P. Padi dialect is remarkable for its general lack of actor voice affixation. Table 7 illustrates actor voice affixation on verbs from five Kerayan dialects, including P. Padi.⁴⁶

⁴⁵ They could be referred to as Kurid dialects, but I prefer to use a neutral term for them until we have a better definition of what characterises a Kurid dialect. According to Sellato the P. Padi dialect is close to Kurid, and the P. Padi people came to the area from Ba Kelalan (Sarawak) six generations ago (Sellato n.d.).

⁴⁶ The Kerayan dialects have not yet been analysed phonemically, and spelling in this paper generally follows the system devised by Blust for Kelabit. The labiodental fricative which is an allophone of /p/ in Lun Dayeh has here been written as [f] since it is an indicator of dialectal difference (see §4.3); [ə] and [ʌ] have both been written as e, as in Kelabit, but given the fact that they contrast in Sa'ban, it is possible that they are separate phonemes in some of the upper Kerayan dialects.

Table 7: Five Kerayan dialects with verbs marked for actor voice

Verb	Lun Dayeh	B. Liku (Lengilo')	Sembudud (Lun Baa')	Sa'ban	P. Padi (Central)
drink	<i>ngirup</i>	<i>ngirop</i>	<i>merop</i>	<i>merop</i>	<i>irop</i>
bite	<i>ngetep</i>	<i>ngetep</i>	<i>metep</i>	<i>ntep</i>	<i>tep</i>
see	<i>nier</i>	<i>niel</i>	<i>miel</i>	<i>nnal</i>	<i>iel</i>
kill	<i>ngaté</i>	<i>ngaté</i>	<i>maté</i>	<i>mataay</i>	<i>ngaté</i>
burn something	<i>ngeseb</i>	<i>ngeseb</i>	<i>meseb</i>	<i>nsep</i>	<i>seb</i>
take	<i>ngalap</i>	<i>ngalap</i>	<i>malap</i>	<i>alaak</i>	<i>alap</i>
fan	<i>nefer</i>	<i>ngefel</i>	<i>mefer</i>	<i>mpel</i>	<i>fel</i>
tear	<i>ngeraak</i>	<i>neraak</i>	<i>meraak</i>	<i>rraak</i>	<i>neniit</i>
cover	<i>ngekeb</i>	<i>ngekeb</i>	<i>mekeb</i>	<i>ngkap</i>	<i>nekeb</i>
wash s.th.	<i>meru?</i>	<i>muro?</i>	<i>muro?</i>	<i>mru?</i>	<i>uro?</i>
shoot with blowpipe	<i>ngefut</i>	<i>ngeput</i>	<i>meput</i>	<i>mpuet</i>	<i>put</i>
do, make	<i>nganau?</i>	<i>naro?</i>	<i>maro?</i>	<i>aro?</i>	<i>aro?</i>

In the languages of Sabah, the typical actor voice prefix *mong-* (*m-pong-*), has been analysed as *m-* signalling 'actor voice' and *-pong-* signalling 'transitivity' (Kroeger 1988:222). In Lun Dayeh and all the Kerayan dialects, including Sa'ban, there is only the one prefix, but why some languages mark actor voice predominantly by *ng-/nge-* and others by *m-/me-* is not obvious. Lun Baa', Sa'ban and the other Kerayan dialects are not alone among Bornean languages in opting for *m-/me-* as the basic actor voice prefix. In Berawan (Sarawak) the largest subgroup of transitive verbs takes *me-/m-* in actor voice, imperfective aspect (Clayre 1994:237), but this prefix (or a homophonous prefix) was also found with intransitive and stative type verbs. For example, *mittéen* 'to stand', *matoong* 'float', *mittam* 'black', *meliram* 'bruised'. A similar situation occurs in Melanau, a coastal language of Sarawak, for example, *maid* 'wipe', *mepit* 'throw away', *mungu* 'sit', *murék* 'climb up', *muaw* 'confused', *mama* 'dirty'.

Among Kerayan dialects, Lun Dayeh appears to be the more unusual in using *nge-/ng-* to signal actor voice, but Lun Dayeh also has a productive stative verb system, marked by the prefix *me-/m-*,⁴⁷ and the two systems, dynamic and stative, are kept separate by the use of *nge-/ng-* to mark actor voice.

3.1.1.2 Perfective aspect in actor voice

Differences in signalling perfective aspect have already been noted between Lun Dayeh, where it is signalled morphologically by the prefix *ne-/n-*, and Sa'ban, where it is signalled by an auxiliary verb, *pi* 'finish'. In Kerayan dialects both forms, the prefix and the auxiliary, are used, and, depending on the dialect, the verb co-occurring with the auxiliary may be marked by the perfective aspect prefix *ne-/n-*, or simply by the actor voice prefix *N-*.

⁴⁷ See §2.1.1.4.

All Kerayan dialects, except P. Sing and P. Kaber, produced at least a few examples of perfective aspect being marked by the prefix *ne-* added to a verb marked for actor voice as illustrated in example (61). In questions such as 'Who did x?', there were examples in some dialects (Kelabit, B. Liku, Binuang and P. Kurid) of verbs being marked either by *ne-* or only by the actor voice prefix, *N-*. Compare examples (61) and (62) both from Binuang.

- (61) *Jé nengekab fě?*
 PIV.who PF.AV:cover water
 'Who covered the water?'
- (62) *Jé muro? tiok nih.*
 PIV.who AV:wash spoon(s) this/these
 'Who washed this/these spoon(s)?'⁴⁸

In addition to signalling perfective aspect by adding the prefix *ne-* to the actor voice form of the verb, many dialects in the north and mid Kerayan areas (for example, Lg Midang, P. Lutut, Lg Padi, Lg Mutan) add the auxiliary verb *pengeh* 'finish' (or the dialectal equivalent) to the construction, as illustrated in example (63) from Lg Mutan.

- (63) *Iek pengeh nemefer apuy.*
 PIV.1S finish PF.AV:fan fire
 'I fanned the fire.'

In other dialects of the same area of the Kerayan (for example, B. Liku, Binuang, T. Karya), the auxiliary *pengeh* may occur either with a verb marked with the perfective prefix *ne-*, or with a verb marked only by the actor voice prefix *N-*. In these dialects the latter is the more common construction. Compare the examples illustrated in (64a and b) from Binuang. In (64a) *pengeh* co-occurs with a verb marked for perfective aspect by the prefix *ne-*, and in (64b) *pengeh* co-occurs with a verb marked only for actor voice.

- (64) a. *Wih pengeh neneraak bekad neh.*
 PIV.1S finish PF.AV:tear shirt that
 'I have torn that shirt.'
- b. *Wih pengeh meré bekad neh ko?-so.*
 PIV.1S finish AV:give shirt that NPNA-2S
 'I gave that shirt to you.'

In the upper Kerayan dialects of P. Sing, P. Kaber and P. Tera, perfective aspect is no longer signalled morphologically, but by the auxiliary verb *pengeh*. The main verb is marked only by the actor voice prefix, just as in Sa'ban. This is illustrated in examples (65–66) from P. Sing. The first clause (65) is in imperfective aspect, and the verb is marked by an actor voice prefix *m-*, the second (66) is a clause in perfective aspect, signalled by the auxiliary *pengeh*, and the main verb is marked by the actor voice prefix *m-*. These may be compared with example (36) from Sa'ban.

⁴⁸ The language helper commented at the time that *nemuro?* was not in common use. Most verbs checked in this construction in Binuang were, however, marked for perfective aspect, the remainder were marked only for actor voice.

- (65) *Wih meraak bekad.*
 PIV.1S AV:tear shirt
 ‘I tear the shirt.’
- (66) *Wih pengeh meraak bekad.*
 PIV.1S finish AV:tear shirt
 ‘I have torn the shirt.’

The Sembudud dialect was unusual as the perfective prefix *n-*, rather than being added to the actor voice prefix *m-*, actually replaced it. For example *meraak* and *neraak*, ‘tear’ and ‘tore’ and *maro?* and *naro?* ‘do’ and ‘did’ (Table 8). Examples (67a–b) from Sembudud illustrate this with the verb *muro?* ‘to wash something’.

- (67) a. *Idan ko muro? seduk neh?*
 when PIV.2S AV:wash spoon that
 ‘When will you wash that/those spoon(s)?’
- b. *Yé nuro? seduk neh?*
 who PF:wash spoon that
 ‘Who washed that/those spoon(s)?’

Occasional examples of a similar usage were noted in T. Karya, where, for example, *mefer* ‘fan’ (IMPF) and *nefer* (PF) occur. In Lun Dayeh this pattern of replacing *m-* by *n-* occurs only with stative verbs, but in Sembudud these are not stative verbs, nor does Sembudud have a stative system comparable to Lun Dayeh. It appears that *m-* in Sembudud signals both actor voice and imperfective aspect, and *n-* actor voice and perfective aspect.

The P. Padi, a dialect that eschewed the use of prefixes in actor voice, generally prefixed verbs with *n-* or *ne-* in perfective aspect. Table 8 gives examples of verbs from five Kerayan dialects, including P. Padi, to illustrate the dialectal variations that occur in the affixes signalling actor voice and perfective aspect.

In actor voice (imperfective aspect) Kerayan dialects fall into two main groups according to the form of the prefix predominantly used to signal actor voice: *nge-/ng-* or *me-/m-*. In Lun Dayeh and Lengilo’, to the north and east of the region, *nge-/ng-* is the predominant form of the prefix, and in Lun Baa’ and Upper Kerayan dialects (except P. Tera) in the west and south of the region, *me-/m-* predominates. The central dialect of P. Padi is unusual in having lost any marking for actor voice.

Perfective aspect marking in actor voice is much less clear, but does show a more north-to-south distinction. In the north in Lun Dayeh the main way of signalling perfective aspect is by the prefix *ne-*, while *pengeh* rarely occurs. In the upper Kerayan area in the south the perfective prefix *ne-* has been lost, and only the auxiliary verb *pengeh* (or a dialectal equivalent) occurs. In between these two extremes, the auxiliary *pengeh* commonly occurs in perfective constructions, sometimes with a verb also marked by the perfective prefix *ne-* and sometimes with verbs marked only by the actor voice prefix *N-*. The Sembudud dialect is idiosyncratic and replaces the actor voice prefix *m-* by the perfective aspect prefix *n-*.

Table 8: The imperfective and perfective forms of verbs in actor voice from five Kerayan dialects

Dialect	verb	Actor voice (imperfective)	Actor voice (perfective)
Lun Dayeh	give fan wash (dish) do/make shoot (with blowpipe) cover	<i>meré</i> <i>meh^her</i> <i>meru?</i> <i>nganau?</i> <i>ngeput</i> <i>ngekeb</i>	<i>nemeré</i> <i>nemeb^her</i> <i>nemeru?</i> <i>nenganau?</i> <i>nengeput</i> <i>nengekeb</i>
L. Mutan (Lengilo')	give fan wash shoot cover tear	<i>meray</i> <i>mefer</i> <i>nemero?</i> * <i>ngeput</i> <i>ngekab</i> <i>ngederaak</i>	<i>nemeray</i> <i>nemefer</i> <i>nemero?</i> <i>nengeput</i> <i>nengekab</i> <i>nengederaak</i>
Sembudud (Lun Baa')	give fan wash do/make tear cover	<i>meré</i> <i>mefer</i> <i>muro?</i> <i>maro?</i> <i>meraak</i> <i>mekab</i>	<i>neré</i> <i>nefer</i> <i>nuro?</i> <i>naro?</i> <i>neraak</i> <i>nekab</i>
P. Padi (Central)	give fan wash do/make shoot	<i>ré</i> <i>fel</i> <i>uro?</i> <i>aro?</i> <i>put</i>	<i>ré</i> <i>nefel</i> <i>nuro?</i> <i>naro?</i> <i>neput</i>
Sa'ban	give fan wash do/make cover shoot	<i>mraay</i> <i>mpel</i> <i>mru?</i> <i>aro?</i> <i>ngkap</i> <i>mpuet</i>	auxiliary verb, <i>pi</i> 'finished', plus main verb marked for actor voice

* All other Lengilo' dialects had *mero?* or *muro?* as the imperfective form of this verb

This variation in usage over such a small area suggests that the morphological marking of perfective aspect, as seen in Lun Dayeh, is giving way to the use of a verb auxiliary as seen in Sa'ban. Table 9 shows the predominant markings for perfective aspect (in actor voice) in Kerayan dialects.

Table 9: The marking of perfective aspect in actor voice declarative clauses in Kerayan dialects

Marking	Dialect
+PF+AV+verb	Lun Dayeh,
Aux +PF+AV+verb	Kelabit, Lg Mutan, Lg Padi (Lengilo'); Lg Midang Lun Baa'); P. Kurid (Central)
Aux +PFR +verb	Sembudud (Lun Baa')
Aux +PF+(AV)verb	P. Padi (Central)
Aux +PF+AV+verb and Aux +AV+verb	B. Liku, Binuang (Lengilo'); T. Karya (Lun Baa')
Aux +AV+verb	P. Kaber, P. Sing, P. Tera (U.Kerayan); Sa'ban

Aux indicates the use of the auxiliary verb, *pengeh*.

PF and AV represent the appropriate aspect and voice prefixes

R indicates that the imperfective prefix is replaced by the perfective prefix

(AV) indicates that actor voice is unmarked

The replacement of actor voice (imperfective) prefix *m-* by actor focus perfective prefix *n-* that occurs in Sembudud contrasts with the usage in the Berawan language of Sarawak. In Berawan *m-* signals actor voice imperfective, and *n-* undergoer voice perfective. For example: *muppo?* 'beat', *nuppo?* 'beaten'; *merrah* 'scatter', *nerrah* 'scattered'; *malaa?* 'take', *nalaa?* 'taken'; *mitéén* 'split', *nitéén* 'split' (Clayre 1994 and fieldnotes). Also some verbs in the Melanau language signal actor voice imperfective by the prefix *m-*, and undergoer voice perfective by replacing *m-* by *n-*. For example, *miib* 'tow' *niib* 'towed'; *mebin* 'carry on back' *nebin* 'carried on back', *makup* 'scoop up' *nakup* 'scooped up' (Clayre 1996:73).

3.1.1.3 The intransitive infix *-em-* in Kerayan dialects

The infix *-em-* that occurs in Kerayan dialects is a reflex of the PAN affix **-um-*. The infix is realised as *-um-* only with the verb *kuman* 'eat'. In all other verbs the affix is realised as *-em-*, or *m-* before vowel-initial stems. This affix, which marks actor focus intransitive verbs, is found in Lun Dayeh, Kelabit and most Lengilo' dialects (Table 10). A metathesised form occurs in Lun Baa' dialects, illustrated by *melanguy* (Table 10). Otherwise in some Lun Baa' dialects, and in all dialects of the central area and upper Kerayan the affix has been lost. In addition, the P. Padi dialect has lost the initial consonant of PKLD **laNuy* 'swim', the dialects of P. Kaber, P. Padi and Sa'ban show loss of the initial syllable of PKLD **ed^hiN* 'arrive', and Sa'ban has lost most of the first syllable of PKLD **tulud* 'fly'.

Table 10: The distribution of the *-em-* affix in Kerayan dialects

Dialect		swim	fly	arrive
PKLD		<i>*laNuy</i>	<i>*tuluk</i>	<i>*ed^hiN</i>
Lun Dayeh		<i>lemanguy</i>	<i>temulud</i>	<i>mecing</i>
Kelabit	Bario	<i>lemanguy</i>	<i>temulud</i>	<i>med^hing</i>
Lengilo'	Binuang	<i>lemanguy</i>	<i>temulud</i>	<i>meséng</i>
	Lg Mutan	<i>temuron #</i>	<i>temulud</i>	<i>meséng</i>
	Lg Padi	<i>temurun</i>	<i>temulud</i>	<i>mesing</i>
	B. Liku	<i>lemangoy</i>	<i>tuluk</i>	<i>mesé:n</i>
Lun Baa'	P. Lutut	<i>melanguy</i>	<i>temulud</i>	<i>mesing</i>
	Pa' Mering	<i>melanguy</i>	<i>tuluk</i>	<i>mesing</i>
	T. Karya	<i>melanguy</i>	<i>tuluk</i>	<i>mesing</i>
	Sembudud	<i>melanguy</i>	<i>tuluk</i>	<i>meséng</i>
	Lg Midang	<i>languy</i>	<i>tuluk</i>	<i>mecing</i>
Central area	Lg Kabid	<i>languy</i>	<i>tuluk</i>	<i>mecing</i>
	P. Padi	<i>anguy</i>	<i>tuluk</i>	<i>sén</i>
Upper Kerayan	P. Tera	<i>langoy</i>	<i>tuluk</i>	<i>mesé:n</i>
	P. Sing	<i>languy</i>	<i>tuluk</i>	<i>mirat</i>
	P. Kaber	<i>langoy</i>	<i>tuluk</i>	<i>si:ng</i>
Sa'ban		<i>langoey</i>	<i>hlut</i>	<i>sieng</i>

In Lun Dayeh *temurun* means 'to descend'

3.1.2 Undergoer voice in Kerayan dialects

3.1.2.1 Imperfective aspect

In Lun Dayeh, verbs in undergoer focus (imperfective aspect) are signalled by the suffix *-en* added to the root (Table 1, and example (4)). In Sa'ban this suffix is no longer in use, but there is fossilised evidence for its former existence in the language (see §2.2.1.3, and examples (43–45)).

In Kerayan dialects this suffix has all but disappeared, it remains only as a fossilised form in three commonly used verbs: *beré* 'give' (Table 11); *kuman* 'eat' and *(t)aru?* 'do' or 'make'. For example, *kenen* 'eaten', the undergoer voice (imperfective) form of *kuman* 'to eat', is found in Kelabit and the Kerayan dialects of Lun Baa' (Lutut and Sembudud), Lengilo' (Lg Mutan, B. Liku, and Binuang), and upper Kerayan (P. Tera, P. Kaber). Example (68) is from Binuang (Lengilo').

- (68) *Kenen so luba? nih.*
 eat:UV.IMPF 2S PIV:rice this
 'Will you eat this rice?'

The undergoer voice form of *beré* ‘give’ shows considerable variation between the dialects in the Kerayan, although the actor voice form of the verb is identical in most dialects (Table 11).

Table 11: Voice and aspect markings of the verb *beré* ‘give’ in Kerayan dialects

(P) indicates the use of a periphrastic construction.

Dialect		Actor voice		Undergoer voice	
		IMPF	PF	IMPF	PF
Lun Dayeh		<i>meré</i>	<i>nemeré</i>	<i>beré.n</i>	<i>biré</i>
Kelabit	Bario	<i>meré</i>	<i>nemeré</i>	(P) <i>meré</i>	<i>biré</i>
Lengilo’	Binuang	<i>meré</i>	<i>meré</i>	<i>beré.n</i> (P) <i>meré</i>	<i>biré</i>
	Lg Padi	<i>meray</i>	<i>nemeray</i>	<i>beré.n</i> (P) <i>meray</i>	<i>biray</i>
Lun Baa’	P. Mering	<i>meré</i>	<i>nemeré</i>	<i>beré</i>	<i>iré</i>
	Sembudud	<i>meré</i>	<i>neré</i>	<i>beré</i>	<i>iré</i>
Central area	P. Kurid	<i>meré</i>	<i>neré</i>	(P) <i>ré</i>	<i>iré</i>
	P. Padi	<i>ré</i>	<i>ré</i>	<i>ré.n</i> (P) <i>ré</i>	<i>iré</i>
Upper Kerayan	P. Sing	<i>meré</i>	<i>meré</i>	<i>ré</i> (P) <i>meré</i>	<i>iré</i>
	P. Tera	<i>meré</i>	<i>meré</i>	(P) <i>ré</i> (P) <i>meré</i>	-*
	P. Keber	<i>me yay</i>	<i>me yay</i>	∅	<i>i yay #</i>
Sa’ban		<i>mraay</i>	<i>mraay</i>	<i>rien</i>	<i>iraay</i>

* The undergoer voice (perfective) form was not recorded in P. Tera.

No undergoer voice (imperfective) form could be elicited in P. Kaber. Note too that in this dialect an initial or intervocalic *r* of other dialects is realised as a voiced velar fricative. P. Kaber is not alone in these respects, the P. Aloon dialect of Sa’ban, not included in this survey, also appears to lack an undergoer voice form of the verb ‘to give’, and in that dialect initial and intervocalic *r* is realised as a voiced velar plosive.

Table 11 shows consistent loss of the initial consonant or the initial syllable of the verb *beré* in central and upper Kerayan dialects. The *beré:n* form that occurs in Lun Dayeh and Lengilo’ dialects is illustrated in (69) from Binuang, and the reduced form *ré.n* is illustrated in (70) from the dialect of P. Padi (central). An example of the Sa’ban reduced form, *rien*, is illustrated in (44).

(69) *Bekad neh beré:n wih koʔ=ieh.*
 PIV.shirt that give:UV.IMPF 1S NPNA=3S
 ‘I will give this shirt to him.’

(70) *Yakad nih ré.n wih ki=kyo.*
 PIV.shirt this give:UV.IMPF 1S NPNA=2S
 ‘I will give this shirt to you.’

In some Lun Baa' dialects (for example, Lg Midang, P. Mering, Sembudud) the suffix appears to have been lost altogether, and undergoer voice is signalled by the verb root *beré*, as illustrated in (71) from P. Mering. In the central and upper Kerayan dialects of P. Kurid, P. Tera and P. Sing not only has the suffix been lost, but also the first syllable of the verb root, producing *ré* as illustrated in (72) from P. Sing. In the dialects of Lg Padi, Lg Mutan, Lg P. Sia, P. Kaber and Sa'ban the final vowel is realised as /ay/, which is closer to the PKLD form **meray*.

(71) *Bekad neh beré wih ki=ieh.*
 PIV.shirt that give:(UV) (GEN)1S NPNA=3S
 'I will give that shirt to him.'

(72) *Bekad nih ré wih wan keh.*
 PIV.shirt this give:(UV) 1S for 2S
 'I will give this shirt to you.'

A periphrastic construction, similar to that described for Sa'ban (§2.2.1.2), was produced for undergoer voice imperfective constructions in many of the Kerayan dialects. In some dialects (Binuang, Lg Padi, P. Padi, P. Sing, and P. Tera), both morphological and periphrastic constructions were offered, as indicated in Table 11, for the verb *beré* 'give'. The periphrastic construction consists of the undergoer voice (imperfective) form of the verb (*t)aru?* 'to do/make' followed by the non-pivot actor (NP) and the main verb. The dialectal variations in the undergoer voice form of the verb (*t)aru?* are listed in Table 12 together with a diagrammatic representation of this periphrastic construction. In this construction any voice markings on the main verb are cancelled following (*t)aru?*.

Table 12: The dialectal forms of the verb (*t)aru?* in the periphrastic undergoer construction in Kerayan dialects

Dialect	(<i>t)aru?</i>	+NP actor	+{AV}:verb
Lun Dayeh	<i>rueŋ?</i>	+NP actor	+{AV}:verb
Kelabit	<i>tuŋen</i>	+NP actor	+{AV}:verb
Sembudud, P. Mering, P. Padi, P. Kaber	<i>uen</i>	+NP actor	+{AV}:verb
Lg Padi, Lg Mutan	<i>ngen</i>	+NP actor	+{AV}:verb
P. Tera, Lg Sing	<i>nen</i>	+NP actor	+{AV}:verb
B. Liku, Binuang, P. Kurid	<i>en</i>	+NP actor	+{AV}:verb

Example (73) shows that this type of construction is also known in Lun Dayeh, although it is not so commonly used as it is in Kerayan dialects. Example (74) is from Bario Kelabit, and (75) is from P. Kurid. Here the main verb is not the actor voice imperfective form *meré*, but a reduced form, *ré*. Example (76) illustrating the periphrastic form from Binuang may be contrasted with (69), the morphological construction from Binuang.

(73) *Ruen mu manak eb^ha luk rupen mu.*
 do:UV.IMPF GEN.2S. {AV}:boil water REL drink:UV.IMPF GEN.2S
 'Do you boil the water that you drink?'

- (74) *Tu?en kuh meré sapa? ineh ngen neh.*⁴⁹
do:UV GEN.1S {AV}:give shirt that to 3S
'I will give the shirt to him.'
- (75) *Akad neh en wih ré wan co.*
shirt this do:UV 1S give for 2S
'I will give this shirt to you.'
- (76) *Bakad neh en wih meré ko? =so.*
shirt this do:UV 1S {AV}:give NPNA=2S
'I will give this shirt to you.'

The periphrastic construction is also used as a polite imperative in Lengilo', and in central and upper Kerayan dialects, as illustrated in (77) from Binuang.

- (77) *En so ngelaak luba? neh.*
do:UV.IMP 2S AV:cook rice that
'Cook that rice!'

The undergoer voice suffix *-en* is in regular use in Lun Dayeh, but it does not occur in Kerayan dialects, except as a fossilised remnant with a very few verbs. These remnants are an indication that the suffix was once in use in Kerayan dialects; thus verbs, such as *beré* 'give', provide useful information on the demise of this part of the voice system in the Kerayan area. They also demonstrate the strategies employed, such as the use of a periphrastic construction, to fill the gap left behind.

Some features noted in the use of undergoer voice (imperfective) verbs in the Kerayan are also known from other Bornean languages. A similar periphrastic construction has been recognised in some of the languages of Sarawak, for example Kayan (Clayre 1995). The use of the simple verb root to signal undergoer voice (imperfective), as illustrated by the occurrence of *beré* in Lun Baa' dialects (71), has been noted in languages such as Penan and Melanau in Sarawak, and in some languages of southeast Kalimantan (Clayre 1996:69, 73, 75). The reduction of *beré* 'give' to *ré* in the dialects of P. Kurid, P. Tera and P. Sing is also not without comparative parallels, since a similar phenomenon was noted in Penan, where *ala:?* 'take' occurred in certain constructions as *la:?* (Clayre 1996:69).

3.1.2.2 *Undergoer focus and perfective aspect*

It has already been noted in Lun Dayeh that the perfective aspect infix *-in-* occurs only with verbs in undergoer voice (which is otherwise unmarked), and that the infix *-in-* occurs as *in-* before vowel-initial roots and as an ablaut *-i-* in roots with a schwa as the penultimate vowel.⁵⁰ Table 13 lists the perfective undergoer voice form of six verbs from twelve Kerayan dialects. It is apparent from this table that several phonological processes are at work in these dialects, the final result of which is that in several dialects the

⁴⁹ Blust (1993) glosses *ngen* as 'a particle preceding personal names and titles' in which case it appears to be similar to the personal noun markers *ni=* or *ki=*, and possibly *i=*, in Lun Dayeh (Blust does not record *i=* in Kelabit). He also glosses it as 'to', presumably a preposition, and Amster (1995) glosses it only as a preposition 'for', 'to' 'at'. In Sa'ban the preposition meaning 'to' is *ngaan*. For the purposes of this paper the occurrence of *ngen* in examples (74) and (84) is interpreted as a preposition.

⁵⁰ See §2.1.1.1

perfective infix becomes a prefix. These dialects are: Lun Baa', central, and upper Kerayan dialects and two Lengilo' dialects (B. Liku and Binuang). Phonological changes concerning this infix are discussed below in §4.1.

Table 13: Six Kerayan verbs marked for undergoer voice and perfective aspect

Dialect		wash	fan	tear	cover	take	do/make
Lun Dayeh		<i>biru?</i> #	<i>bib^her</i>	<i>diraak</i>	<i>kikeb</i>	<i>ilap</i> ** <i>inalap</i>	<i>tinaru?</i>
Kelabit		<i>binuru?</i>	<i>bibher</i>	<i>riraak</i>	<i>kikeb</i>	<i>alap</i>	<i>sinaru?</i>
Lengilo'	Lg Mutan	<i>biro?</i>	<i>bifer</i>	<i>diraak</i>	<i>ikeb</i>	<i>ilap</i>	<i>sinaru?</i>
	Binuang	<i>yuro?</i>	<i>ifel</i>	<i>iraak</i>	<i>ikeb</i>	<i>yalap</i>	<i>yaro?</i>
	B. Liku	<i>yuro?</i>	–	<i>iraak</i>	<i>ikeb</i>	<i>yalap</i>	<i>yaro?</i>
Lun Baa'	P. Lutut	<i>yuru?</i>	<i>ifer</i>	<i>diraak</i>	<i>ikeb</i>	<i>yalap</i>	<i>yaro?</i>
	Sembudud	<i>yuro?</i>	<i>ifel</i>	<i>iraak</i>	<i>ikeb</i>	<i>yalap</i>	<i>yaro?</i>
Central	P.Padi	<i>yuro?</i>	<i>ifel</i>	<i>iraak</i>	<i>ikeb</i>	<i>yalap</i>	<i>yaro?</i>
	P. Kurid	<i>yuro?</i>	<i>ifel</i>	<i>iraak</i>	<i>ikeb</i>	<i>yalap</i>	<i>yaro?</i>
Upper Kerayan	P.Tera	<i>yuro?</i>	<i>ifel</i>	<i>iraak</i>	<i>ikeb</i>	<i>yalap</i>	<i>yaro?</i>
	P. Kaber	<i>iyo?</i>	<i>ipel</i>	<i>iyaak</i>	<i>ikeb</i>	<i>yalaak</i>	<i>ayo?</i>
	Sa'ban	<i>iru?</i>	<i>ipel</i>	<i>iraak</i>	<i>ikap</i>	<i>alaak</i>	<i>aro?</i>

The majority of the Kerayan dialects reflect a verb root with [u] as the penultimate vowel i.e. *buru?*; however, Lun Dayeh, Lg Mutan, Lg Padi and P.Keber all have a schwa as the penultimate vowel in the actor voice imperfective form of the verb, which implies a root *beru?* (see Table 8).

** The two forms, *ilap* and *inalap*, suggest that there may be a variant form of the root with an initial schwa instead of an initial [a].

Undergoer voice in Karayan dialects has more in common with the Sa'ban system than with the Lun Dayeh system. The undergoer voice suffix *-en* of Lun Dayeh has been all but lost in Kerayan dialects, but surviving fossilised remains on verbs such as *beré* 'give' (Table 11) show that it was in use in earlier versions of the dialects. Morphological marking of the undergoer voice (imperfective) form of the verb is replaced, as in Sa'ban, by a periphrastic construction with the verb (*t)aru?* 'do/make'. In perfective aspect an undergoer voice verb continues to be marked morphologically,⁵¹ but the phonological changes which are taking place in Kerayan dialects are resulting in the former infix becoming a prefix.

⁵¹ The auxiliary *pangeh* 'finished' also occurs with verbs marked for perfective aspect in undergoer voice, but it never replaces the affix *-in-/i-* as it does the perfective prefix *ne-* in actor voice in Sa'ban and some Kerayan dialects.

3.1.2.3 Other verb affixes in Kerayan dialects

No other verbal affixes were elicited during this survey. An abortive attempt was made to collect data on stative verbs for comparison with the Lun Dayeh system, but it appears that Kerayan dialects do not have a stative system comparable to that described above for Lun Dayeh.

3.2 Pronoun Sets in Kerayan dialects

Grammatically, there is only one set of pronouns in Kerayan dialects (set I), but this one set can be divided into two groups according to the dialectal form of the pronouns. Pronouns of the first or *wih* group occur in all Lun Baa' dialects, all central dialects, in two Lengilo' dialects (Binuang and B. Liku), and in P. Sing, and P. Tera in the upper Kerayan. Pronouns of the second or *éek* group occur in two Lengilo' dialects (Lg Padi and Lg Mutan), and in three dialects from the upper Kerayan (Lg P. Sia, P. Kaber, Lg Rungan). There is some dialectal variation within each of these two groups of pronouns, so diaforms are used to list them in Table 14.

Table 14: Kerayan pronouns, Set I (diaforms)

	<i>wih</i> dialects	<i>éek</i> dialects
1S	<i>wih</i>	<i>iek/éek</i>
2S	<i>(i)ko</i>	<i>(i)keh</i>
3S	<i>ieh</i>	<i>ieh</i>
1IP	<i>tau</i>	<i>(ki)tam</i>
1EP	<i>kai</i>	<i>kami/amay</i>
2P	<i>(m)uyuh</i>	<i>mé</i>
3P	<i>(i)deh</i>	<i>deh</i>

Pronouns of set I in Kerayan dialects perform the functions of pronouns of set I and II in Lun Dayeh; that is, they function as the actor and undergoer in both actor and undergoer voice clauses, and following a noun, they signal the possessor. Their function in a clause is made clear by their constituent order and, to a lesser extent, by the affix on the verb. Pronouns functioning as actors in actor voice clauses are illustrated in examples (63–66), pronouns functioning as actors in undergoer clauses are illustrated in examples (68–72). Pronouns functioning as undergoers in actor voice clauses are illustrated in example (78) from Sembudud (Lun Baa') and (79) from P. Kaber (Ulu Kerayan).⁵²

(78) *Wih mier ieh.*
 (PIV)1S AV:see 3S
 'I see him.'

(79) *Éek menal ieh.*
 (PIV)1S AV:see 3S
 'I see him.'

⁵² Due to an oversight, the only elicited examples of undergoers in undergoer voice clauses in other Kerayan dialects are of nouns, not pronouns.

Pronouns similar to pronouns of set II in Lun Dayeh were noted in two Kerayan dialects (Lg Padi and Lg Mutan) and in Kelabit,⁵³ but they had no separate function, and in every case could be replaced by a pronoun from the Kerayan set I. Examples (80–82) illustrate this ambivalence from the dialect of Lg Mutan. Note in (81) that in Lun Dayeh a set II pronoun could not occur in this position; it would have to be a set III pronoun (*nemu*) as illustrated in example (24).

- (80) *rumaʔ kuh (or iek)*
house 1S (1S)
'my house'
- (81) *Iek nier mu (or ikeh).*
(PIV)1S AV:see 2S (2S)
'I see you.'
- (82) *Lawid neh ilap iek (or kuh).*
(PIV)fish that UV.PF:take 1S (1S)
'I caught that fish.'

In Lun Dayeh, pronouns of set III marked by the proclitic *ne=/n=* function as undergoers in actor voice clauses, and as obliques, but in Sa'ban, pronouns marked by *ne=/n=* only indicate obliques. In many Kerayan dialects oblique pronouns were marked by *ki=* (or *ke=/ko=*) in constructions parallel to those marked by *ne=/n=* in Lun Dayeh, as shown in examples (64b, 69–71, 76).⁵⁴ Dialects of the upper Kerayan marked obliques by a preposition preceding a pronoun of set I, as in examples (72 and 75) from P. Sing and P. Kurid respectively, but both systems are in use in dialects such as Sembudud, T. Karya and P. Padi.

Bario Kelabit has a set of genitive pronouns similar to those in Lun Dayeh (set II) but, as in the Kerayan dialects, their use seems to be on the wane, and in the examples elicited in 1995, pronouns of set I were also used to signal possession (83), and non-pivot actors (84). The undergoer in actor focus clauses, which is marked by set III pronouns in Lun Dayeh, is simply marked by a set I pronoun in Kelabit (85).

- (83) *rumaʔ iko (or mu di)*
house 2S (or 2S PT)
'your house'
- (84) *Sapaʔ sineh biré uih ngen neh.*
(PIV)shirt this UV.<PF>:give 1S to 3S
'I gave this shirt to him.'
- (85) *Uih nier ieh.*
(PIV)1S AV:see 3S
'I see him.'

⁵³ They were also noted in Lg P. Sia, but the dialect of that language helper was suspiciously close to Lun Dayeh although he used pronouns of the *éek* set.

⁵⁴ Dialects with this proclitic were: Lun Baa' (Sembudud, Lembudud, T. Karya, P. Mering); Lengilo' (Binuang, B. Liku); Central (P. Padi), upper Kerayan (P. Tera).

Pronouns of the *wih* set are phonetically similar to Lun Dayeh pronouns. In turn the Lun Dayeh pronouns have phonetic similarities with pronouns belonging to set I (pivot set) in languages of Sabah. The first person plural forms *tau* and *kai* are phonetically similar to *takau* and *akai* in Timugon, *tokou* and *okoi* in Kimaragang, and *toko* and *kai* in Tombonuo (Clayre 1996:56).⁵⁵ The second person plural pronoun which occurs in the *wih* dialects, *(m)uyuh*, has phonetic similarities only with languages in Sabah, where *muyu* occurs in the nominative and genitive pronoun sets of Ida'an, and *muyun* in the accusative set of the same language; *muyu* and *duyuh* occur in the genitive pronoun sets of Tombonuo and Kimaragang respectively, and *ramuyun* occurs among the non-pivot (non-focus) non-actor set of Timugon (Moody 1991; Clayre 1996:57).

Dialects with the *éek* set of pronouns include two Lengilo' dialects (Lg Padi and Lg Mutan) and dialects of the upper Kerayan, excluding P. Sing and P. Tera. In the *éek* dialects the first person singular pronoun is realised as *iek* in the dialects of Lg Padi, Lg Mutan and Lg P. Sia, and as *éek* in Lg Rungan, P. Kaber, and Sa'ban. Some older Sa'ban recall that their grandparents said *ak*, and this certainly links Sa'ban to the many Bornean languages which have some form of *aku*, but not with the Lun Dayeh and Kerayan dialects which have *wih*. The first person plural pronouns, *(ki)tam* and *(k)ami* occur as *kitam* and *kami* in the Lengilo' dialects of Lg Padi and Lg Mutan, and as *taam* and *amay* in the Lg Rungan dialect and in Sa'ban. These pronouns are quite distinct from the first person plural pronouns in the *wih* set. The links of *(ki)tam* and *(k)ami* are with the Kayan, Murik and Penan languages located to the south of the Kerayan in Sarawak and Kalimantan (Clayre 1996:57). The pivot first person plural pronouns in Kayan are *itam* or *tam* and *kamé?*, in Penan *itam* or *tam* and *amé* or *mé* and in Murik *itam* and *kami?* The second person plural pronoun of the *éek* dialects of Lg Mutan, Lg Padi and Lg P. Sia, is *mé*, which has no immediately obvious similar form in neighbouring languages. In the *éek* dialects of Lg Rungan and Sa'ban this pronoun occurs as *ciem*. Blust (1999) in a paper on Sa'ban suggested that this pronoun, together with the first person singular, *éek*, the second person singular, *ceh*, the first person plural inclusive, *taam*, and some dual and paucal forms of Sa'ban pronouns, may have a Kayan source. The source may have been a language closely related to Kayan, Murik or Ngurik, as the Murik once lived alongside the Sa'ban in the Bahau, before migrating to Sarawak in the nineteenth century.

The second person singular pronoun has a back vowel [ɔ] in *wih* dialects and a central vowel [ə] in *éek* dialects. It is phonetically similar to *ko(h)* or *ikaw* in Tombonuo and Kimaragang Dusun, and *kou* in Timugon, but also to the second person singular forms, *ika* and *iko*, which occur in Kayan and Kenyah dialects (Clayre 1996:56–57).

In the Bario dialect of Kelabit in Sarawak, the pronouns *uih*, *iko*, *ieh*, *tau* and *muyuh* are typical of the *wih* group, but the first person plural exclusive pronoun is *kami*, which is typical of the *éek* group.

3.3 Noun markers

Common nouns have no noun markers in Kerayan languages. Markers for proper nouns were not researched in this survey, but one example of the use of a personal noun marker

⁵⁵ Sellato (pers. comm.) informs me that *tau* or *to* and *kai* occur in Aoheng and other languages of Kalimantan.

to signal an oblique argument occurred spontaneously in a B. Liku example (86). Oblique pronouns in the B. Liku dialect are marked by the proclitic, *ke=*.⁵⁶

- (86) *So meré bakad nih ki=Samwel.*
 (PIV):2S AV:give shirt this PM.NPNA=Samuel
 'You give this shirt to Samuel.'

3.4 Word order

Given that all the Kerayan data were elicited, and that the responses may have been influenced by the Indonesian or Lun Dayeh languages used in the interviews, it would be unwise to attempt to draw firm conclusions concerning the importance of word order in Kerayan dialects. It would appear, however, that in actor voice clauses, the preferred order is SVO. One consistent pattern that does emerge is that the non-pivot core argument always occurs immediately following the verb.

4 Phonological processes observed in Kerayan dialects

It will already have become apparent that, in addition to the grammatical changes that are happening in Kerayan dialects, many phonological changes are also taking place. Analysis of these changes is at a very preliminary stage,⁵⁷ and attention is drawn here to only a few cases relating to the verbs and pronouns discussed earlier.

4.1 The infix *-in-*, *-i-* ablaut, and consonant deletion

The evidence from Kerayan verbs suggests that the processes taking place in verbs with the infix *-in-*, such as *binuru?* (Kelabit) 'washed', *tinaru?* (Lun Dayeh) 'done/made' and *inalap* (Lun Dayeh) 'taken' (Table 13), are as follows: the first stage is the deletion of *n* from the affix to produce *biuru?*, *tiaru?* and *ialap*, the forms *biuru?* and *ialap* were not recorded, but the form *tiaru?* occurs in the Lg Padi dialect of Lengilo'; the next stage is the deletion of the root-initial consonant which would produce *iuru?* and *iaru?*, in which, as in *ialap*, the former *-in-* infix looks like a prefix *i-*; the final stage involves reinterpretation with the loss of syllabicity of initial *i* before vowels to produce *yuru?*, *yaru?* and *yalap*,⁵⁸ and these forms occur in Lun Baa', Lengilo' and central dialects. In the upper Kerayan dialects of P. Sing and Sa'ban, even the *y-* prefix form has been lost and the verb occurs as an unaffixed root, *alaak* and *aro?* (Table 13). In P. Kaber while *yalaak* retains the *y-*, *aʔo?* ('done') has lost any marking. These processes may be diagrammed as in Table 15. Another change taking place in most Kerayan dialects is the lowering of a final high back vowel to a mid back vowel (Table 13).

⁵⁶ See fn. 54.

⁵⁷ I am grateful to Carolyn Rensch of SIL for her comments during preliminary discussions about the phonological changes taking place in the Kerayan.

⁵⁸ A similar loss of syllabicity occurs with *u* in word-final position, for example *uih* '1st person sg.' in Lun Dayeh, is realised as *wih* in Kerayan dialects.

Table 15: Phonological processes associated with the *-in-* infix in Kerayan verbs

Forms in bold occur in at least one dialect

Root	addition of <i>-in-</i>	<i>n-</i> deletion	initial C deletion	<i>i-</i> realignment
<i>buru?</i> 'wash'	<i>binuru?</i>	> <i>biuru?</i>	> <i>iuru?</i>	> <i>yuru?</i>
<i>taru?</i> 'make'	<i>tinaru?</i>	> <i>tiaru?</i>	> <i>iaru</i>	> <i>yaru?</i>
<i>alap</i> 'take'	<i>inalap</i>	> <i>ialap</i>		> <i>yalap</i>

In the case of verbs with the *-i-* ablaut form, such as *biru?*, *bib^her*, *diraak*, and *kikeb* (Table 13), the deletion of the consonant /n/ from *-in-* has already taken place. These verbs also show loss of the initial consonant of the root. This process, however, is not yet complete in all Lun Baa' and Lengilo' dialects since P. Lutut and Lg Mutan both retain some initial consonants, but it is complete in dialects of the central area and upper Kerayan (Table 13). Where this process has taken place the *i-* retains its syllabicity since it precedes a consonant, and the end result is that the perfective infix effectively becomes a prefix. The processes may be diagrammed as in Table 16.

Table 16: Phonological processes associated with the *-i-* ablaut in Kerayan verbs

Root	Schwa to <i>-i-</i> ablaut	Initial C deletion
<i>kekeb</i> 'cover'	<i>kikeb</i>	<i>ikeb</i>
<i>deraak</i> 'tear'	<i>diraak</i>	<i>iraak</i>
<i>beré</i> 'give'	<i>biré</i>	<i>iré</i>

4.2 Palatalisation of /k/, and the case of the second person singular pronoun

Within the Kerayan area the second person singular pronoun has undergone two noticeable sound changes. These are the loss of the initial vowel, and palatalisation. These two processes are illustrated by the dialectal realisations of the second person singular pronouns listed in Table 17. An additional change is seen in the *éek* dialects where the final vowel is realised as a schwa followed by a word-final h.

Table 17: Realisations of the second person singular pronoun *iko* in Kerayan dialects

	Realisation	Dialect
<i>Wih</i> dialects	<i>iko</i>	Lg Nuat, Lg Midang, P. Lutut,
	<i>ko</i>	T. Karya, P. Mering, Sembudud
	<i>k'ò</i>	P. Padi
	<i>ico</i>	Lg Kabid
	<i>co</i>	P. Kurid
	<i>so</i>	B. Liku, Binuang
<i>Éek</i> dialects	<i>s'ò</i>	P. Tera
	<i>ikeh</i>	Lg Mutan, Lg Padi, Lg P. Sia
	<i>keh</i>	Lg Rungan, P. Keber
	<i>ceh</i>	Sa'ban

These examples show /i/ causing palatalisation of a following /k/ and, then, in most dialects being deleted, with its presence being perpetuated by palatalisation.

4.3 The realisation of /p/ and /b^h/ as [f]

In some dialects of Lun Dayeh, the phoneme /p/ occurs in all positions of the word; in other dialects (including the dialect of Lun Dayeh described in this paper) /p/ has a labiodental allophone [f] that occurs in word-initial and word-medial positions. For example, some dialects say [purut] 'dowry' and [tupəd] 'stand' and others say [furut] and [tufəd]. However, [p] occurs in word-final position in all dialects, for example /ngalap/ [ŋalap] 'to take'.⁵⁹ In Kerayan dialects word-medial (intervocalic) /p/ is realised as a voiceless plosive, as for example in *laput* 'cloud', *lipen* 'tooth', *tuped* 'stand'. In word-initial position in most Kerayan dialects /p/ is realised as a voiceless plosive, but in Lun Baa' dialects it is realised as a labiodental fricative [f]. For example /pengeh/ or [fəŋəh] 'finished', /palad/ or [falad] 'palm'(of hand) and /pung/ or [fuŋ] 'wild animal'.

A labiodental fricative that may not be an allophone of /p/⁶⁰ does, however, occur in most dialects of Kerayan as the reflex of what Blust calls the voiced aspirate, */b^h/, of PKLD (Blust 1993:146–148). In Kelabitic languages this phoneme is reflected in three ways: in Kelabit and most dialects of Lun Dayeh its reflex is a voiced aspirate; in three dialects (Lg Sing, P. Keber and Sa'ban) its reflex is /p/; but in most Kerayan dialects its reflex is a voiceless labiodental fricative [f], with an optional bi-labial variant [ɸ].⁶¹ In Table 18, the reflexes of PKLD */b^h/ in Kerayan dialects are illustrated and contrasted with the occurrence of */p/ and */b/.

Table 18: Examples from Kerayan dialects of the reflexes of PKLD */b^h/, */b/ and */p/

Dialect		smoke	to fan	water	hair	tooth	ashes
PKLD		*reb ^h un	*beb ^h er(n)	*eb ^h aq	*eb ^h uk	*lipen	*abuh
Kelabit		reb ^h un	meb ^h er	eb ^h a?	eb ^h uk	lipen	abuh
Lun Baa':	P. Lutut	refun	mefer	fa?	fuk	lipen	abuh
	Sembudud	refun	mefer	fa??	fuk	lipen	abuh
	T. Karya	refun	mefer	fa?	fok	lipen	abuh
Lengilo':	Lg Padi	refun	ngefer	fa?	fuk	lipen	abuh
	B. Liku	defun	ngefel	fě?	fok	–	abuh
Central:	Lg Kabid	fun	mefel	fě?	fok	lipen	abuh
	P. Padi	fun	fel	fě?	fuk	–	abuh
Upper Keryan:	P. Tera	fun	ngefel	fě?	fok	–	abuh
	Lg Sing	depun	mpel	pé?	pok	ipen	abuh
	P. Kaber	lepun	mepel	pei?	pok	ipen	abeu
	Sa'ban	pu.n	mpel	pei?	puek	ipan	abeu

Note that the symbol é represents a front open-mid vowel, and e represents the schwa.

*beb^her is a nominal form, but the Kerayan examples of this word are verb forms

⁵⁹ In the '[f] dialects' when a suffix is added to a verb ending with /p/, the /p/ then becomes intervocalic and is realised as [f]. For example, /ngalap/ 'to take' > [lafen] 'to be taken' (UV.IMPF) and [lafu] 'take!' (UV.IMP).

⁶⁰ No minimal pairs occur in my limited data, but in several dialects such as P. Padi (central), Lg Mutan (Lengilo') and P. Tera (upper Kerayan), there are close contrasts between [p] and [f] in words such as [fun] 'smoke' and [puŋ] 'wild animal', or [fok] 'hair (head)' and [poʔon] 'beginning' or 'trunk of tree'.

⁶¹ Blust (1993:147) lists other dialects and languages in Sarawak with the reflexes, p and f.

In addition to illustrating that [f] is a reflex of PKLD */b^h/ the table also illustrates other sound changes that have been taking place in the Kelabitic dialects of the Kerayan. For example, the loss of initial vowels, illustrated by *eb^haʔ and *eb^huk, the loss of initial consonants as illustrated by *lipen, the loss of initial syllables as illustrated by *reb^hun, and the replacement of final /ɾ/ by /l/ as illustrated in *meb^her.

5 Final comments

This survey based on limited data collected over a very short period nevertheless highlights phonological and morphological changes that are taking place over a small geographical area.

Lun Dayeh, in the north of the area and adjacent to Sabah, is still recognisable as a Philippine type ‘focus’ language with three voices, appropriate verb affixation and separate pronoun sets. Differences between the pronoun sets are, however, evident only in the singular pronouns, the plural pronouns having already been neutralised.

The range of voice types in all other Kerayan dialects is reduced to two: actor and undergoer. Verb affixation and pronoun sets are also reduced. In most dialects pronoun sets are reduced to one set that fulfils all the functions of the three sets in Lun Dayeh. Set III pronouns in Lun Dayeh were marked by a proclitic, but similar proclitic marking is found in only seven Kerayan dialects (Sembudud, Lembudud, T. Karya, B. Liku, Binuang, P. Padi and P. Tera); all other dialects signalled an oblique pronoun by a preposition.

Verb affixes marking the semantic role of the pivot are reduced to a nasal prefix (N-) signalling actor voice, and a form of *-in-* signalling perfective aspect in undergoer voice. All verb suffixes have been lost, including the undergoer voice (imperfective aspect) suffix *-en*. The latter has been replaced by a periphrastic construction but fossilised remnants occurring with a small group of conservative verbs indicate that undergoer voice was morphologically marked in earlier versions of the dialects. The actor voice perfective prefix *ne-* is in the process of being replaced by an auxiliary verb, *pengeh*. The actor voice prefix (N-) itself seems to be under threat in many dialects, such as Sa’ban, where some verbs occur without a prefix, for example *alaak* ‘take’ or *abeh* ‘carry on the back’, and P. Padi where most actor voice verbs had no prefix.

To compensate for the loss of information carried by the pronoun sets and verb affixes, constituent order becomes more important. In Lun Dayeh the verb regularly occurs in clause-initial position, but in Kerayan languages more often the pivot occurs in this position. In Lun Dayeh and all the Kerayan languages, it is the rule that the non-pivot argument always follows immediately after the verb.

The introduction of periphrastic constructions, verb auxiliaries and prepositions indicate that Kerayan dialects are moving away from a morphologically marked language structure to a more isolating kind of language structure.

The survey also shows that many of the phonological and morphological changes that have been noted in Sa’ban are part of a wider picture of language change affecting the whole of the Kerayan region.

Appendix 1: List of language helpers

Dialect	Name	Age	Sex	Interview location	Origin of helper
Lun Dayeh	Andreas Padan	40s	M	Tarakan	Lg Umang
Lun Dayeh	Hendrik Udan	40s	M	Lg Bawan	Lg Sepayang
Lun Dayeh	Tek ^h u Tebari?	70s	M	Kampung Baru	Lg Nuat
Lun Dayeh	Guk ^h ang Tebari?	70s	M	Kampung Baru	Lg Nuat
Lun Dayeh	Hendrik Tadem	50s	M	Lg Bawan	Lg Nuat
Lun Baa'	Mitun Tilo	50s	M	Lg Midang	Lg Midang
Lun Baa'	Isaak Surang	40s	M	T.Karya	T. Karya
Lun Baa'	Markus Petrus	20s	M	Samarinda	T. Karya
Lun Baa'	Sigar Lau	70s	M	Lembudud	Lembudud
Lun Baa'	Elisa Piuk	50s	M	Lg Bawan	Sembudud
Lun Baa'	Fatima Kayeh	30s	F	Lg Bawan	P. Lutut
Lun Baa'	Aurin Yusup	20s	F	Pulau Sapi	P. Mering
Central	Ayub Sigar	40s	M	Lg Bawan	Lg Kabid
Central	Ganit Dolof	40s	F	Lg Bawan	P. Kurid
Central	Moses Dawat	40s	M	Lg Bawan	P. Padi
Lengilo'	Dorce Liu	30s	F	Lg Bawan	Lg Padi
Lengilo'	Nikelas Pengiran	30s	M	Samarinda	Lg Padi
Lengilo'	Juliana Martin	40s	F	Lg Bawan	B. Liku
Lengilo'	Samuel S T Padan	30s	M	Lg Bawan	Binuang
Lengilo'	Simsun Duyung	40s	M	T. Lapang	Lg Mutan
U Kerayan	Ediento Toran	20s	M	Lg Bawan	P. Ibang
U. Kerayan	Sakai Riang	30s	M	Lg Bawan	Lg P. Sia
U. Kerayan	Armi Yudan	30s	M	Lg Bawan	P. Tera
U. Kerayan	Pelipa Abdulah	18	F	Lg Bawan	P. Sing
U. Kerayan	Martinus Jalung	30s	M	Lg Bawan	P. Kaber
U. Kerayan	Ediento Toran	20s	M	Lg Bawan	P. Ibang
U. Kerayan	Darius Kamis	20s	M	Samarinda	Lg Rungan
Sa'ban	Franky Aran	50s	M	Lg Bawan	Lg Tua
Sa'ban	Arsufandi Matius	20s	M	Samarinda	Lg Tua
Kelabit	John Terawe	40s	M	Miri	Bario

References

- Amster, Matthew H., 1995, *Kelabit/English, English/Kelabit glossary*. Kuching: Rurum Kelabit Sarawak.
- Blust, Robert, 1993, Kelabit–English vocabulary. *Sarawak Museum Journal* 65:141–226.
- 1997, Ablaut in northwest Borneo. *Diachronica* 14:1–30.
- 1999, Language, dialect and riotous sound change: the case of Sa’ban. In Graham W. Thurgood, ed. *Papers from the Ninth Annual Meeting of the Southeast Asian Linguistics Society*. Arizona State University, Program for Southeast Asian Studies.
- Bolang, A. and Tom Harrisson, 1949, Murut and related vocabularies. *Sarawak Museum Journal*, 5(1) (NS):116–124.
- Boutin, Michael E., 2002, Nominative and genitive case alternations in Bonggi. In Wouk and Ross 2002:209–239.
- Cense, A.A. and E.M. Uhlenbeck, 1958, *Critical survey of studies on the languages of Borneo*. ‘S-gravenhage: Het Koninklijk Instituut voor Taal-, Land- en Volkenkunde.
- Clayre, Beatrice, 1972, A preliminary comparative study of the Lun Bawang (Murut) and Sa’ban languages of Sarawak. *Sarawak Museum Journal* 20:145–171.
- 1994, Verb affixes in Berawan. In Cecilia Odé and Wim Stokhof, eds *Proceedings of the Seventh International Conference on Austronesian Linguistics*, Leiden 22–27 August 1994. Leiden University: Department of Languages and Cultures of South East Asia and Oceania, Projects Division.
- 1995, A preliminary typology of the languages of middle Borneo. Paper presented at the fourth biennial conference of the Borneo Research Council, Brunei.
- 1996, The changing face of focus in the languages of Borneo. In Steinhauer 1996:51–88.
- 2002, Observations on Lundayeh auxiliaries and the case of aru’. In K. Alexander Adelaar and Robert Blust, eds *Between worlds: linguistic papers in memory of David John Prentice*, 49–63. Canberra: Pacific Linguistics.
- in preparation, The phonology of Sa’ban.
- Crain, Jay B., 1978, The Lun Dayeh. In Victor T. King, ed. *Essays on Borneo societies*, 123–142. Hull Monographs on South-east Asia, Oxford: Oxford University Press.
- Douglas, Reginald S., 1911, A comparative vocabulary of the Kayan, Kenyah and Kelabit languages. *Sarawak Museum Journal* 1(1):75–119.
- Galvin, A. Dennis, 1967, [Kenyah vocabulary], unpublished MS.
- Himmelman, Nikolaus P., 2002, Voice in western Austronesian: an update. In Wouk and Ross 2002:7–16.
- Hudson, Alfred B., 1978, Linguistic relations among Borneo peoples with special reference to Sarawak: an interim report. In Mario Zamora, Vinson Sutlive and Nathan Altshuler, eds *Sarawak: Linguistics and development problems*, 1–44. Williamsburg: Studies in Third World Societies 3, Department of Anthropology, College of William and Mary.
- Kroeger, Paul R., 1988, Verbal focus in Kimaragang. In Steinhauer 1996:217–240.
- 1990, Stative aspect and unaccusativity in Kimaragang Dusun. *Oceanic Linguistics* 29(2):110–131

- Langub, Jayl, 1987, Ethnic self-labelling of the Murut or Lun Bawang of Sarawak. *Sojourn* 2(2):289–299. Singapore: Institute of South-east Asian Studies.
- Lees, Shirley P., 1959, Lun Daye Phonemics. *Sarawak Museum Journal* 13–14:56–62.
- Moody, David C., 1991, Continuity and development in Ida'an narrative discourse. In Stephen H. Levinsohn, ed. *Thematic continuity and development in languages of Sabah*, 137–162. Canberra: Pacific Linguistics.
- Ray, Sidney H., 1913, The languages of Borneo. *Sarawak Museum Journal* 1(4):1–196.
- Roth, Henry Ling, 1896, *The natives of Sarawak and British North Borneo*. 2 vols. London: Truslove and Hanson.
- Schneeberger, Werner F., 1945, The Kerayan-Kalabit highland of central northeast Borneo. *The Geographical Review* 35:544–562.
- 1979, Contributions to the ethnology of central northeast Borneo. *Studia Ethnologica Bernesia* 2. Institute of Ethnology, University of Berne.
- Sellato, Bernard, 1997, Agricultural practices, social organization, settlement patterns, and ethnogenetic processes in East Kalimantan. In K.W. Sørensen and B. Morris, eds *People and plants of Kayan Mentarang*, 27–57. London: World Wide Fund for Nature and UNESCO.
- n.d., Kerayan Settlements. Unpublished MS.
- Southwell, C. Hudson, 1949, Structure of the Murut language. *Sarawak Museum Journal*, 5(1)(NS):104–115.
- 1990, *Kayan-English dictionary* (second edition). Kuching: Sarawak Literary Society.
- Steinhauer, Hein, ed., 1996, *Papers in Austronesian Linguistics* No.3. Canberra: Pacific Linguistics.
- Tuie, Meechang, 1995, *Masyarakat Lun Bawang Sarawak*. Kuching: Dewan Bahasa dan Pustaka.
- Wouk, Fay, 2002, Introduction. In Wouk and Ross 2002:1–4.
- Wouk, Fay and Malcom Ross, eds, 2002, *The history and typology of western Austronesian voice systems*. Canberra: Pacific Linguistics.

3 *The Palu'e passive: from pragmatic construction to grammatical device*

MARK DONOHUE

1 An alternation in Palu'e

Voice constructions are usually associated with changes in the pragmatic status of the arguments of a clause, and so bear a strong resemblance to topic constructions in terms of their information structuring effects and entailments. Importantly, however, a defining criterion of voice alternations is that they morphologically monitor the changing status of the arguments of the verb, in terms of their grammatical function identity. By contrast, topic constructions are held to not affect the grammatical status of the arguments, but to restructure their pragmatic status. I shall present data from Palu'e, an Austronesian language of Central Indonesia, showing evidence that the language that has (recently) begun grammaticalising a topic construction into one member of a grammatical voice system, providing possible insights into the origin of voice systems in pragmatic structuring devices, and the nature of voice systems as constructional oppositions.

Palu'e is an Austronesian language spoken on the island of Palu'e, just off the middle of the north coast of Flores, in southern Indonesia. As with other languages of central Flores it has little bound morphology, with only some aspectual and adverbial clitics joining the language's four genitive clitics as enclitics. There are some incipient proclitics, marking case and agreement for 1SG subject, but they do not concern us here. There are two possibilities for encoding bivalent clauses, which can be illustrated in the alternation between (1) and (2), showing clauses with AVP and PAV orders, respectively.¹ The translations offered are taken from informants' translations into Indonesian. I have called

¹ In addition to A, S and P, which are defined following Comrie (1978) as the most agent-like argument of a lexical predicate, the single argument of a monovalent verb, and the most patient-like argument of a lexical predicate respectively, the following abbreviations are used: 1, 2, 3 first, second and third person; COMP complement(iser); CORE core; EMPH emphasis; GEN genitive; LNKR linker; NOM nominative; PASS passive; PERF perfective; PL plural; PRED predicate; PREP preposition; R realis; RED reduplication; SG singular; v₁ active-like voice (A is subject); v₂ passive-like voice (P is subject); VP verb phrase. Thanks are due to my Palu'e-speaking friends on Batam, and to Kazuya Inagaki.

the different clause types ‘unmarked’ and ‘marked’, based mainly on their frequency of occurrence.²

Unmarked clause type in Palu’e

A V P

- (1) *Ia cube vavi vaʔa.*
 3SG shoot pig that
 ‘He shot that pig.’

Marked clause type in Palu’e

P A V

- (2) *Vavi vaʔa ia cube.*
 pig that 3SG shoot
 ‘That pig, he shot (it).’ OR: ‘That pig was shot by him.’³

There is no doubt that sentences such as (1) represent the unmarked, or ‘basic’, coding choice in Palu’e: this is the structure most frequently encountered in narrative of whatever genre, it is the form given in response to pragmatically neutral translation requests, and it is the form that is translated with unmarked (= active, non topicalised) clause structures in Indonesian. Our question concerns the best analysis of (2): is it better analysed as an instance of topicalisation, bearing a relationship to (1) similar to that which pertains between the first translation given for it, ‘That pig, he shot’, and the translation given for (1), or is it in fact an instantiation of a voice alternation, showing a relationship more similar to that between the second translation of (2) and the translation given for (1)?

Following a short survey of voice systems, and the passive in particular, I shall present various tests for the syntactic status of the arguments in AVP clauses such as (1) and PAV clauses such as (2), and based on this empirical investigation shall discuss the implications of the Palu’e data for our models of voice systems in general, and the historical development of the Austronesian voice system in particular.

2 Voice systems and some atypical passives

All languages utilise some form of diathesis, and often more than one; but they can be hard to tell apart. In this discussion the analysis of the diathesis is problematic, the clearest choices being between a passive(-like) analysis and a topicalisation analysis.

An alternation in diathesis may be grammaticalised, as in the use of a voice system, or more purely pragmatic, such as in the function of topicalisation, which is ‘overlaid’ as a separate pragmatic module on the grammatical structure without affecting, for instance, the

² Palu’e examples are presented in a phonemic transcription. This matches IPA norms, with the following exceptions: *b*, *c*, *j* represent [β, tʃ, dʒ], *mb* and *nd* (and *ŋg*, which does not appear in the data here) are prenasalised (marginal) phonemes, and the accent ´ marks regular bimoraicity for the vowel of a monosyllabic foot. Phonemic CC clusters are broken up with an epenthetic vowel. The hyphen - marks a clitic boundary; there are no affixes in Palu’e, so this distinction does not need to be maintained.

³ For these two sentences, the Indonesian forms given would be: (1) *Dia panah babi itu* (2) *Babi itu dia panah* ~ *Babi itu dipanah dia* ~ *Babi itu dipanahnya*.

assignment of subject and object functions. In many instances these two systems will overlap in a language: most languages with a passive voice, for instance, require the P to be topical, and code it as such. English is a language that does not have this requirement, but does have both a grammaticalised voice system and a productive system of pragmatic topichood.

English: active/passive alternation

- (3) *Cats always chase those rats in the afternoon.*
 (4) *Those rats always get chased by cats in the afternoon.*

English: topic alternation

- (5) *Those rats, cats always chase('em) in the afternoon.*

English: topicalised agent appearing with a passive voice in the same sentence

- (6) *By cats, those rats always get chased in the afternoon.*

The similarity of voice and topicality is often reflected in the morphosyntax of a language. A complement clause in (Singapore) Hokkien may appear without a complementiser if the main clause object is in the VP, as in (7), but requires a complementiser if the object is external to the VP, through either topicalisation or passivisation, seen in (8) and (9), respectively. Thus, regardless of the grammatical status of the VP-external argument, its absence from the VP serves to trigger the requirement for a complementiser, *khi*.

Hokkien: simple complement construction

- (7) *Mama* [VP *kio kinna* [COMP *cia?* *py*]].
 mother tell child eat rice
 'Mother told the child to eat the rice.'

Hokkien: main clause passivised

- (8) *Kinna* [VP [*ho mama*] *kio* [COMP **(khi) cia?* *py*]].
 child PASS mother tell COMP eat rice
 'The child was told by mother to eat the rice.'

Hokkien: main clause topicalised

- (9) *Hi?* *e kinna, mama* [VP *kio* ___ [COMP **(khi) cia?* *py*]].
 That LNKR child mother tell COMP eat rice
 'That child, mother told to eat the rice.'

Comparing the English and Hokkien passive constructions shown in the sentences above we can observe most of the cross-linguistic diversity that is associated with passives, and which is summarised in Table 1. Here the notation *m-X* is used to indicate that some extra of morphology is present on the X, be it inflectional, derivational, adpositional or case.

Table 1: The English and Hokkien passive constructions compared

	English		Hokkien	
	Active	Passive	Active	Passive
word order	A [VP V P]	P [VP V (A)]	A [VP V P]	P [VP A V]
verbal morphology	V	<i>m</i> -AUX <i>m</i> -V	V	V
nominal morphology (A)	NP	(<i>m</i> -NP)	NP	<i>m</i> -NP
Grammatical status of A	SUBJ	OBL	SUBJ	OBL
Grammatical status of P	OBJ	SUBJ	OBJ	SUBJ
Pragmatic status of A	(free)	(free)	(free)	not TOPIC
Pragmatic status of P	(free)	tends to TOPIC	(free)	(free)

In English, and indeed overwhelmingly frequently in languages with a voice construction (Siewierska 1984; Haspelmath 1990), there is marking on the verb, or at least on a verbal auxiliary or the verb phrase, to indicate the passive construction. Similarly the A, optional in the passive construction, is overtly marked usually in a way that is consistent with some sort of adjunct. In Hokkien, on the other hand, we find that the verb is unchanged morphologically from the form seen in the active, and that the only indicators of the passive are the preverbal position of the A (the normal position for adjunct) and the marker *ho* that appears with this NP. Another difference is related to the sole morphological exponence of the passive being the marking on the A: the A is obligatory in this clause, not optional, as in English.

Hokkien: active clause

- (10) *I* [VP *phah hi?* *e kau*].
 3SG hit that LNKR dog
 ‘He hit that dog.’

Hokkien: passive clause

- (11) *Hi?* *e kau* [VP [*ho i*] *phah*].
 that LNKR dog PASS 3SG hit
 ‘That dog was hit by him.’

Hokkien: agentless passive clause

- (12) * *Hi?* *e kau* [VP *phah*].
 that LNKR dog hit
 ‘That dog was hit.’

Unlike a language like English, in which the pragmatic and grammatical tiers are quite separate, allowing for a demoted agent to be topicalised, the passive agent cannot appear topicalised in Hokkien, a fact which sets it apart from other adjuncts.

Hokkien: main topicalised P

- (13) *Hi?* *e kau, i* [VP *phah*].
 that LNKR dog 3SG hit
 ‘That dog, he hit.’

Hokkien: topicalised agent in passive clause⁴

- (14) * [*Ho i*], *hi?* *e* *kau* [_{VP} *phah*].
 PASS 3SG that LNKR dog hit
 'By him, that dog was hit.'
- (15) * [*i*], *hi* *e* *kau* [_{VP} *ho phah*].
 3SG that LNKR dog PASS hit
 'Him, that dog was hit by.'

The differences between the passives in the two languages shown are quite considerable, but commonalities are also clear: the grammatical status of the arguments changes, crucially involving what seems to be a defining feature of voice systems generally, a change in the identity of the argument assigned subject status. The morphology required by the construction can typically range from the multiple highly explicit instances, as in English, to the single NP marker *ho* in Hokkien.

With this background sketch of passive variation (and it is just a sketch; more detailed accounts of the kind of variation encountered cross-linguistically can be found in, amongst others, Foley and Van Valin 1984, Klaiman 1991, Shibatani, ed. 1988, Siewierska 1984, and Van Valin and LaPolla 1997) we can proceed to a syntactically detailed discussion of the AVP:PAV alternation in Palu'e.

3 The Palu'e alternation: syntactic or pragmatic?

We can first examine whether the PAV clauses in Palu'e are in fact instances of simple topicalisation, rather than a distinct clause type, or whether they are simple sentences without any necessary topicalisation, but with another sort of clause-internal alternation.

We know there is a sentence-initial topic position from sentence pairs such as the following. In (16) the goal appears in the usual postverbal position, but in (17) it appears in a sentence-initial topic position. The usual prosodic correlates of topical status are found: an intonationally distinct contour on the topical phrase, the possibility of a pause following the topic, and (for core nominals) the possibility of appearing in the clause with a resumptive pronoun (this will be illustrated later).⁵

- (16) *Ia pana le nata-gu.*
 3SG go PREP village-1GEN
 'He went to our village.'
- (17) *Le nata-gu, ia pana.*
 PREP village-1GEN 3SG go
 'To our village, he went.'

⁴ There is another passive construction in Hokkien, involving a VP-initial passive marker *tio?*. This is a compulsorily agentless passive, allowing sentences such as *Hi? e kau tio? phah* 'That dog was hit', but not (for most speakers) **hi? e kau tio? i phah*.

⁵ Three morphemes, *le*, *lae* and *lau*, will all be glossed simply as PREP for 'preposition'. They are generic prepositions which vary, amongst other factors, in the relative elevation of the NP that they mark (*lae*: lower, *lau*: higher). These factors are not relevant to the discussion here.

It is possible to topicalise on any phrasal element in the clause. In (18) we can see an adaptation of (16) with a topicalised subject; (19) shows the use of a resumptive pronoun inside the clause.

- (18) *Ata laki vaʔa, pana le nata-gu.*
 person male that go PREP village-1GEN
 ‘That man_(,) went to our village.’
- (19) *Ata laki vaʔa, ia pana le nata-gu.*
 person male that 3SG go PREP village-1GEN
 ‘That man, he went to our village.’

We can determine that the topic must appear preceding the clause by examining the distribution of temporal expressions in the clause. Building on (16), examples (20) and (21) show the (myriad) possibilities for the temporal noun *vaicvi* ‘yesterday’.

- (20) *Ia pana le nata-gu vaicvi.*
 3SG go PREP village-1GEN yesterday
 ‘He went to our village yesterday.’
- (21) a. *Ia pana vaicvi le natagu.*
 b. *Ia vaicvi pana le natagu.*
 c. *Vaicvi ia pana le natagu.*

Simply put, the time expression can appear in any position in the clause, as shown in (22).

- (22) A time adverbial may occur anywhere in its clause, not intruding into NPs or PPs

When there is a topicalised element in the sentence we find a constraint on the possible positions for temporal adjuncts: a temporal adjunct may not appear preceding the topicalised phrase. Sentence (24) illustrates the inability of a time expression to precede a topicalised oblique, and in (26) we can see that a time expression cannot precede a topicalised subject (S).

- (23) *Le nata-gu, ia pana vaicvi.*
 PREP village-1GEN 3SG go yesterday
 ‘To our village, he went.’
- (24) a. *Le natagu, ia vaicvi pana.*
 b. *Le natagu, vaicvi ia pana.*
 c. * *vaicvi le natagu, ia pana*
- (25) *Ata laki vaʔa, pana le nata-gu vaicvi.*
 person male that go PREP village-1GEN yesterday
 ‘That man_(,) went to our village yesterday.’
- (26) a. *Ata laki vaʔa, (ia) vaicvi pana le natagu.*
 b. *Ata laki vaʔa, vaicvi (ia) pana le natagu.*
 c. * *Vaicvi ata laki vaʔa, (ia) pana le natagu*

The placement of time expressions clearly delimits the left edge of the clause, and the topic, which occurs sentence-initially preceding all other elements of the clause, can only be followed by a time expression, never preceded by one.

Turning to bivalent clauses, we find a very similar picture. The basic AVP sentence in (27) can be expanded by means of a time expression as shown in (28), with the temporal occurring in all positions, as described in (29). This unproblematically matches the description in (22).

- (27) *Ia cia kami.*
3SG look.for 1PL.EX
'He looked for us.'
- (28) *Ia cia kami vaicvi.*
3SG look.for 1PL.EX yesterday
'He looked for us yesterday.'
- (29) a. *Ia cia vaicvi kami.*
b. *Ia vaicvi cia kami.*
c. *Vaicvi ia cia kami.*

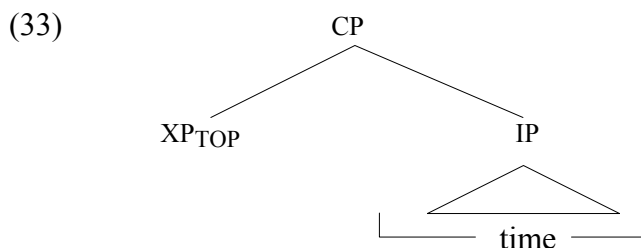
When we examine the PAV construction we find that there is no evidence to indicate that the sentence-initial NP in P role is a topic. No intonation break is required between this NP and the rest of the clause,⁶ and we find that temporal adjuncts may precede this NP. This is shown in (31) and (32). Note particularly that (32c), elaborating on the basic clause in (30), is grammatical. This contrasts strongly with the putatively analogous, but ungrammatical, sentence in (26)c.

- (30) *Kami ia cia.*
1PL.EX 3SG look.for
'He looked for us.' (OR: 'We were looked for by him.')
- (31) *Kami ia cia vaicvi.*
1PL.EX 3SG look.for yesterday
'He looked for us yesterday.'
- (32) a. *Kami ia vaicvi cia.*
b. *Kami vaicvi ia cia.*
c. *Vaicvi kami ia cia.*

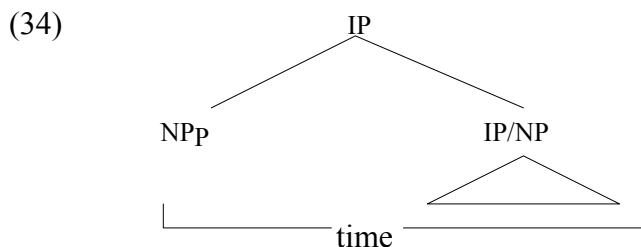
These facts suggest the following different structures, representing a topic structure in (33), and the PAV structure in (34).

⁶ Topicalisation of the P, with behaviour identical to other topics, including the intonation cues, is also possible, but is not the construction being discussed here.

Approximate representation of a topicalisation structure



Approximate representation of a PAV structure



It is clear, then, that the PAV clause in Palu'e does not involve the P appearing in a pre-clausal Topic position. Having established that the PAV construction is not simply an instance of topicalisation, we can now turn to tests that will elicit the syntactic status of the arguments in AVP and PAV clauses.

4 Testing the syntactic status of the alternation

In this section I shall present arguments that the alternation between AVP and PAV orders in Palu'e correlates with a change in the grammatical status of the arguments of the clause. The evidence used comes from three different constructions: the two floated quantifier constructions allow us to identify core arguments, and conjunction reduction and purposive subordination allow us to identify which core argument is the grammatically privileged subject.⁷

4.1 Argument/Adjunct status

The status of the PAV clauses as instances of topicalisation or passives can be partly decided by examining the status of the A. If the A is a core argument, then the topicalisation analysis is strongly supported, since that would indicate no 'demotion'. On the other hand syntactic evidence that the A is not a core argument would strongly support an analysis of this construction as involving a passive voice contrast with the AVP coding option.

Morphologically it is not clear that the A should be treated as an adjunct. It is a general characteristic of adjuncts in Palu'e that they are marked by a preposition (though the

⁷ Donohue (2003), and many others, shows that, in essence, not all constructions are equal for the purposes of determining subject. The constructions selected here appear to be adequately diagnostic for Palu'e.

converse proposition, that prepositions always mark adjuncts, is not true, as will be demonstrated below), and that they follow all subcategorised-for nominals in the clause. Some examples are shown in (35)–(37), illustrating a locative preposition and the instrumental/accompaniment preposition.

Clause with two core arguments and one adjunct (locative)

- (35) *Ia bere kaju lae uma.*
 3SG chop wood PREP garden
 'He chopped the wood in the garden.'

(35)' * *Ia bere kaju uma*

(35)'' * *Ia bere lae uma kaju*

Clause with an instrumental

- (36) *Ia bere kaju no?o tobo.*
 3SG chop wood PREP machete
 'He chopped the wood with a machete.'

(36)' * *Ia bere kaju tobo*

(36)'' * *Ia bere no?o tobo kaju*

Clause with a comitant

- (37) *Aku pana lau Todo no?o ina-gu.*
 1SG go PREP Todo PREP mother-1GEN
 'I went to Todo with my mother.'

(37)' * *Aku pana lau Tobo inagu*

(37)'' # *Aku pana no?o inagu lau Todo*

The presence of prepositional marking, in contrast to the bare NPs that are core arguments, might be thought to be a test for grammatical status. There are, however, some verbs that select for prepositionally marked postverbal NPs (thus, PPs), which are demonstrably not adjuncts, evidenced by their different behaviour when appearing preverbally. An example of this sort of prepositionally marked object in an unmarked postverbal position can be seen in (38). We can see from this example that there are some predicates that, when they take a nominal object, must have it marked with a preposition; *ɣaro* is one such predicate. The use of this preposition is obligatory with this predicate.

Object obligatorily marked by a preposition⁸

- (38) *Aku ɣaro no?o kau.*
 1SG love PREP 2SG
 'I love you.'

(39) * *Aku ɣaro kau*

⁸ The morphosemantic information encoded in the preposition *no?o* can also be coded by the obviously related, but proscribed, incipient case marker *no(?)*-; thus, in the immediately preceding sentences the object of hatred, *no?o ia*, will be realised prescriptively as [nɔʔɔ ia] ~ [nɔʔɔ ja], but in normal speech as [nɔʔia] or [nɔija], these last two variants showing degrees of cliticisation: *no-ia* ~ *noʔ-ia*, rather than being realised as two separate words, *no?o ia*. The syntactic behaviour of the two morphemes in the clause is in all cases identical, and for the sake of brevity only the full preposition *no?o* will be described here.

Other predicates, however, appear with objects that may be marked with a preposition, or with a bare NP. An example of this kind of predicate is *kau* ‘hate’.

Object optionally marked by a preposition

(40) *Aku kau noʔo ia.*
1SG hate PREP 3SG
‘I hate him.’

(41) *Aku kau ia.*
1SG hate 3SG
‘I hate him.’

By contrast, ‘normal’ bivalent predicates such as *cia* ‘look for’, seen earlier, do not allow for a prepositional option. While (27) is grammatical, prepositionally coded objects with this predicate are not: **ia cia (noʔo / le / lau / lae) kami*. When we compare the behaviour of the PP that normally follows a predicate such as *ɲaro* or *kau* with a PP in a clause such as those seen in (35)–(37), we find differences. Sentences with a topicalised adjunct appear with the PPs retaining their prepositions when appearing preverbally, as in (42) and (43), based on (37).

(42) *Noʔo ina-gu, aku pana lau Todo.*
PREP mother-1GEN 1SG go PREP Todo
‘With my mother, I went to Todo.’

(43) *Lau Todo, aku pana noʔo ina-gu.*
PREP Todo 1SG go PREP mother-1GEN
‘To Todo, I went with my mother.’

On the other hand, the preposition must be omitted in sentences based on clauses such as (38)–(41) with a topicalised object. This can be seen in (44)–(47), where only bare NPs are acceptable, regardless of whether or not the verb permits alternation in the appearance of the preposition or not.

(44) *Ia aku ɲaro.*
3SG 1SG love
‘I love him.’

(45) * *Noʔo ia aku ɲaro*

(46) *Ia aku kau.*
3SG 1SG hate
‘I hate him.’

(47) * *Noʔo ia aku kau*

We can see that some prepositionally marked NPs behave as do Ps in bivalent clauses when in a preverbal position, showing that they are not in fact adjuncts, but rather exceptionally case-marked arguments of the verb. This is analogous to English predicates such as ‘look at’, which take prepositionally marked objects (the ungrammaticality of **he looked it at* proves that a phrasal verb analysis is untenable for these predicates). Under passivisation the preposition is ‘left behind’, and does not occur with the new subject: *it was looked at*, **at it was looked*. Palu’e shows similar behaviour in the topicalisation structure shown in (46), but does not allow preposition stranding as English does. These

arguments show that morphological tests alone, such as the presence or absence of a preposition marking an NP, are not sufficient to judge the grammatical status of a participant. The floated quantifier constructions do, however, provide us with a syntactic test that can be appealed to in order to decide whether a nominal is argument or adjunct.

4.1.1 Simple floated quantifiers

The universal quantifier *tetiʔón* 'all' must appear in a clause-final position in Palu'e.⁹ When a monovalent clause appears with a clause-final quantifier, the quantifier can only be interpreted as being restricted to the S of the clause (regardless of the semantic nature of the S; this applies to all the tests illustrated here, though other tests, such as the possibilities available for adverbial clause marking, are sensitive to the unergative/unaccusative distinction).

- (48) *Aku ari-gu nodo tetiʔón.*
 1SG younger.sibling-1GEN sit all
 'All of my younger brothers and sisters are sitting down.'

Even when there is an adjunct closer to the quantifier than the subject, the quantifier cannot be interpreted as being restricted to the adjunct.

- (49) *Konen pana le nua vaʔa tetiʔón.*
 3PL go PREP village that all
 'All of them went to that/those village(s).'
 * 'They went to all of those villages.'

Floated quantifiers are also found with bivalent verbs; in this case the restriction of the quantifier is potentially ambiguous, as the quantifier can be interpreted as being restricted to either of the core arguments (but not both at the same time).

- (50) *Konen bere somu vaʔa tetiʔón.*
 3PL chop garlic that all
 'They chopped all of that garlic.' OR: 'All of them chopped that garlic.'

Notably, in bivalent clauses too the quantifier cannot be interpreted as being restricted to an adjunct nominal, even if it is 'closer' to the quantifier. Only the core arguments of the clause are eligible to control the NP-external quantifier.

- (51) *Konen bere somu noʔo kti vaʔa tetiʔón.*
 3PL chop garlic with knife that all
 'All of them chopped the garlic with those knives.' OR:
 'They chopped all of the garlic with those knives.'
 * 'They chopped the garlic with all of those knives.'

When we examine a PAV construction with a floated quantifier we find that a reading with the quantifier restricted to the A is not possible, as seen in (52).

⁹ It is clear that *konen* '3PL' and *tetiʔón* 'all' must, at least historically, be morphologically complex, involving the use of the third person genitive *-n*. Synchronically, however, there is no alternation and so these lexemes must be treated as unanalysable. Some of the data in this section has been presented as Donohue (2004c).

- (52) *Somu konen bere tetiʔón.*
 garlic 3PL chop all
 ‘They chopped all of the garlic.’ OR: ‘The garlic was all chopped by them.’
 * ‘All of them chopped the garlic.’

The data here indicate that the restriction of the floated quantifier is to non-adjunct arguments; as long as an argument is core, be it an A, S or P, it may be the restriction of the floated quantifier. This applies to monovalent clauses, and to bivalent AVP clauses. In PAV clauses, however, only the P may be the restriction of the quantifier. Either the restriction of this quantifier construction changes in the PAV clause type, or else the grammatical status of the arguments has changed, such that only the P ‘counts’ as a non-adjunct, while the A behaves as an adjunct.

4.1.2 Augmented floating quantification

Floated quantifiers may appear with other nominals, strikingly with non-core nominals being a possibility. In these cases the simple *tetiʔón* construction described in §4.1.1 is not used. Rather, the verb must have an extra cliticised unit, *naba*, and a nominal to which the quantifier is restricted must be reduplicated.¹⁰ If either the reduplication or the clitic *naba* are omitted, then the clause is ungrammatical; if both are omitted, then the only possible interpretation of the restriction of the quantifier is to a core argument of the clause, since this is then a case of simple, rather than augmented, floated quantification. These possibilities (and impossibilities) are shown in (53)–(56).

- (53) *Konen(-konen) vaʔa pana-naba le nata-nata tetiʔón.*
 3PL-RED that go-all PREP village-RED all
 ‘They went to all the villages.’
 * ‘All of them went to the villages.’
- (54) * *Konen vaʔa pana naba le nata tetiʔón.*
- (55) * *Konen vaʔa pana le nata-nata tetiʔón.*
- (56) *Konen vaʔa pana le nata tetiʔón.*
 3PL that go PREP village all
 ‘They all went to the villages.’
 * ‘They went to all of the villages.’

The sentences in (53)–(56) show the behaviour of the quantifier in monovalent clauses. It is not the case that the *-naba* (RED-) construction is used only to express quantification of adjunct participants. This quantifier construction is also found with bivalent clauses without adjuncts, in which case the floated quantifier is unambiguously restricted to the P, not the A. This can be seen in (57).

- (57) *Konen vaʔa ka-naba keʔo-keʔo tetiʔón.*
 3PL that eat-all corn-RED all
 ‘They ate all the corn.’
 * ‘All of them ate the corn.’

¹⁰ Reduplication is an option that is available for indicating plurality of nouns, regardless of the present of quantifiers in the clause. Plural pronouns may be reduplicated, but are usually not, and do not require it.

If a bivalent clause has an adjunct, then ambiguity over the scope of the quantifier arises; the restriction of the quantifier is to either the P in the clause or to an adjunct.

- (58) *Konen bere-naba lambu-lambu noʔo kti(-kti) tetiʔón.*
 3PL cut-all cloth-RED PREP knife-RED all
 ‘They cut all the cloth with knives.’ OR: ‘They cut the cloth with all of the knives.’

So far we have seen instances of the augmented quantifier with both monovalent and bivalent clauses, with the quantifier restricted to an adjunct or the P. This behaviour is different in PAV clauses. This construction can be used with a PAV construction, but the only possible interpretation is that the quantification is restricted to the A, not the P. Again we have evidence of the changed syntactic status of the P.

- (59) *Keʔo(-keʔo) konen vaʔa ka-naba tetiʔón.*
 corn-RED 3PL that eat-all all
 ‘They all ate the corn.’
 * ‘They ate all of the corn.’

It is clear that the simple *tetiʔón* construction is restricted to core arguments, any of A, S or P, as opposed to adjuncts, over which it cannot have scope. The augmented *-naba* (RED-) *tetiʔón* is differently restricted, being able to modify non-core participants or a core P: it is restricted to anything other than an S or an A. This data shows that in the PAV construction, the A cannot be interpreted as a core argument, and shows behaviour similar to the adjuncts of other clause types. The P in a PAV clause, on the other hand, behaves in a similar way to the S or A of the other clause types in not being able to be modified by the floated quantifier. This suggests that the assignment of lexical arguments to grammatical functions is different in the two clause types, the P in the PAV clause behaving like a monovalent clause’s S, and the A of the PAV clause behaving like an adjunct.

4.2 Testing for subject

The tests in the previous section allow us to judge the core status of the A in a PAV clause. We have not, however, judged the functional status of the P with respect to the A to determine which is more syntactically privileged. In other words, we have not evaluated which of A and P should be considered the subject of the clause in AVP and PAV clauses. Two tests are advanced here to investigate this question.

4.2.1 Conjunction reduction

In sentences with coordinated clauses we find that one NP of the second conjunct may be omitted under conditions of co-identity with an NP in the first conjunct. In SV+AVP conjunction the restriction on the identity of the omitted argument and its antecedent is that they must both be either S or an A in their own clause. Sentences illustrating this are shown in (60)–(63).

- S = S
 (60) *Aku pula lae nua lka ___ nodo le kandra.*
 1SG return PREP house and.then sit PREP chair
 ‘I_j came back home, and then Ø_j sat down.’

S = A

- (61) *Ama-gu pana le uma lka — take rero-n.*
 father-1GEN return PREP garden and.then meet friend-3GEN
 ‘My father_j went to the garden, and then Ø_j met his friend.’

A = S

- (62) *Ama-gu kla kaju lka — pula lae nua.*
 father-1GEN split wood and.then return PREP house
 ‘My father_j split some wood, and then Ø_j returned home.’

P ≠ S

- (63) *Ama-gu lie ina-gu lka — nodo.*
 father-1GEN see mother-1GEN and.then sit
 ‘My father_j met my mother_k, and then Ø_{j/*k} sat down.’

From the data above is uncontroversial to assume that there is a constraint that restricts conjunction reduction to members of the S,A grouping, indicating its privileged status in this construction. We do, however, find instances of S = P correspondences, but only when the P occurs preverbally in a PAV clause.

P = S with preverbal P

- (64) *Aku ia balu lka — palu lae nua-n.*
 1SG 3SG hit and.then return PREP house-3GEN
 ‘He_j hit me_k, and then Ø_k returned to his house.’
 * ‘He_j hit me_k, and then (he_j) returned to his house.’

P ≠ S if P is postverbal

- (65) *Ia balu aku lka — palu lae nua-n.*
 3SG hit 1SG and.then return PREP house-3GEN
 * ‘He hit me, and then I returned to his house.’
 ‘He hit me, and then (he_j) returned to his house.’

Other instances of the S,P grouping being the privileged one when the P is preverbal can be seen in examples (66) and (67).

S = P

- (66) *Ina lo?o-gu ia pela lka — mea-?u.*
 mother small-1GEN 3SG watch and.then shy-PERF
 ‘He_i watched my aunt_j, and then (she_j) got embarrassed.’
 * ‘He_i watched my aunt_j and then Ø_i got embarrassed.’

P = S

- (67) *Ama lo?o-de ia pela lka — kau ia.*
 father small-12GEN 3SG watch and.then angry 3SG
 ‘He_i watched my uncle_j, and then (he_j) got angry with him_i.’
 * ‘He_i watched my uncle_j and then Ø_i got angry with him_j.’

It is clear that, whatever syntactic privileges in a cross-clausal deletion construction accrue to S in a monovalent SV clause and the A in an AVP clause are also found with the P in a PAV clause, to the exclusion of those privileges being found on the A.

- (73) *Ia cube vavi tene — mata.*
 3SG shoot pig for — die
 ‘He_i shot the pig in order (for Ø_i) to die.’
 * ‘He shot the pig in order for it to die.’

In this sentence the only grammatical reading is the implausible one, that the A of the first clause, the shooter, is coreferent with the S of the second clause, the entity dying. This indicates that the constraints on cross-clausal coreference are syntactically governed, and not simply pragmatically constrained. We can confirm this impression by examining a similar bivalent–monovalent coordination, with a PAV construction in the first clause in (74).

- (74) *Vavi ia cube tene — mata.*
 pig 3SG shoot for — die
 * ‘He shot the pig in order to die.’
 ‘He shot the pig in order for it to die.’

Here the same semantic constraints on plausibility are operating, but the only possible interpretation has changed. Clearly conjunction reduction in Palu’e is governed by syntactic factors, more than semantic or pragmatic plausibility.

4.2.3 Tests for subject status

The data from coordination and purposive subordination show that there is a clearly privileged argument in both constructions: in both cases, while the S of a monovalent clause is privileged, the privileged argument in a bivalent clause is the A if the clause has AVP order, and the P if it has PAV order. Assuming that conjunction reduction, if restricted, is restricted to a subject, this means that the subject of an AVP clause is the A, and the subject of a PAV clause is the P. These facts, combined with the evidence for valency alternations presented in §4.1, clearly indicate that a voice alternation has applied in the language.

4.3 Constructions with invariant restrictions

The tests in §4.2 have shown that there is morphosyntactic evidence for an alternation in the assignment of grammatical functions to different syntactic roles in the different coding options. In this section I shall show that, if we examine the data from reflexive constructions, we find that there is also evidence for a grammatical relationship between the A and the P remaining the same.

4.3.1 Reflexive binding

A complication in the analysis is found when we examine the data that reflexive constructions allow us to examine. A standard analysis of reflexive binding involves the assumption that the higher argument (in terms of a thematic hierarchy) may bind the reflexive pronoun in a lower argument; conversely, a reflexive in a higher position may not be licensed by a lower argument. Thus given a bivalent clause with two core arguments, an

agent and a patient, an agent may bind a reflexive pronoun for the patient, but not the other way around (Dalrymple 1993). This is shown in (75a) and (75b), representing a grammatical sentence such as *He hurt himself*, and an ungrammatical sentence such as **himself hurt him*.

Reflexive binding: active clause

- (75) a. 'PRED ⟨agent, patient⟩'
 | |
 binder reflexive
- b. * 'PRED ⟨agent, patient⟩'
 | |
 reflexive binder

In a passive clause there is only one core argument; it will, by virtue of its core status, outrank any adjuncts. In this case the only reflexive that may be coded is on the (adjunct) agent, not the (core) patient. The disparity in grammatical functions overrides the difference in semantic roles. The following schema illustrates sentences such as *I was hurt by myself*, and the ungrammaticality of **myself was hurt (by me)*.

Reflexive binding: passive clause

- (76) a. 'PASS.PRED ⟨patient⟩⟨agent⟩'
 | |
 binder reflexive
- b. * 'PASS.PRED ⟨patient⟩⟨agent⟩'
 | |
 reflexive binder

When we examine the data from reflexives in Palu'e AVP clauses, the predictions from (75) are borne out: only the A may bind a P, not the other way round.

- (77) *Aku bere tmbo-gu.*
 1SG chop body-1GEN
 'I chopped myself.'
- (78) * *Tmbo-gu bere aku.*
 body-1GEN chop 1SG
 'myself chopped me'

The data for the PAV construction, however, do not fit the predictions from (76) for passive clauses. Only the P may be coded with a reflexive, bound by the A; the predicted patient binding a reflexive A does not emerge.

- (79) *Tmbo-gu aku bere.*
 body-1GEN 1SG chop
 'Myself, I chopped.' OR: 'Myself was chopped by me.'¹¹
- (80) * *Aku tmbo-gu bere.*
 1SG body-1GEN chop
 'I, myself chopped.' OR: 'I was chopped by myself.'

The reflexive data, then, do not obviously support the view that in the PAV construction the A is demoted to adjunct status. They may, however, be interpreted as suggesting that the A in the PAV construction is just as much a core argument as it is in an AVP construction, which would not be compatible with a passive analysis.

¹¹ Ungrammatical in English, but acceptable in Indonesian (with different grammatical function assignment) as *Diriku kuiris*.

4.3.2 Discussion of the reflexive data

Apart from the conclusions obtained by examining the reflexive data, the analysis of the Palu'e AVP/PAV alternation can unproblematically be described as one showing an active/passive alternation. While there are two core arguments in the AVP construction, with the A being the syntactically most privileged argument, the PAV construction presents the A as a non-core argument, and the P as the privileged argument. Apart from the lack of any morphological marking, this presents itself as a classic case of a passive alternation. The reflexive data, however, do not behave in that way. By comparison, western Austronesian voice systems typically do not involve demotion of the agent to non-argument status in the non-active voice, and so it is with the reflexive data seen in §4.3.1; this will be illustrated below.

When we compare the Palu'e reflexive data with that from other western Austronesian languages with symmetrical voice systems (Tagalog, *Tukang Besi* and Indonesian are used to exemplify these patterns), we find a remarkable congruence in the facts of reflexive binding. In the sentences below the grammatical subjects are shown in bold (the judgments for Palu'e are based on the evidence from the quantifier constructions, conjunction reduction and purposive sentences presented earlier). In the first four examples we can see the predicted pattern of the A binding a reflexive P, while the A is the grammatical subject.¹²

Voice₁, A antecedes reflexive P

- (81) ***Aku*** *pela tmbo-gu.* (Palu'e)
 1SG watch body-1GEN
 'I looked at myself.'
- (82) ***Naka-kita=ako*** *ng=sarili=ko.* (Tagalog)
 V₁:PERF-see=1SG.NOM GEN=self=1SG.GEN
 'I saw myself.'
- (83) ***Te ia*** *no-'ita te karama=no.* (Tukang Besi)
 CORE 3SG 3R-see CORE self=3GEN
 'S/he saw her/himself.'
- (84) ***Dia*** *me-lihat diri=nya.* (Indonesian)
 3SG V₁-see self=3SG.GEN
 'S/he saw her/himself.'

In the alternative voice, morphologically marked in Indonesian and Tagalog, though not in Palu'e, the identity of the grammatical subject is changed, but the conditions on binding remain the same.

Voice₂, A antecedes reflexive P

- (85) ***Tmbo-gu*** *aku pela.* (Palu'e)
 body-1GEN 1SG watch
 'I looked at myself.'

¹² In examples from Tagalog, *Tukang Besi* and Indonesian a distinction between clitics and affixes needs to be made, and so the conventions = to indicate a clitic boundary and - to indicate an affix boundary are used. These conventions differ from the presentation of Palu'e material elsewhere in this paper.

- (86) *Na-kita=ko ang=sarili=ko.* (Tagalog)
 V₂:PERF-see=1SG.GEN NOM=self=1SG.GEN
 'I saw myself.'
- (87) *Te karama=no no-'ita='e te ia.* (Tukang Besi)
 CORE self=3GEN 3R-see=3P CORE 3SG
 'S/he saw her/himself.'
- (88) *Diri=nya di-lihat=nya.* (Indonesian)
 self=3SG.GEN V₂-see=3SG.GEN
 'S/he saw her/himself.'

For Indonesian and Tagalog the explanation of these patterns is that we find the non-subject A binding the subject P in (86) and (88) because, unlike the case of English passives, there is no agent demotion involved in western Austronesian voice systems; the As in (86)–(88) are demonstrably objects, not adjuncts. This analysis is problematic for Palu'e, however, since we have good evidence, from the augmented floating quantifier construction, that the A in PAV clauses behaves as an adjunct, indicating that a common, 'garden variety' passive has applied, as far as the assignment of grammatical functions is concerned in the voice alternation (active subject corresponds to passive adjunct, active object corresponds to passive subject). All we can state with certainty is that while the AVP clauses show a clear alignment of grammatical properties, the PAV clauses show a split in these properties, in that the reflexive data does not indicate a demotion of the A.

5 Conclusions: the Palu'e voice system

We have seen that there is a passive alternation in Palu'e, although some data, here represented by the reflexive construction, do not line up with the prototypical structure that might, based on a cross-linguistic survey, be expected in the behaviour of passives. Table 2 shows which argument displays the most syntactically privileged behaviour in the constructions examined here. The columns are divided according to whether we are discussing the (bivalent) AVP construction, the PAV construction, or a monovalent construction, in which case the single argument must precede the verb.

Table 2: Grammatical evidence:
 the restriction of five different constructions

	AVP construction			PAV construction			S V	
	A	P	OBL	A	P	OBL	S	OBL
Floated quantifiers: I	✓	✓			✓		✓	
Floated quantifiers: II		✓	✓	✓		✓		✓
Conjunction reduction	✓				✓		✓	
Purposive clauses	✓				✓		✓	
Reflexives: antecedent?	✓			✓				

Almost all the evidence points clearly to the PAV construction being best analysed as a coding choice that involves an alternation in voice, compared to the AVP construction. If the PAV construction was simply a topicalised variant of the AVP construction, then we

would not expect to see the variation in behaviour that marks the A as privileged in AVP and the P as privileged in PAV clauses, nor the relative orderings with respect to time adverbials that we examined in §3. The data from all the constructions except reflexive bindings indicates that the voice alternation is an active/passive one, involving a pivotal P in the PAV construction, in which the A is syntactically oblique. The reflexive data are not consistent with this analysis. The reflexive data imply that there is no change in the status of the A and the P, which is incompatible with a passive analysis involving demotion. But the data from the *naba* RED- *tetiʔón* quantification construction clearly points to the A of the PAV construction being best analysed as an adjunct. How can this be resolved?

The answer lies in the morphological form of the Palu'e voice construction. While a typical voice alternation involves the structures seen in the left columns of Table 3, as exemplified by the English passive data in §2, and an active/antipassive pattern would be that shown in the centre columns. The Palu'e voice shows the pattern seen on the right. This is clearly typologically marked with respect to the other two patterns, which both show strong patterns of morphological asymmetry between the two coding options.

Table 3: Passive voice, antipassive voice and voice in Palu'e compared

	Active/passive			Active/antipassive			Palu'e alternation		
	A	P	V	A	P	V	A	P	V
basic voice	NP	NP	V	NP	NP	V	NP	NP	V
non-basic voice	<i>m</i> -NP	NP	<i>m</i> -V	NP	<i>m</i> -NP	<i>m</i> -V	NP	NP	V

There are, of course, languages with voice systems other than those involving passive or antipassive alternations, most notably the voice alternations found in the Algonquian languages or the western Austronesian languages, in which there is no morphological markedness relation between the two (or more) voices in the language; Table 4 compares Palu'e to representations of these language types, arranged for comparison with Table 3.

Table 4: Inverse voice, Philippine-like voice and Palu'e compared

	Inverse voice			Philippine-like voice			Palu'e alternation		
	A	P	V	A	P	V	A	P	V
voice ₁ ('A-centric')	NP	NP	<i>m</i> -V	NP	NP	<i>m</i> -V	NP	NP	V
voice ₂ ('P-centric')	NP	NP	<i>m</i> -V	NP	NP	<i>m</i> -V	NP	NP	V

Comparing these data, we can see that there is more commonality between Palu'e and the other language types, in terms of what Foley (1998) calls 'symmetricality' between the voice alternations in these systems. In both the inverse and the Philippine-type systems the amount of morphological marking on the verb is the same in both voice types, just as it is in Palu'e (of course, the fact that in Palu'e there is no morphology in both instances is significant, as we shall relate). Another point of similarity between the Philippine-type voice alternation and the Palu'e one concerns reflexive antecedency: the Palu'e data on reflexives from §4.3.1 are a challenge to the analysis of the AVP/PAV alternation as involving a passive voice alternation, but, as we saw in §4.3.2, they are consistent with the types of voice systems observed in related languages to the west. These languages,

however, do not employ passive voices,¹³ while we have seen that Palu'e does. Can we reconcile the evidence for a passive voice alternation in Palu'e with the fact that the A and P in the reflexive construction does not show a change in syntactic status?

An unambiguous example of a language with a passive construction, involving demotion of the A to an adjunct function and yet retaining the A as the antecedent of a reflexive in either passive or active voice, is Marathi (Joshi 1993). In Marathi the verb is marked as passive, and there is a special case marker for the (optional) *by*-phrase agent in these clauses, all indicating an unproblematic passive. Yet at the same time the antecedent of the *aapan* reflexive is restricted to only the *by*-phrase agent (which Joshi calls the logical subject), not the grammatical subject. Similar conditions apply to reference to gapped arguments in *un* adverbial clauses. This arises because of a condition in the grammar of Marathi that requires these constructions to refer to the argument that is highest-ranked on the thematic hierarchy from the verb's subcategorisation frame. While the agent in the sentences is clearly an adjunct, marked by the postposition *kadun*, the higher thematic role that it bears is enough to license it, and only it, being the antecedent of the reflexive, regardless of the changes in grammatical function assignment. These data are proof that it is possible for a construction to be analysed as a passive while the reflexive data behave very differently from the expected pattern. The behaviour of reflexives indicates that, at least optionally, they can best be regarded as being constrained by the relative positions of the arguments in argument structure, regardless of any grammatical-function changing operations (such as passive, or potentially other voice) that have applied to the clause.¹⁴ The reflexive data, in short, are not incompatible with a passive analysis of the voice system, though they do represent a highly unusual pattern.

What, then, of the lack of passive-marking morphology? While there are examples of languages lacking verbal (or auxiliary) morphology to indicate the passive, there is usually at least some indication of the passive, either as a VP-level marker (such as the Mandarin *bèi*, and other passive markers) or on the NP itself (such as Hokkien *ho* described in §2, or Manggarai *le* [Arka and Kosmas, this volume]). Is there a precedent for a voice alternation with no morphological marking at all?

A case similar to the Palu'e one can be found in Lango (Noonan & Bavin-Woock 1978; Foley & Van Valin 1984; Noonan 1992), in which only the order of the A and the P indicates the choice of voice. In (89) the grammatical subject is *dákó*, while in (90) it is *lócà*. The only morphological or phrase-structural difference between the two clauses is the position of the P in the clause.

¹³ See Schachter (1976, 1977) for the 'classic' presentation on these issues, Kroeger (1993) or Falk (2000a, 2000b) for more recent formal treatments. Indonesian is the exception. In Indonesian in addition to an 'A-centric' and 'P-centric' voice, which behave as described in §4.3.2 without demotion, there is a passive voice, which employs the same verbal morphology as the P-voice, but with additional nominal-marking morphology, suggesting that the importance of nominal marking, hinted at in the Hokkien examples seen in §2, is also salient in Austronesian languages. The equivalents of (88) in this true passive voice would be the clause seen in (i) below. Note that (ii) is ungrammatical, confirming the oblique status of the *oleh* phrase. See Arka and Manning (1998) for further discussion.

(i) <i>Dia di-lihat oleh diri=nya.</i> 3SG V ₂ +see by self=3GEN 'He was seen by himself.'	(ii) * <i>Diri=nya di-lihat oleh=nya.</i> self=3GEN V ₂ +see by=3SG 'Himself was seen by him.'
---	---

¹⁴ This predicts that a 'quirky reflexive' such as that in Palu'e or Marathi should be possible in a language with a non-passive voice alternation (an antipassive, for instance). To my knowledge this has not been attested.

- Voice₁: active
- (89) *Dákó ò-jwát-ò lócà.* (Lango)
 woman 3SG-hit-3SG man
 ‘The woman saw the man.’
- Voice₂: passive
- (90) *Lócà dákó ò-jwát-ò.* (Lango)
 man woman 3SG-hit-3SG
 ‘The man was seen by the woman.’

In Lango too the A is the antecedent of the reflexive in both voices, and the marking of agreement on the verb shows clearly that there is no demotion of either argument. Some controversy has been associated with the analysis of the Lango alternation shown above as a voice alternation (e.g. Woolford 1991). In *Tukang Besi*, which we have seen in examples (83) and (87), there is no *dedicated* voice morphology, but the alternation in voice is indicated by a (potential) change in case marking on the NPs, and a change in the amount of pronominal agreement found on the verb (Donohue 1999:51–54, 461–490; 2004a).

Indonesian also shows what appears to be a purely word-order defined voice alternation, but only for first and second person As, and only with that class of verbs that do not show regular voice marking (though see Chung 1978 for a caution on the analysis of bare ‘stem’ verbs as invariably representing non-active clauses in Indonesian; Cartier 1984 also provides relevant discussion). Consider the following sentences using the verb *makan*, which in this construction shows no verbal marking in the active or non-active voices, and has a first person singular A and a P that vary only in its positions, not in any NP-marking.¹⁵

- Indonesian
- Voice₁: active
- (91) *Saya makan nasi itu.*
 1SG eat rice that
 ‘I ate that rice.’
- Voice₂: ‘objective’ / ‘inverse’
- (92) *Nasi itu saya makan.*
 rice that 1SG eat
 ‘I ate that rice.’ OR: ‘That rice was eaten by me.’

This alternation seems to be identical to that found for Palu’e. The only significant difference involves the non-oblique status of the A in (92) compared to the Palu’e translation in (93), in which *aku* is oblique (as demonstrated earlier in this paper).

- (93) *Lama vaʔa aku ka.*
 rice that 1SG eat
 ‘I ate that rice.’ OR: ‘That rice was eaten by me.’

¹⁵ The lack of alternation in *makan* reflects the fact that the morphology which is functionally cognate with the active *mey-* prefix — seen earlier in (84) — is here frozen onto the lexical root as **ma-*, attached to the historical root **kaən*. This historical prefix is present even in the non-local A non-active voices, where marking with *di-* is obligatory: *di-makan*, **di-kan*.

More significantly, the apparently zero-morphological alternation seen in the Indonesian examples only manifests itself in a highly restricted set of circumstances: the A must be first or second person, and the verb must (irregularly) not take any active marking for there to be no morphological alternation. In Palu'e, on the other hand, this alternation is regular for all persons and for all verbs: any combination of A and P can appear, and no verbs are marked in either the active or the passive (other than for aspect and some forms of subordination).

In summary, the Palu'e PAV construction that we have examined can be productively analysed as a passive alternative of the AVP construction, and the apparently aberrant reflexive data are, while unusual, not unprecedented.

6 Implications for our understanding of 'voice'

The Palu'e voice is an unusual exemplar of a voice system. In analysing it, when compared to more prototypical voice systems, we must conclude that it is either the very beginning of a voice system, or the very end of one. The two competing hypotheses run as follows:

1. Modern Palu'e voice is an inceptive 'proto-voice' system
 - the contemporary voice system is a recent innovation in Palu'e, and as such has not yet acquired all the hallmarks of a fully 'mature' voice alternation, including morphological marking (on nouns and on the verb), nor has the reflexive construction yet adapted to the presence of a function-changing construction in the language;
2. Modern Palu'e voice is the relic of an earlier more elaborate system
 - an earlier stage of the language possessed a more 'complete' voice system, presumably with both morphological and syntactic characteristics that were more 'neat' with respect to their characterisation in a typology of voice.

Of course, in a real sense these two hypotheses are not 'in competition' — there is nothing to preclude both of them from being applicable and true descriptors of the Palu'e situation, and I shall return to this point later in this section. Nonetheless, the implications of these two positions are quite different, and are most easily examined in isolation from each other. If we assume position 1, then we are claiming that it is possible for a passive alternation to exist in a language in the absence of any morphological marking. In effect, we are claiming that there is such a thing as a passive construction, independent of any morphological 'load-bearers'. While this is not necessarily a bad claim, it certainly is one that is awkward for most of the widespread theoretical approaches to morphosyntax. In a theory, such as Lexical-Functional Grammar, which claims that operations such as passive, causative, applicative etc. are derived in the lexicon, the absence of any morphological material means that we must assume a zero-derivational process such as that which is assumed by some to apply to noun-verb alternations in, for instance, English. The extensive precategoriality that characterises Palu'e would lend support to this view of zero derivation (for instance, *tusu* can be used in a sentence as either a referent translatable as 'breast', or as a predicate translatable as 'suckle (on a breast)', and *kɪi*, translated here in referential functions as 'knife', can equally be used predicatively with the sense 'cut off,

sever (as using a knife)'). The data is more problematic for those theories (such as most Chomskyan-derived models) that assume that the alternation in grammatical behaviour between the active and the passive is due to a change in the available Case roles: we could imagine that the Hokkien-style passive, in which only the demoted A is marked with an extra morpheme, could be analysed as having this morpheme absorbing the object-assigning Case role, and then being realised through movement prepositionally on the preverbal oblique A; the now caseless P then moves away from the verb and acquires subject properties through its new structural position. This is obviously not possible by means of any overt morpheme in Palu'e, and the analysis can be salvaged only if we posit a phonologically-null morpheme, or movement at LF. Adopting this first position would also be the same as claiming that many of the diagnostics that are used to test kinds of voice systems, particularly the reflexive construction as a test of core or oblique status, belong together only in the passive 'construction' artefactually: that they are, in fact, independent variables that coincide only as what we recognise as voice 'types' (passive, antipassive, inverse, Philippine-type) over a long period of time.

If we were to adopt the second stance, then we would be assuming that some of the diagnostics of a 'construction' are more or less stable than others. Assuming that Palu'e previously possessed a Philippine-type voice alternation (since we have attested examples of non-demoting voice in western Austronesian languages, and a reflexive construction in Palu'e that behaves just as the reflexive behaves in these related languages), we would be claiming that the amount of morphological material slowly reduced (similar to a language-wide extension of the particular Indonesian construction seen in (91) and (92)), but that even as the morphological clues to the construction are dropping the construction is changing. The motivation for the change towards an active–passive alternation, rather than the A-voice–P-voice alternation can only be guessed at.¹⁶ The implication, however, is that syntactic change is possible without any morphological grammaticalisation: rather than morphology instigating the syntactic reinterpretation, the departure of morphology would have to be held responsible for the reinterpretation. We are also dealing, if we attempt to model this historical picture in morphologically based frameworks such as described above, a sequence of two different null morphemes. Clearly a model that admits the existence of various grammatical constructions is preferable to one that requires morphological markers to drive the syntactic derivations that it assumes. We also require the model to allow a relationship to exist between the A and the P in both the AVP and the PAV constructions that is identical, in order to drive the reflexive binding data. That is, there must be something constant in the representation of both the AVP and the PAV construction, even though there are clearly changing patterns of grammatical behaviour with respect to, for instance, floating quantifiers and conjunction reduction. This implies that we are best off considering a theoretical model that allows us to deal with separate modules of structure at the same time, such as the argument structure distinct from functional and constituent structures in Lexical Functional Grammar (e.g. Bresnan 2001).

Up to this point I have been using the label 'construction' simply as shorthand to refer to different syntactic phenomena that might just as easily have been described as 'structures' or 'derivations' or whatever other theory-specific term one might prefer. The

¹⁶ Though it is worth noting that a change from non-demotional to demotional voice types is attested in Malay/Indonesian varieties as they occur in the east of the archipelago, away from their western homeland and presumably influenced by other eastern languages (Donohue 2004b).

data we have seen, however, has an important part to play in determining the legitimacy of the label 'construction' as the most appropriate one to describe the whole nature of voice alternations. The Palu'e voice alternation cannot be easily described as being a derivation, since there is no morphology involved in its creation: both the active and the passive are unmarked by any non-lexical morphology. We might appeal to non-overt morphology, but this would be simply applying an analysis to suit the theoretical constructs that we already have in place, and not empirically examining the data. Similarly we might, in a movement-based model of syntax, assume that there was a null operator that (for instance) absorbed the accusative case assigned by the verb to the P, and so forced that argument to a VP-external position (following standard Chomskyan analyses of passives — see, for instance, Jaeggli (1986) or Baker, Johnson and Roberts (1989)). I believe that the Palu'e data best support an analysis of construction primacy. The 'construction' of voice alternation is clearly what is being manipulated historically here, not any particular morpheme. There may well have been overt voice-marking morphology in Palu'e's past, and there might be overt morphemes associated with voice alternations in the future, but the contemporary language has an unmarked passive. If it has changed from an inverse(-like) system to the modern system, then we have to accept that there are some morphosyntactic markers of the voice construction that are more resistant to change than others — the behaviour of reflexives, for instance, has not changed, even though the other evidence examined here suggests that the voice alternation is not inverse, but passive. On the other hand we must also accept that there are some aspects of the passive construction that are not as central as others: the voice alternation does not exist in Equi or raising constructions, where only active interpretations can be found. If 'passive' is a construction, this must be a non-essential part of the construction. While the only model that allows us to accurately think about the voice alternation is that of unified constructions, the idea of the 'construction' as a monolithic entity has been challenged by the inconsistencies in the behaviour of the construction.

As an epilogue, it is worth considering the analytical passive in Palu'e, which is formed with the verb *coma* '(be) affect(ed)'. This construction does not allow for the expression of the A, but also acts to change grammatical functions, as seen in the coordination data in (95).

Analytical passive in Palu'e

P V

- (94) *Vavi vaʔa coma cube.*
 Pig that affect shoot
 'That pig was shot.'

- (95) *Vavi vaʔa coma cube lka — laju lae uta.*
 pig that affect shoot and.then — run PREP bush
 'That pig was shot and then ran to the bush.'

An analytical passive such as that seen in these two sentences is clearly an innovative development in Palu'e, since most other instances of passives in Austronesian languages, and certainly the more archaic versions, involve morphological marking.¹⁷ This might imply that systemically the language is filling a 'need' for a voice, which might

¹⁷ The striking exception to this concerns the numerous analytical passives that can be found in non-standard varieties of Malay, using *kena* 'affect' or *dapa* 'receive'. These, too, seem to appear in varieties of Malay that have lost the synthetic *di-* prefix in a passive use on verbs (Donohue 2004b).

conceivably lend some weight to the second proposition above: as the earlier voice system deteriorates, a new analytical passive construction is born. But without extensive comparative data from other Flores languages, we can only speculate.

References

- Arka, I Wayan and Christopher D. Manning, 1998, Voice and grammatical relations in Indonesian: a new perspective. In Miriam Butt and Tracey King, eds *On-line LFG proceedings*. Available at <http://www-csli.stanford.edu/publications/LFG3/lfg98-toc.html>.
- Baker, Mark, Kyle Johnson and Ian Roberts, 1989, Passive arguments raised. *Linguistic Inquiry* 20:219–251.
- Bresnan, Joan, 2001, *Lexical functional syntax*. Oxford: Blackwell.
- Cartier, A., 1984, De-voiced transitive verb sequences in formal Indonesian. In Frans Plank, ed. *Objects: towards a theory of grammatical relations*, 161–183. New York: Academic Press.
- Chung, Sandra L., 1978, Stem sentences in Indonesian. In S.A. Wurm and Lois Carrington, eds *Second International Conference on Austronesian Linguistics: proceedings*, 355–365. Canberra: Pacific Linguistics.
- Comrie, Bernard, 1978, Ergativity. In Winifred P. Lehmann, ed. *Syntactic typology: studies in the phenomenology of language*, 329–394. Sussex: The Harvester Press.
- Dalrymple, Mary, 1993, *The syntax of anaphoric binding*. Stanford: CSLI Publications.
- Donohue, Mark, 1999, *A grammar of Tukang Besi*. Grammar Library series No. 20. Berlin: Mouton de Gruyter.
- 2003, The ingredients of grammatical functions: sorting subjects and pivots in Tukang Besi. Unpublished MS, National University of Singapore.
- 2004a, Voice oppositions without voice morphology. In Paul Law, ed. *Proceedings of AFLA 11, ZAS, Berlin 2004*. University of Newcastle. ZASPiL Nr. 34, October 2004:73–88. Berlin: Zentrum für Allgemeine Sprachwissenschaft, Typologie und Universalienforschung, Berlin. Also available at http://www.zas.gwz-berlin.de/index.html?publications_zaspil.
- 2004b, Voice varieties in Indonesian/Malay. Paper presented to the conference *Diversity and Universals in Language: The Consequences of Variation*. Stanford University, California, 21–23 May 2004.
- 2004c, Floating quantifiers and universal grammar. In Christo Moskowsky, ed. *Proceedings of the 2003 conference of the Australian Linguistics Society*. University of Newcastle. Published online at <http://www.newcastle.edu.au/school/lang-media/news/als2003/proceedings.html>.
- Falk, Yehuda, 2000a, Pivots and the theory of grammatical functions. In Miriam Butt and Tracy Holloway King, eds *Proceedings of the LFG00 conference, University of California at Berkeley*, 123–139. Stanford: CSLI Publications.
- 2000b, Philippine subjects in a monostratal framework. In Carolyn Smallwood and Catherine Kitto, eds *The proceedings of the Austronesian Formal Linguistics Association VI*, 133–136. Toronto Working Papers in Linguistics.

- A more complete version of this paper is available at
<http://pluto.mscc.huji.ac.il/~msyfalk/PhilippineSubjects.pdf>.
- Foley, William A., 1998, Symmetrical voice systems and precategoriality in Philippine languages. Unpublished MS, Department of Linguistics, University of Sydney.
- Foley, William A. and Robert D. Van Valin, 1984, *Functional syntax and universal grammar*. Cambridge Studies in Linguistics 38. Cambridge: Cambridge University Press.
- Haspelmath, Martin, 1990, The grammaticization of passive morphology. *Studies in Language* 14 (1):25–72.
- Jaeggli, Osvaldo, 1986, Passive. *Linguistic Inquiry* 17:587–633.
- Joshi, Smita, 1993, Selection of grammatical and logical functions in Marathi. PhD dissertation, Stanford University.
- Klaiman, M.H., 1991, *Grammatical voice*. Cambridge: Cambridge University Press.
- Kroeger, Paul R., 1993, *Phrase structure and grammatical relations in Tagalog*. Stanford: CSLI publications.
- Noonan, Michael, 1992, *A grammar of Lango*. Berlin: Mouton De Gruyter.
- Noonan, Michael and Edith Bavin-Woock, 1978, The passive analog in Lango. In *Proceedings of the fourth annual meeting of the Berkeley Linguistics Society*, 128–139. Berkeley: Berkeley Linguistics Society.
- Schachter, Paul, 1976, The subject in Philippine languages: Topic, Actor, Actor-Topic or none of the above. In C.N. Li, ed. *Subject and Topic*, 491–518. New York: Academic Press.
- 1977, Reference-related and Role-related properties of Subjects. In Peter Cole and Jerrold M. Sadock, eds *Syntax and semantics 8: grammatical relations*, 279–305. New York: Academic Press.
- Shibatani, Masayoshi, ed., 1988, *Passive and voice*. Amsterdam: John Benjamins.
- Siewierska, Anna, 1984, *The passive: a comparative linguistic analysis*. London: Croom Helm.
- Van Valin, Robert D. and Randy J. LaPolla, 1997, *Syntax: structure, meaning and function*. Cambridge: Cambridge University Press.
- Woolford, Ellen, 1991, Two subject positions in Lango. In Kathleen Hubbard, ed. *Proceedings of the seventeenth annual meeting of the Berkeley Linguistics Society, February 15–18, 1991: special session on African language structures*, 231–243. Berkeley: Berkeley Linguistics Society.

4 *Passive without passive morphology? Evidence from Manggarai*

I WAYAN ARKA AND JELADU KOSMAS

1 Introduction

This paper deals with a passive-like construction in Manggarai¹ which appears to be typologically unusual because it has no specific verbal passive morphology on the verb.² Rather than verbal marking, the passive in Manggarai is marked analytically on the Agent argument, i.e. by means of the preposition *le*, which can be shortened to *l=*.³ This is illustrated by the sentences in (1): sentence (1a) is a canonical sentence with the Agent

¹ Manggarai is a language belonging to the Central-Malayo Polynesian subgroup of Austronesian languages (Blust 1993), spoken by around 400,000 speakers in western and north central Flores (Indonesia). It has four main dialects (Verheijen 1991:315): the dialects of East Manggarai, Central Manggarai, S>H Manggarai (named after the /s/ and /h/ correspondence, also called Rego), and West Manggarai. Unless otherwise stated the data reflect the dialect of Central Manggarai as described in Kosmas (2000). The second author is a native speaker of Manggarai.

² We thank Bill Foley, Mark Donohue, Adrian Clynes, and Erik Zobel for their comments on the earlier version of this paper when it was presented at the Ninth International Conference on Austronesian Linguistics, Canberra. Special thanks must go to Laurie Reid, Bernard Comrie, Jane Simpson and Malcolm Ross who have read the paper thoroughly and given us detailed and useful comments. All errors are, however, ours.

³ The preposition *le* may appear as *l=* and *li* (and sometimes *ali*). These allomorphs are partly phonologically and partly lexically determined. *Le* is used for common nouns and pronouns beginning with a consonant (e.g. *le polisi* ‘by the police’, *le hau* ‘by you’, *le meu* ‘by you’). Clitic *l=* is used when the following noun begins with a vowel, e.g. *l=aku* ‘by me’, *l=ami* ‘by us (inclusive)’, *l=iha* ‘by him/her’. *Li* is used when the following noun object is a proper name (or kin terms used as such), e.g. *li John* ‘by John’, *li kakak* ‘by our/your elder brother/sister’. It should be noted that *li=* is replaced by *le=* in some dialects (e.g. in Rego).

coming before the verb, and sentence (1b) is a pragmatically marked structure with the backgrounded Agent coming after the verb and marked by *l=*.⁴

- (1) a. *Aku cero latung=k.*
 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.'

In this paper we argue that sentence (1b) is indeed syntactically passive. That is, (i) the Patient *latung*, which was object⁵ in (1a), is subject in (1b);⁶ and (ii) the Agent *aku* marked by prepositional clitic *l=* in (1b) is syntactically a non-core argument. We will present the evidence shortly to support the proposal that sentence (1b) is an instance of passive despite the fact that the verb has the same form as in (active) sentence (1a). We argue that the non-typical characteristics of the *le* passive in Manggarai are independently motivated by Manggarai's language-specific property as an isolating language.

The paper is organised as follows. First, basic surface clause structures in Manggarai are presented in §2, followed by a brief discussion on clitic sets in §3. Evidence for passive constructions without passive morphology is given in §4. A typological note on the analysis is discussed in §5. Finally, §6 forms the conclusion.

2 Basic clause structures

Morphologically, Manggarai is isolating (i.e. words in Manggarai are typically monomorphemic). Structurally, it could be classified as a head-initial language. Its clause structure, for example, is typically VOS or SVO and its phrasal syntax has its head coming first (e.g. prepositional phrase instead of postpositional phrase). Subject (and in certain circumstances, object) can be expressed by an NP which is also cross-referenced by a pronominal enclitic that agrees with it. Pronominal clitic agreement associated with subject is exemplified by *aku* (subject NP) and *=k* (enclitic) in (1a). Agreement associated with object is given in (7) below.

⁴ The following abbreviations are used in this paper: 1, 2, 3 first, second and third person; A actor, APPL applicative; ART article; ASP aspect; AV agentive voice; DET determiner; GEN genitive; IN inchoative; IRR irrealis; NOM nominative; NUM number; OBJ grammatical object; OBL oblique; OV objective voice; p plural; subscript p polite; P patient; PASS passive; PERS person; POSS possessive; R realis; s singular; S (intransitive) subject.

⁵ The terms 'subject' and 'object' are used in this paper to refer to 'surface' grammatical relations. A semantic role can be manifested in different grammatical relations, e.g. a Patient can be grammatically an object or subject.

⁶ It should be noted, however, that the object in (1a) and the subject in (1b) are not semantically the same: one is indefinite, the other definite. Sentence (1a) is a natural active sentence with an indefinite object but it would be acceptable also with a definite object. Strictly speaking, (1b) cannot therefore be a 'true derivation' of (1a). Indeed, there is evidence from western Austronesian languages that passive is not derived from active or vice versa; see footnote 46.

We will show later that an enclitic alone or an NP alone can function as subject (or object). When both the enclitic and the free NP are present, we argue that the enclitic is the ‘real’ subject or object.⁷ Unlike subject and object (which are core arguments), an oblique can never be cross-referenced by a pronominal enclitic. Moreover, an oblique is more mobile than subject and object (further discussed in §4.1).

The following are some important features of Manggarai syntax.

First of all, the sentence-initial free NP is in fact the default TOP(ic) and it is also by default the grammatical subject of the sentence (which is possibly clause-external in an extended⁸ clause structure). The inference that it is the default TOP comes from the restriction that it must be definite. For example, omitting the article *hi*⁹ of *hi enu* ‘the girl’ from the subject NP in sentence (2a) below gives rise to a bad sentence as shown by (2b):

- (2) a. *Hi enu cebong sili tiwu lewe.*
 ART girl bath in pool long
 ‘The girl took a bath in the long pool.’
 b. * *Enu cebong sili tiwu lewe.*
 girl bath in pool long.

The TOP NP can be ‘cross-referenced’ by an enclitic pronoun.¹⁰ This is already exemplified by the transitive sentence in (1a), in which enclitic =*k* agrees with *aku*. More examples are given in (3), where the sentences are intransitive. In (3a), enclitic =*i* (third person singular) agrees with *hia* and in (3b) it agrees with *hi Kode*. In (3c), enclitic =*s* (third person plural) agrees with *ise*. In contrast to (3c), sentence (3d) is bad because agreement is violated (i.e. =*i* is incompatible with *ise*).

- (3) a. *Hia pa’u eta mai bubung mbaru hitu=i.*
 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.’
 c. *Ise lonto musi mai ami=s.*
 3p sit behind from 1pi=3p
 ‘They sat/were sitting behind us.’
 d. * *Ise lonto musi mai ami=i.*

⁷ That is, the enclitic is the argument that bears subject or object relation whereas the free NP may simply carry the pragmatic function of TOP. This TOP NP is also identified as the subject/object due to its ‘anaphoric’ relation with the enclitic.

⁸ An ‘extended clause’ is a clause structure that includes an extra clausal NP in the TOP (possibly adjoined) position; to be discussed shortly with reference to the tree diagram in (5) below.

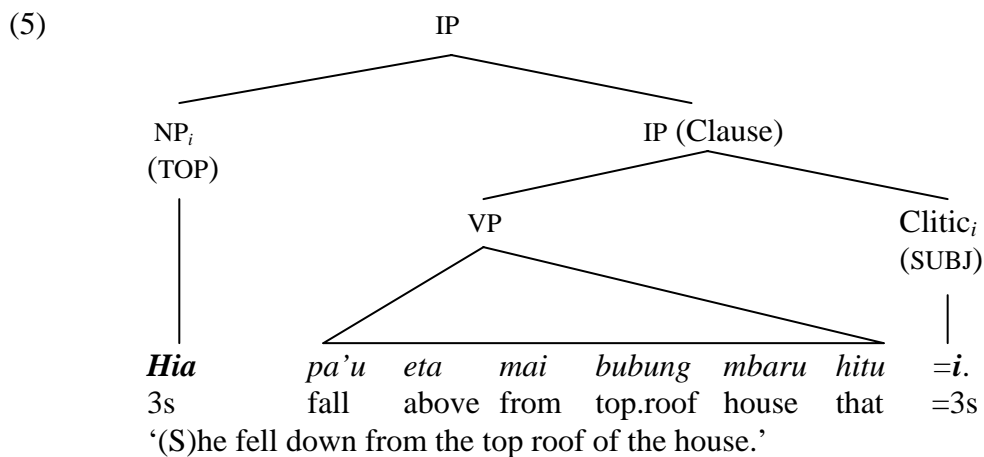
⁹ The determiner (e.g. *hitu* ‘that’) also has the same function as the article.

¹⁰ The cross-reference of a clitic by a free pronoun/NP in a clause is optional. However, when it is present, it may be associated with topicality (where the clitic appears to act like a resumptive pronoun), or it may also be accompanied by certain emphasis or additional meaning (typically ‘surprise’ or ‘disbelief’), which is absent when there is no cross-reference. Semantic and pragmatic aspects of enclitic agreement are discussed in detail in Arka (to appear.a).

The enclitic pronoun may function as an anaphoric pronoun. The antecedent of the enclitic may be within the same (extended) clause. This type of anaphoric relation is exemplified by =*i* and =*s* in (3). However, the antecedent of the enclitic may be in a different clause or sentence preceding the clause containing the enclitic. Therefore, the enclitic in this context generally functions to maintain topic continuity in a stretch of sentences. For example, in the following quotation from a story, the enclitic =*i* in (4), which is the subject in the first sentence and an object in the second one, refers to a child (who gets lost in the jungle and becomes a sago tree), mentioned in the preceding clause:

- (4) *Iti wa tanah ghitu=i. Ghau peke waeng neteng loho=i.*
 that down soil that=3s 2s must water every day=3s
 'He is down there inside the ground. You must water him every day.'

Let us be specific about the structural position of the sentence-initial NP that acts as the antecedent for the enclitic pronoun. It appears that it occupies a clause-external position of TOP position within an extended clause structure (see Bresnan 2001:116). To illustrate the point, consider sentence (3a), which can be analysed as having the structure shown in (5). In this structure, the verb *pa'u* and the following PP *etamai bubung* are within VP; and then the VP and the enclitic =*i* form the clause. (For simplicity, the internal structure of VP is not shown.) The clause is represented as IP, which represents a projection of the category I, where I is a functional category that carries Tense or Aspect information, i.e. traditionally the auxiliary. (IP corresponds to the traditional symbol S (=sentence)). The I category is not realised in this structure and is not shown in the tree.¹¹ The (pronominal) clitic subject is within the clause, represented as within the (lower) IP. The extended clause structure is represented as the higher IP, which includes a TOP position outside the lower IP. The TOP position is an adjunction position (i.e. a position that is adjoined to IP to form another higher IP).¹² The TOP NP and the subject enclitic are co-indexed by a subscript *i* to show that they must agree in their referential features (PERS and NUM), otherwise the structure is not acceptable (e.g. (3d)).



¹¹ Words belonging to this I category include *reme* 'PROG', *paka* 'must' and *kudut* 'FUT'. These auxiliaries precede verbs in Manggarai.

¹² Alternatively, one could adopt an analysis where the TOP NP position occurring sentence-initially is not adjoined to the IP. It is perhaps in the specifier position of the CP (complement phrase). We leave this for further research.

Supporting evidence for the analysis shown in (5), where the enclitic pronoun is subject, comes from the structural properties of the enclitic. In particular, it can be attached to the final word of different kinds of syntactic units contained in the VP, that precede it within the clause. For example, it can be attached to the object NP of a transitive verb as in (1a), to a PP agent as in (1b), and, if the verb is intransitive, to the last word of an oblique/adjunct PP as in (5).

In real texts, either the free NP TOP alone or the enclitic pronoun alone can appear as subject. This gives the impression that either of them is optional. However, they cannot both be absent in a given clause (unless the subject is controlled in an embedded clause). Hence, sentences (6a, b) are both fine, but sentence (6c) is not (as an independent clause). The fact that the two nominals which are involved in the agreement do not have to show up together at the same time suggests that the nature of the ‘agreement’ is not grammatical (i.e. it is not like subject–verb agreement in English).

However, in certain circumstances, the enclitic pronoun can function as an object (but not an oblique). Consider examples (7). Sentence (7a) is a simplified version of the second clause in (4), where =*i* is the patient object of the verb *waeng*. The enclitic (=i) cannot be understood to refer to the subject *ghau* ‘2s’ (because =*i* and *ghau* have different referential features). In (7b) enclitic =*i* is cross-referenced by the (Patient) NP *ata hitu* ‘that person’, which follows it. Note that the Patient NP cannot be extraposed sentence-initially (i.e. unacceptability of (7c)). The unacceptability of sentence (7c) suggests that the enclitic object agreement is highly constrained: the enclitic that agrees with the sentence-initial TOP is restricted to the subject function, i.e. it cannot be object.¹³

- (6) a. *Hia onggga aku.*
 3s hit 1s
- b. *Ongga aku=i.*
 hit 1s=3s
- c. * *Ongga aku.*
 ‘(S)he hit me.’
- (7) a. *Ghau peke waeng=I ...* (Rego)
 2s must water=3s
 ‘You must water him ...’
- b. *Aku kawe=i ata ghitu.*
 1s look.for=3s person that
 ‘I’m looking for that person.’
- c. * *Ata ghitu Aku kawe=i.*

Furthermore, it appears that the distribution of the enclitic pronoun functioning as object is also constrained by the PERS category of the object. For example, the first and second person object enclitics are not possible as illustrated by the contrast in (8). Sentence (8a) is

¹³ The presence of the object clitic appears to exclude the presence of the subject clitic. That is, the construction of the type * *kawe=i=k* ‘look.for=3s=1s’ ‘I’m looking for him’ is not acceptable. It is not clear at this stage why this is the case.

bad because the Patient object is =*k*, the first person pronominal clitic. To make this sentence acceptable, a free (i.e. non-clitic) pronoun (*aku*) must be used (8b).¹⁴

- (8) a. * *Ghia onnga=k*. (Rego)
 3s hit=1s
 ‘(S)he hit me.’
- b. *Ghia onnga aku*.
 3s hit 1s
 ‘(S)he hit me.’

While subject and object arguments in Manggarai may be expressed as enclitics and/or receive enclitic agreement, oblique arguments may not. Thus, in contrast to (9a) where the Goal oblique is expressed by *hia* (a free pronoun), sentence (9b) is bad because the oblique is expressed by the corresponding clitic form, =*i*. The contrast in (9c–d) further shows that an oblique argument cannot receive enclitic agreement.

- (9) a. *Aku tombo agu hia*.
 1s talk with 3s
 ‘I talked with him/her.’
- b. * *Aku tombo agu=i*.
- c. *Aku tombo agu hi Joni*.
 1s talk with ART name
 ‘I talked with Joni.’
- d. * *Aku tombo agu=i hi Joni*.

To conclude, enclitic agreement is a property of core (i.e. subject and object) arguments in Manggarai (bearing in mind that object enclitics have certain restrictions). Obliques do not share this property.

3 (En)clitics in Manggarai

There are three kinds of clitics in Manggarai that are of interest. The three are shown in Table 1: pronominal subject/object (or core) enclitics (column 3), pronominal GEN(itive) enclitics (column 4), and the POSS(essive) *d(e)* clitic shown here with free pronominal forms (column 5).¹⁵ For simplicity’s sake, only the GEN and POSS clitics are discussed in this subsection.

¹⁴ The exact condition of the distribution of the enclitic pronoun that may function as object needs further study. At this stage, bearing in mind that there are restrictions of the distribution of the object enclitic agreement, it is enough to point out that pronominal enclitic agreement is a property of core in Manggarai because a non-core argument never gets this property. That is, while the relation between the clitic and its antecedent is anaphoric in nature (i.e. not syntactic), the distribution of the clitic itself is syntactically constrained because it must be a core argument.

¹⁵ We do not label the *d(e)* clitic as Genitive to underline the fact that the clitic itself is not inflected and is formally distinct from the pronominal (inflected) GEN clitic. It is not clear at this stage whether historically *d(e)* is a Locative marker corresponding to *di* found in other Austronesian languages of Indonesia. A locative in contemporary Manggarai is expressed by an unrelated prepositional form, *one*. For this reason, *d(e)=* is labelled as POSS rather than Locative even though we do not exclude the possibility that the possessive meaning might be historically derived from the locative meaning (e.g. ‘something is with/in/at me’). (This point was raised for us by Laurie Reid.)

The pronominal GEN clitic is inflected for PERS and NUM whereas the POSS clitic is not. The POSS clitic may appear, however, as either syllabic *de* or non-syllabic (i.e. only consonant) *d=*.¹⁶ Both GEN and POSS clitics are used to express possession and nominalisation. Nominalisation examples are given in (11)–(12) below.

Table 1: Free pronouns and pronominal clitics in Manggarai

(1)	FREE (2)	SUBJ ¹⁷ (3)	GEN (4)	POSS.(with free pronouns) (5)
1s	<i>aku</i>	= <i>k</i>	= <i>g</i>	<i>d=(aku)</i>
2s	<i>hau</i>	= <i>h</i>	= <i>m</i>	<i>de=(hau)</i>
3s	<i>hia</i>	= <i>i</i>	= <i>n</i>	<i>d=(iha)</i>
1p.e	<i>ami</i>	= <i>km</i>	= <i>gm</i>	<i>d=(ami)</i>
1p.i	<i>ite</i>	= <i>t</i>	= <i>d</i>	<i>d=(ite)</i> (plural polite form)
2p	<i>meu</i>	= <i>m</i>	= <i>s</i>	<i>de=(meu)</i>
3p	<i>ise</i>	= <i>s</i>	= <i>d</i>	<i>d=(ise)</i>

Another difference between the two kinds of clitics is illustrated in (10). The pronominal GEN enclitic is attached to the possessed noun whereas the POSS *d(e)* clitic is generally attached to the pronominal possessor. The pairs in (10) illustrate alternative ways of expressing possession in Manggarai.

- (10) a. *buku=k* *buku* *d=aku* ‘my book’
 book=1s.GEN book POSS=1s
- b. *buku=m* *buku* *de=hau* ‘your(s.) book’
 book=2s.GEN book POSS=2s
- c. *buku=t* *buku* *d=ite* ‘our book’
 book=1pi.GEN book POSS=1pi
- d. *buku=s* *buku* *de=meu* ‘your (p.) book’
 book=2p.GEN book POSS=2p

Examples (11)–(14) show that the POSS and GEN clitics are used to mark the ‘possessor’ in nominalisation. The crucial point to note is that, using Dixon’s/Comrie’s terminology (Comrie 1978; Dixon 1979; 1994), the possessor can be S (the sole core argument of the intransitive verb), A or P (Actor and Patient core arguments of the transitive verb).¹⁸ Examples in (11) show that POSS *de=* NPs are associated with S while the ones in (12) show they are associated with A (12a) and P (12b).

¹⁶ In other dialects such as Kempo, *d=* becomes *g=* in some members of the clitic group: *g=aku* ‘1s.GEN’, *g=au* ‘2s.GEN’, *g=ami* ‘1p.e.GEN’, *g=emi* ‘2p.GEN’. It should be noted that the clitic *d=* may attach to the preceding unit, hence it shows up as *=d*, especially in its use in nominalisation (see §4.2).

¹⁷ In the Kempo dialect, the pronominal enclitics functioning as subject enclitics and GEN enclitics are not distinguishable, see Semiun (1993).

¹⁸ The fact that the enclitics in Manggarai can be associated with the three core arguments (S, A and P) suggests that morphologically speaking nominalisation does not provide evidence for accusative or ergative system alignment.

- (11) a. *mata de keraeng tu'a*
die POSS master old
'the death of the older brother/sister'
- b. *mai de dokter hitu*
come POSS doctor that
'the coming/arrival of the doctor'
- (12) a. *weli de ende*
buy POSS mother
'the buying (of/by) Mother'
- b. *pika=d kaba situ*
sell=POSS buffalo that
'the selling of the buffalo'

Pronominal GEN clitics can also be associated with an S possessor as in (13a), a P possessor as in (13b), and an A possessor as in (14). There are two things to be noted here. First, the GEN enclitic must agree with the following nominal: (13c) is unacceptable because the third person singular clitic =*n* does not agree with the plural NP *ata situ*. Second, the property of the bound pronominal as a clitic is evident from the fact that its distribution is not constrained by the grammatical category of the host (Zwicky 1985, 1987; Zwicky & Pullum 1983). For instance, in (14a), it appears attached to the head verb whereas in (14b) it is cliticised to the Patient NP (of the VP). It should be noted that the first clitic =*n* agrees with the NP *acu hitu* (i.e. the Agent) in the examples in (14) whereas the second one appearing on the predicate *dilem-dilem* 'deep-deep' agrees with the nominalised V(P) *akit (indus)=n*.

- (13) a. *lako=n ata hitu*
go=3s.GEN person that.s
'the leaving of the person'
- b. *weli=n mbaru hitu*
buy=3s.GEN house that
'the buying of the house'
- c. * *lako=n ata situ*
go=3s.GEN person that.p
'the leaving of the **persons**'
- (14) a. *Akit=n acu hitu dilem-dilem=n.* Kempo
bit-3s.GEN dog that.s deep-deep=3s (Semiun 1993:77)
'The dog's bite is very deep.'
- b. *Akit indus=n acu hitu dilem-dilem=n.*
bit cat-3s.GEN dog that.s deep-deep=3s
'The dog's bite on the cat is very deep.'

To sum up, Manggarai is poor in verbal morphology and makes use of clitics to show alternative syntactic expressions of the argument (S, A or P) of the verb.

4 The passive *le*

There is evidence that the *le* construction that gives rise to a backgrounding effect is syntactically passive. That is, we argue that the alternation shown in (1), repeated as (15), involves (a) syntactic demotion of the Agent to non-core status and (b) syntactic promotion of the Patient to subject status.

- (15) a. *Aku cero latung=k.*
 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.'

In what follows, we present evidence that the *le* Agent is not subject. There is convincing evidence that it is an oblique.

4.1 The Agent marked by *le* is not subject

First, evidence for the idea that the *le* Agent is not subject comes from enclitic agreement. Recall that TOP-subject can have pronominal agreement. In (15a), *aku* is TOP and the real subject is the enclitic *=k*. In (15b), there is a change in agreement triggered by the promotion/ foregrounding of the Patient NP *latung hitu*: this NP now agrees with the enclitic *=i*. This is what is expected on the analysis that sentence (15b) is a passive structure. That is, the Patient is indeed grammatically subject of this sentence and the *le* agent is not subject. Then, we expect that the *le* Agent can no longer have clitic agreement. This is confirmed as shown by the unacceptability of (16a, b):

- (16) a. **Latung hitu cero=k l=aku.* (cf. (15b))
 b. **Latung hitu cero l=aku=k.*

Second, further evidence comes from structural positions. Given the analysis that the *le* Agent has non-core status, it is expected that it is adjunct-like and is therefore more mobile than the NP agent. We expect that it may appear in several positions. This is confirmed. For example, the *le* Agent can appear after the subject enclitic as in (17a), after a sentence adverbial like *meseng* 'yesterday' as in (17b), or it can even appear before the subject enclitic as already observed in (1b, 15b), repeated here again for comparison as (17c).

- (17) a. *Kala situ toto=s l=ise.*
 betel.nuts that.p show=3p by=3p
 'The betel nuts were shown by them.'
- b. *Latung hitu cero meseng l=aku.*
 corn that fry yesterday by=1s
 'The corn was fried by me yesterday.'
- c. *Latung hitu cero l=aku=i.* (= (1b/15b))
 corn that fry by=1s=3s
 'The corn is (being) fried by me.'

In contrast, the Agent NP typically appears in two fixed positions: sentence-initially before a verb phrase (VP) or sentence-finally after a VP. Crucially, it cannot come within a VP (i.e. it cannot intervene between verb and its Patient object NP). Thus, the following contrast is expected in the present analysis where the *le* Agent is not subject:

- (18) a. *Cero l=aku latung hitu.* V – PP(OBL) –NP(SUBJ)
 fry by=1s corn that
 ‘I fried the corn.’
- b. **Cero aku latung hitu.* V – NP(SUBJ) – NP (OBJ)
 fry 1s corn that
 ‘I fried the corn.’

Sentence (18a) (acceptable) is a passive structure with the Patient coming sentence-finally as subject while the oblique *le* Agent appears (possibly inside the VP) before the Patient NP. In contrast, structure (18b) is not allowed in Manggarai because it would be interpreted as an active structure with its Agent (subject) NP intervening in the VP sequence (i.e. between the verb and the Patient object).

Note that our analysis of Manggarai clause structures claims: (i) that the Agent NP and the *le* Agent PP are syntactically different arguments (i.e. the TOP/subject NP vs the Agent oblique PP) and (ii) that structurally they occupy different positions. Further evidence to support these claims comes from the contrast shown in (19). Sentence (19b) is bad because the *le* Agent (*l=ise*), which is an oblique, is forced to appear in the (default) TOP/subject position before the verb.¹⁹

- (19) a. *Ise onnga ata hitu.*
 3p hit person that.s
- b. **L=ise onnga ata hitu.*
 by=3p hit person that.s
 ‘They hit the person.’

Third, the *le* construction appears to be driven by the same (syntactic) motivation as drives a passive structure in some other languages, namely the need for the Patient to be manifested as subject to meet some syntactic requirement such as control, in which case the Agent is then forced to be demoted or backgrounded. Manggarai is like some other Austronesian languages of Indonesia in preferring the subject argument to be controlled. Constructions that involve control are, among others, purposive clauses, participial clauses and relative clauses. In what follows, we exemplify each of these clause types in turn.

Consider the purposive clauses (marked by *te*) in (20). (The controlled or gapped argument position is indicated by a dash.) The purposive clause in (20a) is intransitive with its sole argument controlled and gapped. The clause in (20b) is (active) transitive, its Agent being the subject argument. Control is fine. Sentence (20c) shows a failed attempt

¹⁹ However, sentence (19b) can be made acceptable by nominalising the head verb as in shown below:

L=ise onnga-n ata hitu.
 by=3p hit-GEN3s person that.s
 ‘They hit the person.’ (lit. ‘The hitting of the person was by them.’)

This sentence is quite different from sentence (19b). It is not a passive counterpart of (19a); rather it is the one where the preposition (*le*) acts as the head predicate and the nominalised verb acts as the subject argument of the predicate. We will not discuss the issue of nominalisation in this paper any further.

to control the Patient argument because the Patient is the object argument of the purposive clause. Finally, sentence (20d) shows the acceptable version of (20c) where a passive structure is used allowing the Patient argument to be expressed as subject. Note that the Agent of this purposive clause must appear in its oblique function, marked by *le*.

(20) Controlled clauses: purposive clauses:

- a. *Ame mo one lo'ang [te __ toko].* Kempo
 father go Loc bedroom *te* sleep (Semiu 1993:93)
 'Father went into the bedroom to sleep.'
- b. *Ame mo le uma [te __ weri mawo].*
 father go to garden *te* plant rice
 'Father went to the garden to plant rice.'
- c. * *Ame mo wa kota [te dokter priksa __].*
 father go Loc town *te* doctor examine
 'Father goes to town to be examined by (the/a) doctor.'
- d. *Ame mo wa kota [te __ priksa le dokter].*
 father go Loc town *te* examine by doctor
 'Father goes to town to be examined by the doctor.'

Participial adjuncts (or perhaps 'nominalised' complements) involving control are exemplified in (21).²⁰ Sentence (21a) (acceptable) shows that the Agent of an active participial adjunct is controlled. Sentence (21b) (unacceptable) shows an attempt to control the Patient of an active participial adjunct (i.e. object-controlled). In order to have an acceptable version of (21b) where the Patient is controlled, the Patient must be manifested as subject and the Agent is demoted and backgrounded to oblique as shown in

²⁰ Unlike the purposive clauses, the participial adjuncts involving intransitive verbs have no 'clear' control structure as shown by the unacceptability of sentence (a) below. Instead, intransitive verbs are obligatorily nominalised; both GEN structures are possible as shown by (b) and (c).

- a. * *Aku porong hia lako.*
 1s see 3s walk
 'I saw him walking.'
- b. *Aku porong lako d=iha.*
 1s see walk GEN=3s
 I saw him walking.' (lit. 'I saw his walking.')
- c. *Aku porong lako=n.*
 1s see walk=GEN.3s
 'I saw him walking.' (lit. 'I saw his walking.')

The difference between participial adjuncts and purposive clauses with respect to intransitive verbs is perhaps due to the marking. Purposive clauses have a special marking to show subordination, namely *te*, whereas participial clauses have no such special marking. Manggarai, then, extends GEN marking for this purpose. However, it is still not clear why 'a gap' strategy in the subject position is ruled out because a transitive verb allows this strategy as shown in (21a). (Other Austronesian languages of Indonesia such as Indonesian and Balinese allow both gapping and nominalisation strategies.) It should be noted that the GEN strategy is also possible for a participial adjunct with a transitive verb as shown by the following example:

- d. *Aku ita tengo acu-e d=ia.* (Kempo, Semiu (1993), ex. 252)
 1s see hit dog-? GEN=3s
 'I saw him hitting the dog.' (lit. 'I saw his hitting (of) the dog.')

(21c). This is just like the purposive clause exemplified in (20c–d), where the passive structure must be used.

(21) Controlled clauses: participial adjunct:

- a. *Aku ita hia* [___ *emi seng hitu*]. (A is controlled)
 1s see 3s take money that.s (active participial)
 ‘I saw him taking the money.’
- b. * *Aku ita hia* [*polisi deko* ___]. (P is controlled,
 1s see 3s police arrest Active participle)
 ‘I saw him arrested by the police.’
- c. *Aku ita hia* [___ *deko le polisi*]. (P is controlled)
 1s see 3s arrest by police (passive participial)
 ‘I saw him arrested by the police.’

The same restriction with respect to control is also observed in relativisation. The relative markers differ depending on the dialects, e.g. *se* (Rego) and *ca* (Kempo). The following examples (from Rego) show that non-subject arguments cannot be controlled:

- (22) a. *Ata molah* [*se* ___ *ita aku*] *ghitu rebao ngo gi*. (Rego)
 person girl REL see 1s that just.now go already
 ‘The girl [who saw me] has just gone.’
- 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.’
- d. * *Ata molah* [*se* *aku ita le* ___] *ghitu rebao ngo gi*.
 person girl REL 1s see by that just.now go already
 * ‘The girl [that I was seen by ___] has just gone.’

Relativisation of the Agent argument of the active structure is fine (22a). No change in syntax of the embedded clause is observed (except the gapping of the Agent subject position). Again, an attempt to relativise/control the Patient object gives rise to a bad sentence (22b). To enable the Patient to be relativised, it should be manifested as subject, which triggers a passive structure, in which case the Agent is expressed as an oblique (22c). Relativisation of an oblique also leads to a bad sentence as exemplified by (22d).

Relative clause structures are also used to express contrastive FOC, in which case the same constraint applies. For example, when the patient argument is given contrastive FOC, a passive structure is used as shown in (23b).

- (23) a. *Joni ca* [___ *tengo acu*]. (Kempo)
 name REL hit dog Semiu (1993, ex. 278)
 ‘It is John who hit the dog.’
 ‘John is the one who hit the dog.’

- b. *Joni ca [__ tengo le²¹ polisi].* (Kempo)
 name REL hit by/of police Semiun (1993, exx. 279, 280)
 ‘It is John who was hit by the police.’

The evidence presented so far indicates that the PP Agent (marked by *le*) is not syntactically subject. Let us now turn to evidence from reflexive binding, optionality, and grammatical encoding which suggests that the PP (*le* Agent) is indeed an oblique.

4.2 The Agent marked by *le* is an oblique argument

Evidence from reflexive binding strongly indicates that the *le* Agent has non-core status. Reflexive binding in Manggarai is sensitive to a grammatical relations hierarchy. One such hierarchy (observed across languages and also relevant in Manggarai) is the Subject>Object>Oblique hierarchy (Keenan & Comrie 1977, among others; Bresnan 2001), where notation ‘>’ means ‘... more prominent than ...’.²² The evidence comes from the following examples:

- (24) a. *Hia_i mbele weki ru-n_i* <*hia_i, self_i*> (active)
 3s kill body self-3s.GEN Subj. Obj.
 ‘S/he killed himself/herself.’
- b. *Hia_i mbele le ru-n_i* <<*hia_i*><*self_i*>> (passive)
 3s kill by self-3s.GEN Subj. Obl.
 ‘S/he was killed by himself/herself.’

In the active structure (24a), the reflexive (*weki run*) is thematically Patient and syntactically object. Since it is object, it is grammatically outranked by *hia* (subject). Binding of object (*run*) by subject (*hia*) is therefore fine. In (24b), the syntactic ranking of the Agent and Patient is reversed: the Patient is syntactically subject (the highest ranked syntactic function) whereas the Agent now gets marked by *le*. The evidence of the reversal of ranking comes from the fact the Patient of the passive construction in (24b) (i.e. *hia*) can bind the *le* Agent (*le run*). That is, the Agent reflexive *run* marked by *le* must be grammatically outranked by (i.e. lower than) the Patient *hia/wekin*, otherwise binding would not be possible. We take this evidence of successful binding of the *le* Agent by sentence-initial Patient *hia* to suggest that sentence (24b) is indeed a passive sentence where the *le* Agent is an oblique.

²¹ The POSS marker *de* can be also used here instead of *le*. According to Semiun (1993:83) the difference between *le* and *de* is the semantics of the verb: *le* implies an action-like situation, whereas *de* implies a state-like situation.

- a. *Aku cumang Joni ca deko le polisi.*
 1s meet name REL catch by police
 ‘I met John, who was caught by the police.’
- b. *Ho’o loce ca nanang de ine.*
 this mat REL plait GEN mother
 ‘This is the mat that was plaited by Mother.’

²² However, a close look at argument structures across languages, the following ranking is also attested: (a) Core Arguments outrank Non-core arguments, (b) within sets of cores/non-cores, prominence reflects semantic prominence (Manning 1996a, 1996b; Wechsler & Arka 1998; also Arka 1998, 2003).

Now, if the subject argument of the passive structure is reflexive, then there should be a problem in having *le* Agent bind the reflexive. This is because the reflexive would syntactically outrank its intended binder. We expect binding to fail. The expectation is borne out. Consider (25) below. The intended binding (i.e. reading (ii)) fails. The sentence is acceptable only in its first reading, where the Agent must be someone else. In reading (i), the reflexive form encodes emphasis only (i.e. an emphatic reading). In short, the failure of binding shown by reading (ii) is what is expected on the passive analysis of the *le* agent.

- (25) *Weki ru-n_i mbele le hia*_{i/j}.* <<self_i><hia_j/*_i>>(passive)
 body self-3s.GEN kill by 3s
 (i) ‘(S)he_i (himself/ herself) was killed by him_j /her_j.’
 (ii) *‘(S)he was killed by himself/herself.’

Further evidence from reflexive binding is illustrated by the active-passive alternation in examples (26)–(27). In (26a), the reflexive *wekin* precedes the Goal oblique *hi ase* ‘little sibling’ (marked by *kamping*), whereas in (26b) it follows the oblique. Notice that in both cases *wekin* can be bound only by subject (*hi ema* ‘father’, index *i*), not by oblique (*kamping hi ase*, index *j*). This suggests that linear precedence does not play any role in reflexive binding in Manggarai. What matters is syntactic prominence: the binder must be syntactically more prominent than its bindee. Thus, we expect that if the Goal *hi ase* is manifested as the grammatical subject by means of passivisation then binding is possible. This is indeed the case, as shown by (27), where the goal *hi ase* is the only possible binder of reflexive *wekin*. The fact that the passive Agent (i.e. PP Agent *li ema*) now fails to bind the theme *wekin* suggests that the Agent is now an oblique. That is, if the *le* Agent were a core argument (i.e. not oblique) it would be able to bind the (object/theme) reflexive *wekin*. Note that, as (24a) shows, when the Agent and the Theme are both core arguments, the Agent is a possible binder because it is thematically higher than the Theme/Patient (cf. Manning 1996a, 1996b).

- (26) a. *Hi ema_i toto weki-n_i/*_j one kaca kamping hi ase_j.*
 ART father show refl-3 at mirror to ART little.sibling
 b. *Hi ema_i toto kamping hi ase_j weki-n_i/*_j one kaca*
 ART father show to ART little.sibling refl-3 at mirror
 ‘Father showed himself in the mirror to (the/my) little sibling.’
- (27) *Hi ase_j toto weki-n_j/*_i li ema_i one kaca.*
 ART little.sibling show refl-3 by father at mirror
 ‘The/my little sibling was shown himself by Father in the mirror.’

However, one might reject the passive analysis on the ground that the *le* Agent oblique in many cases appears to be obligatory. This is unusual on the passive analysis since a passive Agent in other languages is generally optional. One would take it as evidence that the *le* Agent may be a core argument rather than an oblique.

We argue that such an analysis is untenable for the following reasons. First, in theory, there is no reason why an oblique should always be obligatory. Certain obliques may be obligatory for various reasons. For example, the locative oblique of the verb *put* in English is obligatory. This is a lexically specific constraint of the verb *put*.

Second, our study reveals that the requirement that in Manggarai the oblique Agent be obligatorily present in Manggarai is due to an independent, language-specific constraint associated with grammatical-relation marking of this language. In other languages, where passivisation is marked morphologically on the verb, the Agent oblique is indeed generally optional because verbal morphology has done the job of marking the passive structure. In contrast, in isolating languages like Manggarai, there is no such verbal marking strategy.²³ Simply reversing the word order would result in confusion of the linking (Agent vs Patient). Prepositional marking with *le* is the only means to encode the passive structure. Since it is the only marker, it is understandably obligatory; otherwise no passive structure would be recognised.

Third, related to the second point just mentioned, we expect that, when the verb has received certain marking to encode a passive structure, which is by means of cliticisation in Manggarai, the *le* agent is optional.²⁴ This is confirmed. Consider (28):

- (28) a. *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).’)
- b. *Nia=s na’a=d bao surak situ (le hau).*
 where=3p place=POSS just.now letter that.p (by 2s)
 ‘Where were the letters placed just now (by you)?’ (lit. ‘Where were they, the placing of the letters just now (by you)?’)
- c. *Toem tiba=n tegi d=ite (l=ise).*
 not.exist accept-3GEN demand POSS=1p.i by=3p
 ‘Our demand was not accepted (by them).’ (lit. ‘The acceptance of our demand (by them) did not exist.’)

In (28a) and (28b), the verbs (*emi*, *na’a*) are nominalised by the POSS marker =*d*. The head predicates of these sentences are *poli* and *nia*, which host the subject enclitic =*s* agreeing with the grammatical subject *surak situ*. Crucially, the demoted Agent is not obligatorily present. (This is indicated by putting the Agent within brackets.) If it is present, it is marked by the oblique marker *l(e)*. In (28c), the verb is nominalised by the inflected pronoun =*n*, and again the Agent — which must be marked by *l=* (if present) — is optional.

Fourth, although uncommon, Agentless passive structures are indeed attested in texts. This kind of passive is typically used when the Patient (of a semantically transitive verb) is maintained as the topic along a stretch of clauses. The agent is either well understood from the context or is deleted because it is clear or considered unimportant in a given context. Consider the underlined clause in the following quotation:

²³ Verheijen (1977) concludes that there appears to be no clear evidence that Manggarai was historically a suffixing language. (However, given Manggarai is an Austronesian language, some earlier stage must have had affixation.)

²⁴ However, it should be noted that the cliticisation does not have to be passivisation (cf. examples (12)–(14)).

- (29) *Sai one puar, ita le ghia bakuk ghang d-anak-n*
 arrive in forest see by 3s basket rice POSS-child-3.GEN
gho ata ngai weo kin one lobo ghaju (Rego)
 this REL PROG hang still in tip wood
 ‘Arriving in the forest, he saw the rice basket of his child (still) hanging
 on a tip of a branch (of a tree). (lit. ‘(He) arrived in the forest, (and) the
 rice basket of his child which was still being hung on a tip of a tree was
 seen by him.’)

In this example, the semantically transitive verb *weo* ‘hang’ is used intransitively, i.e. it is used in a passive structure without an Agent. Recall that only subject can be relativised in Manggarai. This excludes the analysis that the relativised Patient of the verb *weo* is object. In other words, the underlined relative clause cannot be understood as an active clause with its Agent subject missing, rather as a passive clause with its subject Patient controlled and its Agent unexpressed.

Finally, the idea that the *le* agent is an oblique gets support from its grammatical category expression. That is, obliques in Manggarai are always marked by prepositions. Core arguments (i.e. subject and object) are always bare NPs, not marked by a preposition. The contrast in (30) shows that a Goal/Locative oblique must be PP (30a), otherwise the sentence is not acceptable (30b). Then, when the Goal/Locative argument is promoted to subject (31a), in which case it is a core argument, the promotion to core status necessarily requires a change in marking. That is, the Goal argument must be NP; otherwise the structure is bad (31b). Also, in this passive construction, the passive Agent must be PP because it is an oblique; it cannot be NP (hence the unacceptability of (31c)).

- (30) a. *Aku puci ngger one lo’ang.*
 1s enter to in room
 ‘I entered (into) the room.’
 b. * *Aku puci lo’ang.*
 1s enter room
 ‘I entered the room.’
- (31) a. *Lo’ang hitu puci le ata tako.*
 room that enter by person steal
 ‘The room was entered by a thief.’
 b. * *Ngger one lo’ang hitu puci le ata tako.*
 to in room that enter by person steal
 ‘*To the room was entered by a thief.’
 c. * *Lo’ang hitu puci ata tako.*
 room that enter person steal
 ‘The room was entered by a thief.’

In short, the basic function of the prepositions, e.g. *le* (marking the Agent), *ngger one* ‘into’ (marking Goal) — is to mark non-core status. Core arguments do not get preposition marking. Therefore, we rule out the analysis which treats the PP (*le*) Agent as a core argument. The *le* Agent should be analysed as an oblique.

Finally, the claim regarding *le* as a passive marker should be slightly modified. This claim needs to be revised given the fact that Manggarai *le* is not uniquely associated with the passive construction. It also has locative, goal and instrumental function in Manggarai, reinforcing the oblique nature of *le*, e.g. *le puar* ‘in the forest’, *le pasar* ‘to the market’, and *le wase nol* ‘with a (piece of) rope (called) *nol*’.²⁵ Thus, a better or more correct claim is to say that *le* is simply an oblique marker.²⁶

To summarise, there is evidence to support the following ideas. First, backgrounding/demotion of the Agent argument is also accompanied by foregrounding/promotion of the Patient argument. Second, the backgrounded *le* Agent is not subject and the foregrounded Patient is subject. Evidence comes from agreement (examples (15)–(16), structural positions (17)–(19), and control (20)–(23). Third, the backgrounded *le* Agent should be analysed as an oblique argument for the following reasons:

- The *le* Agent, unlike the bare Agent NP, may appear in different positions (examples (17)–(19). This suggests that a PP Agent is structurally an Adjunct-like unit, hence a non-core or oblique, because subject (or a core argument) has a fixed position.
- The *le* Agent cannot bind a core Theme/Patient reflexive argument, which it thematically outranks (examples (24b) and (27)). This is not expected on the analysis that the *le* Agent is a core argument because if the Agent and the Theme/Patient are both core arguments, the Agent must be able to bind the Theme.
- The *le* Agent may be optional or unexpressed in certain circumstances (example (28)).
- In Manggarai, *le* is a preposition marking an oblique/adjunct; a core argument is never realised by a PP in this language (cf. examples. (30)–(31)).

5 Discussion

Typologically speaking, passive is a category of voice that is generally associated with verbal forms having certain passive morphology. That is, a particular affix on the verb is identified as a passive morpheme. Let us call this type of passive the ‘morphological’ passive. Haspelmath (1990) adopts a strong view claiming that the verbal morphology associated with the passive construction is an essential part of the passive. He claims that ‘passive constructions without passive morphology do not exist’ (1990:27).²⁷ (In his view,

²⁵ The following example illustrates the instrumental function of *le*:

Kempet le haju=n.
clamp with wood=3
‘It’s clamped with a piece of wood.’

²⁶ *Le* has been so far roughly glossed with ‘by’ for ease of the exposition of passives. It should perhaps be better glossed as either OBL (for Oblique), Pcase (i.e. prepositional case (marker)), or simply LE (i.e. as a marker). This avoids the necessity of determining what its exact case marking function is, or what its word class is, matters which are crucial to a full accounting of the syntax of the language. The fact that it can alternate with *de* in some contexts is an interesting case that deserves further studies.

²⁷ He makes his claim slightly weaker in the summary of the paper by saying that they are ‘rare or non-existent’ (1990:62).

other morphological devices used in passive constructions such as (prepositional) agent markers which are not specific to passive constructions are not considered ‘passive morphology’. We shall see below that Manggarai poses a challenge to his view.) He adopts the following definition of passive constructions:

- (32) A construction is called passive if:
- i. the active subject corresponds either to a non-obligatory oblique phrase or to nothing; and
 - ii. the active direct object (if any) corresponds to the subject of the passive; and
 - iii. the construction is somehow restricted vis-à-vis another unrestricted construction (the active), e.g. less frequent, functionally specialised, not fully productive.

Based on a genetically stratified sample of 80 languages (the Gramcats sample), only 31 were found to have a passive. Haspelmath notes that there are at least six expression types of morphological passives, as shown by Table 2. Among them, the ‘stem affix’ passive, i.e. the passive with its passive affix being attached directly to the verb stem, inside inflectional affixes (i.e. aspect, tense, and person markers), is by far the most common type. The English type where the passive is expressed by an auxiliary plus a participle verb is not a typical one.

Table 2: Expression types of passive morphemes in the Gramcats sample (Haspelmath 1990)

Expression types	Number of languages
additional stem affix	25
auxiliary verb (+participle)	6
extra-inflectional affix	3
differential-subject person markers	2
alternate stem affix	1
particle	1

We will not discuss and exemplify each of the types in this paper. Rather, we want to discuss the typology of passives in relation to the Manggarai data, taking into account Haspelmath’s claim that there are no such things as ‘passives without passive morphology’ because, according to him, the passive is essentially a verbal morphological category (diachronically) motivated by ‘inactivation’ of the verbal situation.²⁸ He further claims that inactivisation is the original function of the passive, not the participant backgrounding or foregrounding. (The latter is a consequence of inactivisation.) The evidence for this, according to him, comes from the grammaticalisation path, where the case of inactive auxiliaries is probably the most common source of passive morphology across languages; for example, just to mention some languages outside Austronesian languages here, Korean

²⁸ By this, he means, ‘non-agentive’ (i.e. ‘the opposite of active, actional’). All state verbs are ‘stative’; hence they are typically ‘inactive’. However, inactive verbs may express ‘dynamic’ situations. For example, *Maria was kissed by Juan* is as dynamic as the corresponding active. Inactive verbs are typically intransitive but they can be transitive (e.g. ‘undergo’ and ‘suffer’).

passive affix *-ji* (from (<) the verb *ji-* ‘fall’), Tamil *-paṭ* (< the verb *paṭu* ‘fall, happen’), Equadorian Quechua *-ri* (< the verb *ri-* ‘go’), Turkic *-il/-ıl* (< the verb *ol-* ‘be’), and Japanese *-ar(e)* (< the verb *aru* ‘be’) (see Haspelmath 1990 for more examples and further details).

Morphological passives in Austronesian languages of Indonesia are typically found in the Indonesian-type languages dominating the western part of Indonesia. Examples (33)–(39) show some Indonesian languages that have passive morphology (Indonesian, Balinese, Makasarese, Javanese, Menó-Meneé Sasak, and Bima). Some of them may have more than one passive affix. Indonesian *di-* and *ter-* (33) and Javanese *di-* and *ke-* (36) differ in volitionality of the event, e.g. *ter-/ke-* implying an accidental event. Balinese *ka-* and *-a* (34) differ in register, with *ka-* being used in high register (see Arka 1998 and Arka to appear, for further details). *Di-* and *ra-* in Bima (39) differ in mood (realis vs irrealis).

- (33) a. *Anak itu di-gigit (oleh anjing).* (Indonesian)
 child that PASS-bite by dog
 ‘The child was bitten by a dog.’
 b. *Mereka ter-tabrak ((oleh) mobil).*
 3p PASS-hit by car
 ‘They (accidentally) got hit by a car.’
- (34) a. *Ipun ka-icen jinah antuk Bapak Guru.* (Balinese)
 3 PASS-give money by father teacher
 ‘(S)he was given money by the teacher.’
 b. *Nyoman beli-ang-a nasi teken I Meme.*
 NAME buy-APPL-PASS rice by ART mother
 ‘Rice was bought for Nyoman by Mother.’
- (35) *Ni-jakkalak-i ri pulisi.* (Makasarese)
 PASS-arrest-3 by police
 ‘He was arrested by the police.’ (Manyambeang, Mulya, and Nasruddin 1996)
- (36) a. *Klambi-ne di-kumbah aku/kowe/Siti.* (Javanese)
 shirt-DEF PASS-wash 1s/2s/Name
 ‘The shirt was washed by me/you/Siti.’ (Sawardi 2001)
 b. *Klambi-ne ke-kumbah aku/kowe/Siti.*
 shirt-DEF PASS-wash 1s/2s/Name
 ‘The shirt was accidentally washed by me/you/Siti.’
- (37) *Aku te-gitaq isiq Ali.* (Menoó-Mené Sasak)
 1s PASS-see by NAME
 ‘I was seen by Ali.’ (Austin 2002)
- (38) *‘U-to-kiki’i na iko’o* (Tukang Besi)
 2s.R-PASS-bit NOM you
 ‘You were bitten.’ (Donohue 2002)
- (39) a. *Sia ra-ha’a ba ngao ede.* (Bima)
 3s PASS.REAL-bit by cat that
 ‘(S)he has been bitten by the cat.’ (Jauhary 2000)

- b. *Wela ede di-weli ba La Amir.*
kite that PASS.IRR-buy by ART NAME
'The kite will be bought by Amir.'

While some Austronesian languages outside Flores show morphological passives, other Austronesian languages in Flores do not even have passives. Although we talk about 'passive' in Manggarai, Manggarai does not have a typical passive because the head verb bears no verbal morphology.

We have provided evidence for the analysis that the construction with the (backgrounded) *le* agent is passive even though the verb form of this construction morphologically has the same form as the active counterpart. Indeed, the *le* Agent construction in Manggarai also passes Haspelmath's passive criteria formulated in (32). The first criterion — optionality of an oblique — is satisfied because the *le* agent is indeed omissible²⁹ in certain limited contexts; see discussion in §4.2, examples (28) and (29). The second criterion—the passive subject corresponds to the active object — is also satisfied. The Patient of the *le* (passive) construction is indeed grammatically subject. The evidence comes from a number of tests (agreement, structural positions, reflexive binding and control) discussed in §4.1. As for the third criterion — markedness and functionality — we can say for certain that the passive is pragmatically marked (i.e. functionally specialised). The distribution of the passive in texts is slightly less frequent than the corresponding active construction in a small sample of texts that we have looked at.³⁰ As seen in Table 3, the use of active and passive constructions vary across texts (intransitive

²⁹ Passives with obligatory agents are reported to exist (Siewierska 1984:35). Indonesian is said to have this kind of passive based on the analysis of Chung (1976). This is the Indonesian non-agentive construction with pronominal clitic on the verb, e.g. *buku itu ku=baca* 'book that 1s=read'. However, current research (Arka & Manning 1998) suggests that this type of structure is not really syntactically passive because the verb is syntactically transitive with *buku itu* 'the book' being the subject argument and the Agent clitic *ku=* being a core non-subject argument (i.e. not an oblique). In our definition of (syntactic) passive, the syntactic shift of the Agent and Patient of the transitive verb is crucial. This includes the most important criterion where the Patient becomes subject and the Agent loses its subject status. Typically, the Agent becomes an (optional) oblique but the Agent can also be totally suppressed from the argument structure of the derived passive verb (i.e. the Agent cannot show up). Some languages, however, allow a construction where the Agent is not subject but it is still a core argument. This includes, for example, the inverse voice found in Algonquian, a language in north America (Mithun 2001:222–228) and the non-agentive voices found in Indonesian languages such as Balinese (Arka 1998, 2003) and the Philippine languages such as Tagalog (Kroeger 1993). There is good evidence to consider that this type of construction is not passive for syntactic and pragmatic reasons, e.g. the Agent is still very much prominent syntactically and pragmatically (see, among others, Arka and Manning (1998), Foley (1998), Mithun (2001:225) and Purwo (1989)). The Indonesian *buku itu ku=baca* 'book that 1s=read' is then more like an inverse rather than a passive. This Indonesian construction is different from the passive construction in Manggarai in at least two respects. First, the Agent clitic *ku=* in Indonesian is always obligatory whereas the *le* Agent in Manggarai may be optional in certain circumstances. And second, crucially, there is evidence that the *le* Agent, unlike the Indonesian Agent clitic *ku=*, is not core.

³⁰ It should be noted that text 4 *Sungke Cahap* appears to show a slightly greater number of passives (16) to actives (12). A closer look at this story, however, reveals that it is about a variety of social and physical problems suffered by a person who has a bad habit of talking while sleeping. The focus of the story is 'Patient-oriented'; i.e. how this person undergoes a variety of bad experiences with other people. The high proportion of passive verbs is expected. We believe that the genre has a distinct effect on the type of structures used as text 4 shows, which may be inconsistent with the generalisation that actives are more common than passives across text types.

verbs such as ‘come’ and ‘sit’ are excluded). Out of a total of 94 transitive verbs, around two thirds of them (62 verbs, or 66%) appear in active constructions whereas 26 are in passives with agents and only 6 in agentless passives.³¹

There seems to be evidence from Austronesian languages of Indonesia to support Haspelmath’s observation that an inactive verb may get grammaticalised (and morphologised) to become a passive marker. For example, the ‘informal’ Indonesian passive *ke-* could have come from the verb *kena* ‘be adversely hit by, suffer’.³² Both of them (i.e. sentences (40a) and (40b) are used interchangeably nowadays. In standard/formal Indonesian, the *di-* passive (40c) is used, however.³³

Table 3: The frequency of actives and passives in some Manggarai texts

	Text titles	Actives	Passive with agents	Passive without agents	Total
1	<i>Wendong ata manusia le darat</i> ‘The abduction of a man by a spirit’	19	6	1	
2	<i>Tombo tora mangan rana mese</i> ‘The story of the lake Rana Mese’	11	8	–	
3	<i>Sungke cahap</i> ‘The story about talking while sleeping’	12	11	5	
4	<i>Tombo Ka agu Kode</i> ‘The crow and the monkey’	20	1	–	
	Total	62	26	6	94

- (40) a. *Ia kena tipu.*
3s hit cheat
‘(S)he was/got cheated.’
- b. *Ia ke-tipu.*
3s PASS-cheat
‘(S)he was/got cheated.’
- c. *Ia di-tipu.*
3s PASS-cheat
‘(S)he was/got cheated.’

³¹ The figures given in Table 3 are based on limited texts. While they are indicative that the active construction is more common than the passive construction, further quantitative research based on a larger corpus that consists of a variety of text types in Manggarai is needed. This might give us a better understanding of the distribution of different voices. One of the crucial issues worth investigating is the distribution of the agentless passive. For example, the figure in Table 3 which indicates that 6 out 94 examples are of an agentless passive might suggest that there could be something more than simply a ‘performance’ fact.

³² Apparently, this ‘informal passive’ in Indonesian is related to Javanese *ke-*. It is unclear, however, whether the Javanese *ke-* also came from an inactive verb ‘suffer or hit’. Malcolm Ross (pers. comm), however, doubts this: the prefix *ke-* is probably derived from the PAN/PMP **ka-*.

³³ The historical origin of the passive prefix *di-* is a matter of debate. There are at least three competing hypotheses. It could come from (i) a captured (locative) preposition *di-*, or (ii) a third person pronoun *dia*, or (iii) the PAST affix *ni-* (see Ross (2004) for detailed discussion).

It is not clear at this stage whether Balinese passive *ka-* illustrated in (41) also came from an inactive verb.³⁴ Balinese *ka-*, unlike the informal Indonesian *ke-* or Javanese *ke-*, has no accidental or adversative meaning.

- (41) *Gumi-ne ka-prentah (antuk/teken bangsa gelah).*
 country-DEF PASS-govern by people own
 ‘The country is governed (by our own people).’

The historical origin of *le* in Manggarai is unclear at this stage. Particularly, we do not know for sure whether *le* used to be a verb, but it most probably never was. The apparently related (cognate?) words in other western Austronesian languages are all prepositions (e.g. Indonesian *oleh*, Balinese *olih* and Acehnese *lé*). In contemporary Manggarai, *le* is a preposition. Indeed, Verheijen’s (1967) dictionary lists it only as a preposition. No other entries of the same form are listed as verbs of the ‘inactive’ type meaning ‘suffer’ and the like. (Recall that Haspelmath (1990) claims that the common historical path of the passive marker is from an inactive verb.)

If Manggarai *le* was never a verb, then Manggarai is indeed a language that has a passive without passive morphology, contrary to Haspelmath’s generalisation, because *le* is not a unique marker of the passive construction. As noted earlier in §4.2, the preposition *le* is also used to encode locative, goal, and instrument. It is worth discussing other languages that have been cited to have a Passive without passive morphology (which potentially challenge Haspelmath’s claim). In what follows, we briefly discuss Chinese and Acehnese cases.³⁵

Like Manggarai, Chinese is an isolating language. Like Manggarai, it employs the same verb form (i.e. without any additional affix) in both active and passive constructions.³⁶ For example, the verb form *zhemo* appears in the active sentence (42a) as well as in the passive counterpart (42b). The passive construction is marked by (what is believed now to be) the preposition *bei*. *Bei* may appear without its Agent complement NP as shown by example (42c).

- (42) a. *Zhou Hua zhemo Gao Qiang.* (Shi 1997)
 NAME torment NAME
 ‘Zhou Hua tormented Gao Qiang.’

³⁴ Balinese *ka-* was possibly borrowed from (Old) Javanese (Clynes 1989), and the shift in meaning might have taken place in Old Javanese.

³⁵ Also relevant here is evidence from creole languages such as the pair below from Jamaican Creole (Bernard Comrie, pers. comm.):

- a. *Dem sel-aaf di bota.* ‘They have sold all the butter.’
 b. *Di bota sel-aaf.* ‘The butter has all been sold.’

As with Manggarai, the virtual lack of inflectional morphology seems to be a relevant factor. These examples show a syntactic shift (of the patient *di bota* ‘the butter’) that is not coded by the verbal morphology (i.e. the verb in (b) has no passive morpheme). A detailed comparison of Manggarai with creole languages is beyond the scope of this paper.

³⁶ It should be noted that the functions performed by the passive are more usually performed in Chinese by topicalising the patient of an otherwise default clause. Such a structure may appear with or without an agent. I thank Malcolm Ross for pointing this out to me. In any case, however, the verb has the same form.

- b. *Gao Qiang bei Zhou Hua zhemo.*
 NAME BEI NAME torment
 ‘Gao Qiang was tormented by Zhou Hua.’
- c. *Wo bei liyongle.*
 I BEI use.ASP
 ‘I was used.’

Haspelmath (1990) argues that the Chinese *bei* passive is not a counterexample to his claim because there is (historical) evidence, discussed in (Bennett 1981; Hashimoto 1988; Zhang 1990), that *bei* was not a preposition, but a verb.³⁷ In Old Chinese, it was a verb meaning ‘to receive’ (Bennett 1981), ‘to sustain’ (Hashimoto 1988) or ‘to cover’ (which later changed to ‘suffer’) (Zhang 1990). The double functions of *bei* (as a preposition and a verb-like/passive marker) have led to competing analyses of the exact status of this morpheme as reviewed in Shi (1997). In short, the data that seems to be against Haspelmath’s claim turns out, from a historical perspective, to support his analysis that a passive marker originates from an ‘inactive’ verb.

Turning to the Acehnese case, we have a different situation. Consider the examples in (43),³⁸ which show that the verb form *geu-côm* ‘3_p-kiss’ in the alleged passive sentence (43b) also employs the same verb form as the active (43a). Sentence (43b) is claimed to be passive by Lawler (1977) but is disputed by Durie (1988). Durie (1988) argues that the so-called passive in Acehnese illustrated by (43b) is in fact not passive, but rather a word-order variant of the active structure (43a) (i.e. sentence (43b) is a kind of Patient preposing). The controversy over the analysis boils down to whether the notion of ‘Subject’ (on which the notion of ‘Passive’ is based) can be proved to be relevant in Acehnese.³⁹ In particular, the validity of the passive analysis of sentence (43b) very much depends on the evidence that the sentence-initial NP (in this case *lôn*) is grammatically ‘Subject’. Durie (1988) argues that it is not. In fact, Durie (1987) argues that no surface grammatical relations of subject and object can be identified in Acehnese. Lawler (1988) in his reply to Durie (1988) points out that such notions as ‘subject’ (and also ‘Passive’) are appropriate for Acehnese, at least according to a more recent grammar of Acehnese by Asyik (1987),⁴⁰ a native speaker of the language.

- (43) a. *Gopnyan ka geu-côm lôn.*
 she_p IN 3-kiss 1_p
 ‘She kissed me.’
- b. *Lôn ka geu-côm lé-gopnyan.*
 1_p IN 3-kiss - she_p
 ‘I was kissed by her.’

³⁷ Evidence that it is not an ordinary preposition comes from the fact that it does not behave like a preposition in modern Chinese; e.g. *bei* can be left stranded as in (42c) whereas a real preposition cannot (Shi 1997:46; Hashimoto 1988).

³⁸ The subscript ‘p’ means ‘polite’ and IN means ‘INCHOATIVE’.

³⁹ However, one could embrace a framework or analysis that does not make use of subject and object metalanguage but can still talk about passives (e.g. Foley & Valin 1984; Shibatani 1985).

⁴⁰ We have not been able to consult this work.

It should be noted that Acehnese, unlike Manggarai and Chinese, is not strictly speaking an isolating language. Acehnese crucially differs from these two languages in having an agreement system marked on the verb that is sensitive to the verbal semantics or argument-roles of the verb. For example, the bound pronoun *geu-* in (43) agrees with the Agent argument; hence it agrees with *Gopnyan*, irrespective of whether the free NP Agent appears before the verb as in the active sentence (43a) or after the verb as in the alleged passive structure (43b), where the free NP agent must be marked by a preposition-like marker *lé*. (This marker is not glossed in (43b) and poses a potential problem in Durie's analysis⁴¹). The agreement is particularly clear in the case of intransitive verbs where a split is recognised: an action-like intransitive verb such as 'run' gets agreement like the transitive Agent whereas a Patient-like intransitive verb like 'fall' does not get this agreement (see Durie 1987 for details).

Recall that Manggarai also shows (en)clitic agreement. However, the enclitic agreement in Manggarai significantly differs from the agreement in Acehnese in that Manggarai clitic agreement is indeed sensitive to surface grammatical relations: it is associated with subject, which is therefore not restricted to the Agent role. This has been discussed in subsections (4.1 and 4.2) to prove that there is passivisation in Manggarai (i.e. object–subject alternation of the Patient of the transitive verb).

Another difference between Manggarai and Acehnese relates to the grammatical status of the Agent argument of the semantically transitive verb. We have argued at length that the Agent marked by *le* in Manggarai is oblique. This is not easily resolved in Acehnese due to the presence of the pronominal *geu-* on the verb. A bound pronoun that precedes a verb (or occurs morphologically on the verb) and receives cross-reference with a floating PP (typically after the verb) is in fact not unique to Acehnese. Balinese and Sasak are other such languages. This type of structure poses a problem in the analysis of, for example, Balinese voice. However, current research on Balinese (particularly with evidence from binding and other properties as discussed in Arka (1998, to appear)) has led to progress in the understanding of such a structure. For this reason, we briefly discuss the Balinese examples and then turn to Sasak.

In Balinese, there is evidence for grammaticalisation of the bound pronoun as a passive marker. Consider examples in (44) where the third person bound pronoun *-a* can be made explicit by a following PP *teken Wayan* as in (44c). In contemporary Balinese, this is unambiguously a syntactic passive, as argued in (Arka 1998, to appear). However, the PP Agent in Balinese, like the one in Acehnese, can be absent as in (44b). This is an ambiguous structure. On one hand, the (bound) pronoun *=a* can be understood as an argument of the verb, in which case the verb is associated with Objective Voice (OV). This kind of voice, typical in western Austronesian voice systems, reflects an Undergoer-oriented voice (where the Patient is subject) but, unlike passive voice, the Agent is still a core argument. On the other hand, it may be simply a variant of passive (44c) with the Agent being unexpressed. Arka (1998; to appear) argues that in the presence of the PP Agent (i.e. of the type (44c), the construction is indeed syntactically passive because the evidence overwhelmingly shows that the Agent is an oblique. In short, in contemporary Balinese, the bound pronoun lives a double life (as a bound pronoun and as a 'passive' marker). The evidence for the analysis of different voices in Balinese is much clearer than

⁴¹ He analyses it as an 'ergative' marker, which seems to add more controversy, rather than clarification, to the Acehnese analysis.

that in Acehnese thanks to the verbal voice morphology in Balinese, particularly the marking of the Active Voice (AV) by a nasal prefix. That is, we are absolutely sure that the Patient argument in (44b) is grammatically subject. A similar claim with absolute certainty apparently cannot be made for the Acehnese data in (43b).

- (44) a. *Wayan maang Nyoman pipis.*
 NAME AV.give NAME money
 'Wayan gave Nyoman money.'
- b. *Nyoman baang=a pipis.*
 NAME OV.give=3 money
 '(S)he gave NYOMAN (not someone else) money.'
- c. *Nyoman baang-a pipis (teken Wayan).*
 NAME give-PASS money by Wayan
 'Nyoman was given money (by Wayan).'

A similar example from Menó-Mené Sasak is given in (45) (Austin 2002), where bound clitic =*n* is cross-referenced by PP *isiq Herman*. While Austin explicitly claims that this type of sentence is transitive (not passive), he is not explicit in saying whether it is 'active' or 'agentive' voice. (Note that the Balinese OV verb illustrated in (44b) is also syntactically transitive but it is not 'agentive/actor-oriented'.) This is presumably because Menó-Mené Sasak does not distinguish grammatical relations other than core versus oblique (Austin 2002, fn. 6). In this case, then, the PP Agent construction in Menó-Mené Sasak is closer to that in Achenese than to that in Balinese. Unlike Acehnese, however, Menó-Mené Sasak clearly has a passive construction with verbal passive morphology (*te-*, example (37) repeated here as (45b)). Note that the same preposition (*isiq*) is used to mark the Agent NP.

- (45) a. *Yaq=n gitaq kanak-kanak=no isiq Herman.* (Austin 2002)
 fut=3 see child-child=tjat by NAME
 'Herman will see the children.'
- b. *Aku te-gitaq isiq Ali.*
 1s PASS-see by NAME
 'I was seen by Ali.'

The points that we want to highlight with the data from other (Austronesian) languages of Indonesia are these. First, the analysis of a passive without passive morphology in other languages such as Acehnese, for example, is hard to demonstrate because the presence of the bound pronominal Agent clitic or prefix on the verb may indicate that the verb is still transitive (i.e. no change in transitivity as is the case if we have passivisation). Other processes in Acehnese such as reflexive binding appear to be sensitive to a semantic role hierarchy (Durie 1987), so they do not help much. In contrast, the situation in Manggarai is much clearer because there are language-specific properties such as the agreement clitic and reflexive binding that help us to show that the grammatical relation of the bare Agent NP is different from that of the PP (*le*) Agent. Secondly, an apparently similar (pronominal) bound form that might be encountered in other constructions might have been grammaticalised to become a different voice marker. In other words, on the basis of the same morphology (e.g. Balinese bound form *-a*), we cannot jump to the conclusion that the form (and the verb/construction) is associated with a single voice. It may be

ambiguous between two different voices, Objective Voice (OV) and Passive voice. In Manggarai, we do not have this ambiguity; we can claim with certainty that we have a clear distinction between active and passive constructions.

Finally, a close association of the preposition that marks an Agent (e.g. *le* in Manggarai) with the passive construction needs a brief comment here. The preposition would not be conventionally regarded as a passive marker. Thus, the Indonesian *oleh* ‘by’ or the Balinese *antuk/teken* ‘by’ is the preposition that marks the oblique Agent but is not considered a passive marker because the verb has passive morphology, *di-* (Indonesian) and *ka-* (Balinese). These verbal affixes are the passive markers because they indicate that the clauses are passives. (The PP agents with *oleh/teken* may be absent.) However, the preposition *le* in Manggarai is the only indication by which a passive construction is recognised. In isolating languages, due to the lack of passive morphology on the verb, any ‘marker’ (outside the verb) that indicates a passive understandably serves as a significant cue. This is confirmed by research from language processing (in other languages of a similar type). For example, Li, Bates, and MacWhinney (1993) reports that, while there is evidence for the significance of animacy in Chinese speakers’ use of different cues in sentence processing, *bei* is a powerful cue, helping the speakers select the second noun more than any other cue. Indeed, even if the prepositional marker for oblique is not generally considered a passive marker (e.g. in languages like English), it turns out to be a crucial cue. A report on language comprehension and production of the passive voice among ten severely prelingually deaf boys and girls (Power & Quigley 1973) indicates that more than half of them correctly understood passive sentences and less than half correctly produced such sentences. Crucially, they interpreted passive sentences in terms of the surface word order and the preposition *by* was the only passive marker for most of these children. It appears that much emphasis has been placed on the use verbal morphology to identify the (so-called) Passive in the linguistic literature on voice but verbal voice morphology might not be as significant or powerful a marker as we initially thought.⁴²

6 Conclusion

Manggarai is an instance of a language that has a passive construction without passive verbal morphology. This passive construction appears with *le*, signalling the oblique status of the Agent argument of the (semantically transitive) verb. We have presented evidence to support the analysis that the *le* Agent construction passes typical tests for passive. However, since *le* is not unique to the passive (i.e. it is also used to mark locatives, instrumental, and goals), it cannot be claimed that it *is* the passive marker in Manggarai. It is simply a prepositional ‘case’ of non-core status, signalling a passive construction. (It is also used in other construction types.)

Typologically speaking, it is not really surprising to have a passive without passive morphology. If one looks at how passives get expressed cross-linguistically, as Haspelmath points out, they are of numerous types (morphological, analytical, or a combination of these) with the voice morphology, if any, not originally used to mark passives. An earlier

⁴² It appears that the ‘whole construction’ is recognised as passive (which is also true with the deaf children). As far as the Manggarai data is concerned, the Agent *le* may be missing and the context could play a role in the recognition of the construction as passive.

study by Andersen (1989) also reveals the same conclusion and he further concludes that 'the existence of passive morphology is neither necessary nor sufficient for the definition of the passive'. In other words, passive morphology is not central to the characterisation of the passive construction.

The question then is what characterises passives cross-linguistically. Siewierska (1984:39–40) lists a number of properties that distinguish passives from actives (the order of NPs, special verbal morphology, etc.) but further suggests that none of them can be used for defining passives in all passive constructions. One way of characterising passives cross-linguistically is to abstract away from overt or surface expressions by referring to certain (abstract) relations such as 'subject' and 'object', which is in fact a practice that has been a long tradition in linguistics.⁴³ Such abstract relations are taken up again in contemporary syntactic theories, notably Relational Grammar (Perlmutter & Postal 1977) and Lexical Functional Grammar (e.g. Bresnan (1982)). An attempt to formulate the universal characterisation within Relational Grammar is given in Perlmutter and Postal (1977), which is contested in O'Grady (1980) but is defended in Dryer (1982).⁴⁴ (Note that the characterisation of passives from Haspelmath cited in (32) makes use of the notions of 'subject' and 'object' as well.) The advantage of using abstract relations such as subject and object is that it allows us to draw generalisations without saying anything about passive morphology (or other overt markings).

However, a morphological characterisation of passives seems to be assumed in at least two current theories. This seems to be a drawback because it leads to a problem. In LFG (Alsina and Mchombo 1990; Alsina and Mchombo 1993; Bresnan and Kanerva 1989; Bresnan 2001; Dalrymple 2001, among others), passive alternations are dealt with under alternative linking within LMT (Lexical Mapping Theory), in which certain morphology on the verb (i.e. a passive affix) signifies a marked linking where the Agent is either totally suppressed from the argument structure of the verb or prevented from being mapped onto subject. The passive in Manggarai would pose a problem to this conception of passivisation as a morpholexical process. In this view, passivisation necessarily involves morphological derivation (or affixation).⁴⁵ However, as we have seen, passivisation in Manggarai involves no such derivation. Likewise, it is arguably hard to deal with the Manggarai passive in the Case-theoretic Chomskyan transformational model of grammar (see for example Webelhuth 1995; Guilfoyle, Hung, and Travis 1992), where passivisation is also associated with morphological verbal marking because the basic assumption of this model is that passivisation involves movement of an NP (i.e. the Patient) motivated by Case (absorption). It is assumed that the passive affix is said to 'absorb' Case that would be otherwise assigned to the Patient by the transitive verb. As result, the Patient has to move to a position where it may receive Case (i.e. to the subject position). Again, we cannot appeal to such a Case-theoretic explanation to account for the passive in Manggarai

⁴³ Another way of doing it is to analyse different kinds of voice in terms of prototypes, in which pragmatic, semantic, syntactic and morphological properties are considered (see Shibatani 1985 for details.)

⁴⁴ As noted in Siewierska (1984), when we speak of a universal characterisation of passive we do not mean that the passive is a universal phenomenon, because it is obvious that many languages lack this voice category. For example, Austronesian languages of eastern Indonesia (e.g. Taba (Bowden 2001)) typically show no passives. Based on the Gramcats sample, Haspelmath (1990) remarks that 'it is more likely for a language to lack a passive than to have one'.

⁴⁵ Analysing passives within a version of LFG e.g. Andrews and Manning (1999), where passives can be viewed as complex predicates might solve the problem. We leave this for further research.

because there is no affix on the verb that can absorb Case in the first place.⁴⁶ Discussing the theoretical issues of Manggarai passive in detail is certainly beyond the scope of the present paper.

References

- Alsina, Alex and Sam Mchombo, 1990, The syntax of applicatives in Chichewa: problems for a theta theoretic asymmetry. *Natural Language and Linguistic Theory* 8:543–505.
- 1993, Object asymmetries and the Chichewa applicative construction. In S. Mchombo, Stanford, ed. *Theoretical aspects of Bantu grammar*, 17–45. CSLI Publications.
- Andersen, Paul Kent, 1989, Is there a passive morphology? (Gibt es passivmorphologie?). *Linguistische Berichte* 121:185–205.
- Andrews, Avery and Christopher Manning, 1999, *Complex predicate and information spreading in LFG*. Stanford: CSLI.
- Arka, I Wayan, 1998, From morphosyntax to pragmatics in Balinese. PhD dissertation, University of Sydney, Sydney.
- 2003, *Balinese morphosyntax: a lexical-functional approach*. Canberra: Pacific Linguistics.
- to appear a, Aspects of core pronominal enclitics in Manggarai. In G. Reesink and B.K. Purwo, eds *Thinking and speaking, in memory of John Verhaar*. Amsterdam: John Benjamins.
- to appear b, Voice and the syntax of =a/-a verbs in Balinese. In P.K. Austin and S. Musgrave, eds *Voice and grammatical relations in Austronesian languages*. Stanford: CSLI.

⁴⁶ The idea of NP movement within this Case-theoretic passive analysis may have problems from a typological perspective with respect to cases: (i) where passivisation does not appear to affect the grammatical role of the patient nominals (e.g. the passive in Mayan languages) (Shibatani, ed. 1988:2–3; Shibatani 1985), or (ii) where passives do not seem to have active counterparts as in Indonesian (examples below, pointed to us by Yassir N. Tjung, pers. comm.) and in Nootka-Nitinaht (Rose & Carlson 1984).

- (i) *Uang-nya di-beli-kan mobil oleh Amir.* (passive)
 money-DEF PASS-buy-APPL car by NAME
 ‘The money was used to buy a car by Amir.’
- (ii)?? *Amir membeli-kan uang-nya mobil.*
 NAME AV.buy-APPL money-DEF car
 ‘Amir bought the car with the money.’

In a lexically based model such as LFG (within the linking theory) (Bresnan & Kanerva 1989; Bresnan 2001; Alsina & Mchombo 1990; Alsina & Mchombo 1993; Dalrymple 2001, among others), the two problems just mentioned can be easily dealt with because active and passive verbs may share argument structures. Crucially, however, there is no requirement that the Patient that is linked to subject (in the passive verb) must be always linked to object first. That is, there is no one-to-one relation between active and passive constructions. Arguments for the idea that passivisation are a lexically-based relational change (not a syntactic transformational process) is discussed, among others, in Bresnan (1978, 2001).

- Arka, I Wayan and Christopher Manning, 1998, Voice and grammatical relations in Indonesian: a new perspective. In Butt and King, eds *The Proceedings of the LFG '98 Conference*, Stanford: CSLI, <http://csli-publications.stanford.edu/LFG/3/lfg98-toc.html>.
- Asyik, Abdul Gani, 1987, A contextual grammar of Achenese sentences. Dissertation, University of Michigan.
- Austin, Peter, 2002, Clitics in Sasak, eastern Indonesia. Paper read at Paper presented at the Tenth National Congress of the Indonesian Linguistics Society Conference, Denpasar.
- Bennett, Paul A., 1981, The evolution of passive and disposal sentences. *Journal of Chinese Linguistics* 9:61–89.
- Blust, Robert, 1993, Central and central-eastern Malayo-Polynesian. *Oceanic Linguistics* 32 (2):241–293.
- Bowden, John, 2001, *Taba: description of a south Halmahera language*. Canberra: Pacific Linguistics.
- Bresnan, Joan, 1978, A realistic transformational grammar. In M. Halle, J. Bresnan and G.A. Miller, eds *Linguistic theory and psychological reality*, 1–59. Cambridge, MA: MIT Press.
- 2001, *Lexical functional syntax*. London: Blackwell.
- Bresnan, Joan, ed., 1982, *The mental representation of grammatical relations*. Cambridge, Massachusetts: the MIT Press.
- Bresnan, Joan and Jonni M. Kanerva, 1989, Locative inversion in Chichewa: a case study of factorization in grammar. *Linguistic Inquiry* 20:1–50.
- Chung, Sandra, 1976, On the subject of two passives in Indonesian. In Charles N. Li, ed. *Subject and topic*, 57–99. New York: Academic Press.
- Clynes, Adrian, 1989, Speech styles in Javanese and Balinese: a comparative study. MA thesis, The Australian National University, Canberra.
- Comrie, Bernard, 1978, Ergativity. In W.P. Lehman, ed. *Syntactic typology: studies in the phenomenology of language*, 329–394. Austin: University of Austin Press.
- Dalrymple, Mary, 2001, *Lexical functional grammar*. San Diego: Academic Press.
- Dixon, R.M.W., 1979, Ergativity. *Language* 55:59–138.
- 1994, *Ergativity*. Cambridge, England: Cambridge University Press.
- Donohue, Mark, 2002, Voice in *Tukang Besi* and the Austronesian focus system. In F. Wouk and M.D. Ross, eds *The history and typology of western Austronesian voice systems*, 81–99. Canberra: Pacific Linguistics.
- Dryer, Matthew S., 1982, In defense of a universal passive. *Linguistic Analysis* 10:53–60.
- Durie, Mark, 1987, Grammatical relations in Acehese. *Studies in Language* 11:365–399.
- 1988, The so-called passive of Acehese. *Language* 64:104–113.
- Foley, William A., 1998, Symmetrical voice systems and precategoriality in Philippine languages. Sydney: University of Sydney.
- Foley, William A. and Robert Van Valin, 1984, *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.

- Guilfoyle, Eithne, Henrietta Hung and Lisa Travis, 1992, SPEC of IP and SPEC of VP: two subjects in Austronesian languages. *Natural Language and Linguistic Theory* 10:375–414.
- Hashimoto, Mantaro J., 1988, The structure and typology of the Chinese passive construction. In Shibatani, ed. 1988:329–354.
- Haspelmath, Martin, 1990, The grammaticization of passive morphology. *Studies in Language* 14:25–72.
- Jauhary, Eddy, 2000, Pasif bahasa Bima. MA thesis, Udayana University, Denpasar.
- Keenan, Edward L. and Bernard Comrie, 1977, Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8:63–99.
- Kosmas, Jaladu, 2000, Argumen aktor dan pemetaannya. MA thesis, S2 Linguistik, Universitas Udayana, Denpasar.
- Kroeger, Paul, 1993, *Phrase structure and grammatical relations in Tagalog*. Stanford: CSLI Publications.
- Lawler, John M., 1977, A agrees with B in Achenese: a problem for relational grammar. In P. Cole and J.M. Sadock, eds *Grammatical relations (Syntax and semantics, 8)*, 219–48. New York: Academic Press.
- 1988, On the questions of Acehnese ‘passive’. *Language* 64:114–117.
- Li, Ping, Elizabeth Bates and Brian MacWhinney, 1993, Processing a language without inflection: a reaction time study of sentence interpretation in Chinese. *Journal of Memory and Language* 32:169–192.
- Manning, Christopher, 1996a, Argument structure as a locus for binding theory. In Miriam Butt and Tracy Holloway King, eds *Proceedings of the LFG '96 Conference, Rank Xerox Research Centre, Grenoble, France, August 26–28*. Stanford: CSLI (<http://csli-publications.stanford.edu/LFG/1/toc-lfg1.html>).
- 1996b, *Ergativity: argument structure and grammatical relations*. Stanford: CSLI.
- Manyambeang, A.K, Abdul Kadir Mulya and Nasruddin, 1996, *Tatabahasa Makassar*. Jakarta: Departemen Pendidikan dan Kebudayaan.
- Mithun, Marianne, 2001, *The languages of native north America*. Cambridge University Press: Cambridge.
- O’Grady, William D., 1980, The universal characterization of passivisation. *Linguistic Analysis* 6:393–405.
- Perlmutter, David M. and Paul M. Postal, 1977, Toward a universal characterization of passivisation. In K. Whistler, et al. eds *Proceedings of the Third Annual Meeting of the Berkeley Linguistics Society*, 394–417. California: University of California Berkeley.
- Power, D.J. and Stephen P. Quigley, 1973, Deaf children’s acquisition of the passive voice. *Journal of Speech and Hearing Research* 16:5–11.
- Purwo, Bambang Kaswanti, 1989, Voice in Indonesian : a discourse study. In B.K. Purwo, ed. *Serpih-serpih telaah pasif bahasa Indonesia*, 344–442. Yogyakarta: Kanisius.
- Rose, Suzanne M. and Barry F. Carlson, 1984, The Nootka-Nitinaht passive. *Anthropological Linguistics* 26:1–12.

- Ross, Malcolm D., 2004, Notes on the prehistory and internal subgrouping of Malayic. In J. Bowden and N.P. Himmelmann, eds *Papers in Austronesian subgrouping and dialectology*, 97–109. Canberra: Pacific Linguistics.
- Sawardi, F.X., 2001, Kesimetrisan objek bahasa Jawa. *Nuansa Indonesia* 16 (15):16–22.
- Semiun, Agustinus, 1993, The basic grammar of Manggarai: Kempo subdialect. MA thesis, La Trobe University, Melbourne.
- Shi, Dingxu, 1997, Issues on Chinese passive. *Journal of Chinese linguistics* 25:41–70.
- Shibatani, Masayoshi, 1985, Passives and related constructions: a prototype analysis. *Language* 61:821–848.
- Shibatani, Masayoshi, ed., 1988, *Passive and voice*. Amsterdam: John Benjamins.
- Siewierska, Anna, 1984, *The passive*. London: Croom Helm.
- Verheijen, J.A.J., 1967, *Kamus Manggarai I (Manggarai-Indonesia)*. 's-Gravenhage: Koninklijk Instituut voor Taal- Land- en Volkenkunde.
- 1977, The lack of formative in affixes in the Manggarai language. *NUSA* 5:35–37.
- 1991, *Manggarai dan Wujud Tertinggi*, Jakarta: LIPI-RUL.
- Webelhuth, Gert, ed., 1995, *Government and Binding theory and the Minimalist Program*. Oxford: Blackwell.
- Wechsler, Stephen and I Wayan Arka, 1998, Syntactic ergativity in Balinese: an argument structure based theory. *Natural Language and Linguistic Theory* 16:387–441.
- Zhang, Hongming, 1990, Metaphor or Function? Bleaching or Gaining? *Mid-America Linguistics Conference papers*: 446–460.
- Zwicky, Arnold M., 1985, Clitics and particles. *Language* 61:283–305.
- 1987, Suppressing the Zs. *Journal of Linguistics* 23:133–148.
- Zwicky, Arnold M. and Geoffrey K. Pullum, 1983, Cliticization vs. inflection. *Language* 59:502–513.

5 *Two passive-like constructions in Tongan*

YUKO OTSUKA

1 Introduction

Polynesian languages divide into two groups in terms of case marking: accusative and ergative. The former consists of Eastern Polynesian languages, while Tongic and many of the Samoic-Outlier languages are ergative.¹ It is also well known that the accusative languages in the Polynesian family have a passive construction, typically involving the passive suffix *-Cia*, whereas the ergative languages apparently lack such a construction. It has been shown in the literature that cognates of *-Cia* exist in Tongan, although they do not function as passive morphemes, nor is such affixation productive (Chung 1978; Tchekhoff 1973). However, some have claimed that there are constructions that are typically interpreted as passive, namely (a) VOS constructions and (b) agentless transitive constructions (Churchward 1953; Lynch 1972; Tchekhoff 1973). This paper re-examines these passive-like constructions in Tongan. It will be shown that agentless transitive and VOS constructions are not syntactically passive in that the ABS-marked argument in these constructions does not exhibit the properties of subjects² in Tongan. There are four syntactic phenomena that are associated solely with subjects: (a) alienable possessive pronouns, (b) clitic pronouns, (c) *mo* coordination, and (d) control. None of these applies to the ABS-marked arguments in VOS and agentless transitive constructions. The current study proposes that (a) VOS constructions are derived as a result of scrambling, and (b) agentless transitive is an instance of *pro*-drop, permitted in Tongan with some restrictions.

¹ Some do not agree on this generalisation. For instance, Gibson and Starosta (1990) argue that Maori, an Eastern Polynesian language, is ergative.

² By 'subject' I refer to A and S in Dixon's (1979) terms.

2 Passive in Polynesian

This section discusses the so-called passive construction in Polynesian. First, the three construction types found in the accusative languages (i.e. Eastern Polynesian) and those found in the ergative languages such as Tongan and Samoan will be compared. On the surface each of the three constructions found in the former has a corresponding construction in the latter. However, only Eastern Polynesian is said to have a passive construction. It will be shown that despite the superficial similarities, there are some significant differences between the passive construction in Eastern Polynesian and the active transitive construction in the ergative languages.

2.1 Three constructions in Polynesian

It is generally accepted that a passive exists in Eastern Polynesian. Eastern Polynesian has three constructions: intransitive, transitive, and passive, as illustrated below.³

(1) Maori (Biggs 1969, cited by Clark 1976:42, 67)

Intransitive	TNS V NP	<i>Ka haere te wahine ki te whare.</i> TNS go DEF woman to DEF house 'The woman goes to the house.'
Transitive	TNS V NP i/ki NP	<i>Ka inu te tangata i te wai.</i> TNS drink DEF man ACC DEF water 'The man drinks the water.'
Passive	TNS V-Cia NP e NP	<i>Ka inu-mia te wai e te tangata.</i> TNS drink-Cia DEF water AGT DEF man 'The water is drunk by the man.'

Passive constructions have a uniform structure in Eastern Polynesian languages. The verb is affixed by *-Cia*,⁴ the underlying object appears in NOM, and the underlying subject appears as an oblique NP, marked by the agent marker *e*.

Other Polynesian languages, which are ergative, such as Tongan also have three constructions.

³ Abbreviations used in this paper include: ABS absolutive; ACC accusative; AGT agent; ALIEN alienable; DEF definite; d. dual; ERG ergative; FUT future; INALIEN inalienable; INDEF indefinite; NOM nominative; PRED predicate; POSS possessive; PRS present; PST past; s. singular; TNS tense; 1 first person; 3 third person. The asterisk * indicates that the sentence is ungrammatical.

⁴ *-Cia* is the label generally used to represent variable forms of the so-called Polynesian passive suffix. C stands for the variable consonant, which is lexically determined. For example, in Maori there are eleven variants: *-a*, *-ia*, *-hia*, *-kia*, *-mia*, *-ŋia*, *-ria*, *-tia*, *-ina*, *-kina*, *-whina*. Pawley (2001) agrees with Clark (1976) that Proto Polynesian **-Cia* was not passive but transitive, and gives the following analysis: *-i* comes from the Proto Oceanic short transitive suffix and *-a* from the Proto Oceanic object clitic, which was required when the object of a transitive verb was a specific noun.

(2) Tongan

Intransitive	Tns V ('a) NP	<i>Na'e ha'u 'a e tangata.</i> PST come ABS DEF man 'The man came.'
Middle	Tns V ('a) NP <i>i/ki</i> NP	<i>Na'e sai'ia 'a e tangata 'i</i> PST like ABS DEF man in <i>he tamasi'i.</i> ⁵ DEF boy 'The man liked the boy.'
Transitive	Tns V 'e NP ('a) NP	<i>Na'e inu 'e he tangata 'a</i> PST drink ERG DEF man ABS <i>e vai.</i> DEF water 'The man drank the water.'

The striking resemblance between the three constructions in Eastern Polynesian and those found in Tongan is immediately noticeable. The Tongan middle construction is identical to the Maori transitive construction in that the subject is unmarked⁶ and the other NP is marked by a preposition *i* 'in, at, on' or *ki* 'to'. The choice of preposition is lexically determined by the verb. Middle constructions are so called because they are semantically transitive, requiring two participants, but syntactically intransitive, taking only one core argument. The NP marked by *i/ki* is considered an oblique object. Similarly, transitive constructions in Tongan and passive constructions in Maori are very alike except that the verb is always suffixed by *-Cia* and the *e*-marked NP occurs sentence-finally in the latter.

In this discussion passive is considered to be a productive syntactic operation that applies to a transitive construction and detransitivises it.⁷ Specifically, passivisation has the following effects: (a) the underlying object appears as the surface subject in terms of case marking and other syntactic properties; and (b) the underlying subject appears as an oblique NP if they are present at all. Passive in Maori illustrated above conforms to this definition, as far as the case marking is concerned. Specifically, passivisation in Maori has the following characteristics: (a) *-Cia* attaches to a transitive verb; (b) the underlying object (marked by *i* in the transitive construction) is unmarked, suggesting that it is in NOM; and (c) the underlying subject (unmarked in the transitive construction) is marked by *e*, suggesting that it is an oblique NP.

⁵ The difference between the two definite articles *e* and *he* are morphophonological. The latter is used immediately following the ergative case marker 'e or prepositions 'i 'in, at', *ki* 'to', and *mei* 'from'.

⁶ 'Unmarked' here does not necessarily mean the absence of noun markers. In Tongan, it is marked by 'a. However, this element can be considered to be 'unmarked' for the following reasons: (a) compared to the distribution of 'e, which is limited to the transitive subject, the distribution of 'a can be characterised as default; and (b) it can be, and often is, dropped when it is followed by a definite article *e*. Note also that in other languages that show a similar pattern such as Samoan and East Futunan, ABS is marked by zero-morpheme.

⁷ Thus, stative verbs that are translated as passive in English such as Tongan *lavea* 'injured' and *ngalo* 'forgotten' are not considered to be passive in the following discussion, as they do not involve any syntactic operation.

2.2 Tongan -Cia

Passive constructions similar to those of the Eastern Polynesian languages cannot be found in Tongan. It is true that *-Cia* also exists in Tongan and some verbs ending in *-Cia* do seem to have a passive meaning. However, *-Cia* in Tongan does not function as a passive suffix that induces passivisation as a syntactic operation. Let us assume that a passive affix has at least the following two properties: (a) affixation is productive, and (b) it changes the valency of a (transitive) verb to which it attaches with the agent argument demoted to an oblique NP. Churchward (1953:242) notes that *-Cia* verbs in Tongan divide into three subclasses according to their meaning: (a) durative, (b) passive, and (c) courtesy speech. The class of ‘intransitive verbs that may appear to be passive’ includes verbs ending in *-a*, *-fia*, *-hia*, *-kina*, *-mia*, *-ngia*, *-sia*, or *-ia*. The *-Cia* verbs in Tongan are typically intransitive, but the root may be a transitive verb, adjective, or noun. In addition, as noted by Chung (1978), the affixation of *-Cia* to transitive verbs is not a productive process. Of the 94 *-Cia* verbs included in the *Tongan Dictionary* (Churchward 1959), 71% have passive meaning.⁸ However, intransitive verbs that have a passive meaning and are derived from a transitive verb make up only 3% of the total. What is more, we find some instances in which the derived word is a transitive verb. In other words, *-Cia* does not necessarily detransitivise a verb. Indeed, in some cases, it actually transitivises an intransitive verb.

If *-Cia* is not a passive morpheme, then, why is it that so many of the *-Cia* verbs have passive meaning? It seems that the Tongan *-Cia* bears a feature [+affected]. In fact, 87% of the *-Cia* verbs included in the dictionary differ from their stem in terms of the feature [+affected].⁹ Consider the following examples.

- (3) a. *maluuluu* ‘moist, soft’ – *maluuluungia* ‘moistened, softened’
 b. *’ataa’ataa* ‘free, not busy’ – *’ataa’ataaina* ‘to be freed, cleared’
 c. *’uha* ‘to rain’ – *’uheina* ‘to be caught in the rain’
 d. *’anuhi* ‘to spit’ – *’anuhia* ‘to mess up by spitting on’

The affected entity is the subject if it is an intransitive verb or adjective, or the object if it is a transitive verb. Thus, it seems that Tongan *-Cia* is a morpheme that assigns a semantic feature [+affected] to its internal argument, rather than a syntactic feature [+passive]. Since the passive voice is necessarily associated with affectedness, a sentence containing a *-Cia* verb is likely to be translated as passive in English. Hence, we may argue that the passive meaning associated with *-Cia* verbs arises due to the semantic property [+affected] and not as a result of passivisation.

3 Passive-like constructions in Tongan

We have seen that *-Cia* is not a passive morpheme in Tongan and that passive constructions similar to those found in Eastern Polynesian do not exist in Tongan.

⁸ Note, however, that some of the verb roots in any case have a passive meaning: e.g., *malu* ‘to be sheltered’ vs. *malungia* ‘to be shaded, overshadowed’.

⁹ A similar observation has been made by Cook (1988) regarding *-Cia* in Samoan, which is also an ergative language.

However, there are constructions that are claimed to have passive interpretation in Tongan. These are the agentless transitive (VO) and VOS constructions.

3.1 Agentless transitive

In natural Tongan discourse, we frequently encounter sentences with a transitive verb, but with only one argument. Churchward (1953) claims that such sentences are interpreted as passive. Consider (4) below. The verb, *'ave* 'take' is a transitive verb and usually takes two arguments, as illustrated by (4a). However, (4b) contains only one argument, which is marked as ABS.

- (4) a. *Na'e 'ave 'e Sione 'a e tamasi'i ki he fale mahaki.*
 PST take ERG Sione ABS DEF child to DEF house sick
 'Sione took the boy to the hospital.'
- b. *Na'e 'ave 'a e tamasi'i ki he fale mahaki.*
 PST take ABS DEF boy to DEF house sick
 'The boy was taken to the hospital.'

By definition, ABS-marked NPs are either intransitive subjects or direct objects. Thus, it is possible to analyse (4b) as an intransitive construction with the ABS-marked NP being the subject.¹⁰ In this view, the verb *'ave* in (4b) is considered to be intransitive as a result of passivisation, presumably by the affixation of a zero-passive morpheme. This is the analysis proposed by Lynch (1972).¹¹ In contrast, Churchward (1953) and Tchekhoff (1973) consider the agentless transitive to be an instance of argument-drop, an operation freely permitted in Tongan.

3.2 VOS constructions

While the unmarked order is VSO, Tongan freely allows VOS sentences. Churchward (1953) notes that the VOS order is used when the emphasis is on the object. He argues that VOS sentences are translated as passive because in English it is the way to encode the emphasis on the object.

- (5) a. *Na'e 'ave 'e he faiako 'a e tamasi'i.*
 PST take ERG DEF teacher ABS DEF boy
 'The teacher brought the boy.'
- b. *Na'e 'ave 'a e tamasi'i 'e he faiako.*
 PST take ABS DEF boy ERG DEF teacher
 'The boy was brought by the teacher.'

¹⁰ Tchekhoff (1973) observes that Tongan verbal constructions with only an ABS-marked NP can be divided into three classes: a) those in which the ABS-NP can only be interpreted as the patient, such as *tamate* 'to kill', b) those in which the ABS-NP can only be interpreted as the agent, such as *'ave* 'to take', *fa'ele'i* 'to bear a baby', and c) those in which the ABS-NP can be interpreted either as the agent or the patient, such as *taki* 'to lead'. However, Tchekhoff does not consider the case (b) as passive.

¹¹ To be precise, Lynch (1972) claims that Tongan is an accusative language, arguing that *'a*, *'i*, and *'e* are the NOM-marker, the ACC-marker, and the Agent marker respectively, as in Eastern Polynesian.

On the other hand, Lynch (1972) proposes that the VOS construction is syntactically passive. He argues that the NP marked by ‘*a*’ is the (derived) subject and the NP marked by ‘*e*’ is the agent appearing in oblique, ‘*e*’ being the agent marker.¹² In the next section, we will see that the ABS-marked NPs in VOS and the agentless transitive are not syntactic subjects, and that these constructions should not be regarded as passive.

3.3 Are the ABS-NPs in passive-like constructions ‘subject’?

Subjects, whether ABS or ERG, are distinguished from objects in Tongan in a number of syntactic phenomena: e.g. the use of possessive pronouns, the use of clitic pronouns, control, and *mo*-coordination. In this section, we will examine whether the ABS-marked NPs in VOS and agentless passive constructions behave like subjects in terms of these syntactic phenomena.

3.3.1 Possessive pronouns

Tongan has two sets of possessive pronouns: alienable (‘*e*-class’) and inalienable (*ho*-class). When an alienable possessive pronoun precedes a verb, it refers to the subject of the verb. In contrast, an inalienable possessive pronoun preceding a verb refers to the object. Thus, inalienable possessive pronouns cannot occur with an intransitive verb.¹³ See (6) below. As illustrated by (6a, b), the subject of an intransitive verb must be represented by an alienable possessive pronoun. Example (6c) shows that when used with a transitive verb, an alienable possessive pronoun refers to the subject. In contrast, (6d) shows that an inalienable possessive pronoun must refer to the object when used with a transitive verb.

¹² Lynch’s analysis is questionable in some crucial respects. First, empirical evidence does not support his assumption that Tongan is an accusative language. Secondly, transitive verbs such as ‘*ave*’ ‘take’ never occur in what Lynch calls accusative construction, as illustrated below.

Na‘e ‘ave ‘a e faiako kia Sione.
 PST take ABS DEF teacher ACC Sione
 Intended meaning: *‘The teacher took Sione.’
 Actual meaning: ‘(Someone) took the teacher to Sione.’

Lynch proposes that some verbs in Tongan may only occur in intransitive passive, but not in active transitive.

¹³ Stanley Starosta (pers. comm.) points out that this is an instance of nominalisation and therefore cannot be used to make hypotheses about clause structure. While his point is acknowledged, it should be noted that it only poses a problem if the derived noun has an argument structure distinct from the base verb. However, Tongan zero-derived nominalisation keeps the argument structure of the base verbs. It is always the subject that appears with ‘*e*-possessive’. In addition, they can also take the object, as illustrated below. In this respect, it is rather similar to gerunds in English.

‘ene kai ‘a e ika
 ALIEN.POSS.3.s. eat ABA DEF fish
 ‘his eating the fish’

For this reason, I consider the use of alienable possessive pronouns to be a phenomenon associated with subjects in Tongan.

- (6) a. *'ene foki*
ALIEN.POSS.3.s. return
'his returning'
- b. * *hono foci*
INALIEN.POSS.3.s. return
- c. *'ene taki*
ALIEN.POSS.3.s. lead
'his leading (someone)'
- d. *hono taki*
INALIEN.POSS.3.s. lead
'his being led (by somebody)'

The ABS-marked NP in VOS and agentless transitive constructions behaves the same as objects rather than subjects: it cannot occur as an alienable possessive pronoun. Consider (7) below.

- (7) a. *'ene 'ave ki he fale mahaki*
ALIEN.POSS.3.s. take to DEF house sick
* 'his being taken to the hospital'
'his taking (someone) to the hospital'
- b. *hono 'ave ki he fale mahaki*
INALIEN.POSS.3.s. take to DEF house sick
'his being taken to the hospital'

As (7a) shows, *'ave* 'take' cannot occur with an alienable possessive pronoun and have a passive meaning. Note that the phrase itself is grammatical, as long as the passive interpretation is not intended. Furthermore, (7b) confirms that the ABS-marked argument in (4b) as well as (5b) is the object.

3.3.2 Clitic pronouns

Tongan has a set of clitic pronouns, which appear in the position between the tense marker and the verb. The use of clitic pronouns is restricted to subjects. A pronominal object must be realised as an independent pronoun. See (8) below.

- (8) a. *Na'a ku 'alu ki ai.*
PST 1.s. go to there
'I went there.'
- b. *Na'a ku 'ave 'a e tamasi'i ki ai.*
PST 1.s. take ABS DEF boy to there
'I took the boy there.'
- c. * *Na'a ku 'ave 'e he faiako ki ai.*
PST 1.s. take ERG DEF teacher to there
'The teacher took me there.'

- d. *Na'e 'ave au 'e he faiako ki ai.*¹⁴
 PST take 1.s. ERG DEF teacher to there
 'The teacher took me there.'

As shown by (8a), a pronominal subject in an intransitive construction occurs as a clitic. Thus, if the ABS-marked argument in VOS and agentless transitive constructions is in fact the subject, we would expect a construction similar to (8a) when the relevant argument is pronominal. The following data argue against this hypothesis.

- (9) a. *Na'a ne 'ave ki he fale mahaki.*
 PST 3.s. take to DEF house sick
 'He took (someone) to the hospital.'
 *'He was taken to the hospital.'
- b. * *Na'a ne 'ave 'e he faiako ki he fale mahaki.*¹⁵
 PST 3.s. take ERG DEF teacher to DEF house sick

As illustrated by (9a), the ABS argument of an agentless transitive cannot be realised as a clitic pronoun. Note that the sentence is grammatical, but the clitic pronoun must be taken as the subject of the verb 'ave 'take'. Similarly, the ABS argument of a VOS construction cannot occur as a clitic pronoun, as shown by (9b). The sentence is ungrammatical, for it contains two subjects, the clitic *ne* and the ERG-marked NP, *he faiako* 'the teacher'. These data show that the ABS argument is not the subject, but the object in VOS and agentless transitive constructions.

3.3.3 Control

The third test involves control. I assume that the embedded clause in control constructions such as (10) below contains a phonetically null argument PRO that is coreferential with an argument of the matrix verb.

- (10) a. John wants [PRO to go].
 b. John persuaded Mary [PRO to go].

It is attested that, universally, PRO can occur in the subject position of the embedded clause, but never in the object position (Chomsky 1981). For example, consider the following English sentences.

¹⁴ Note that pronominal objects appear in the position immediately after the verb and without the ABS marker, rather than in the regular position following the subject. Otsuka (2000) proposes that this is due to a rule that requires incorporation of pronominal object into the verb. In marked speech, independent pronouns can occur in the regular VSO order, both as a subject and as an object. However, in that case, they must be preceded by an appropriate case marker, as illustrated below.

Na'e 'ave 'e ia 'a e tamasi'i ki he fale mahaki.
 PST take ERG 3.s. ABS DEF boy to DEF house sick
 'He took the boy to the hospital.'

¹⁵ Sentence (9b) is ungrammatical because clitic doubling is not permitted in Tongan. Clitic pronouns cannot co-occur with a coreferential NP.

- (11) a. John wants [PRO to be praised].
 b. *John wants [someone to praise PRO].

Sentence (11a) is grammatical because PRO is the subject of the embedded clause, whereas (11b) with PRO as the object is ruled out, as it violates the above-mentioned universal restriction on control.

Our question is whether VOS and/or the agentless transitive can occur in a control construction with PRO in place of the ABS argument. If such a construction is permissible, then the ABS-marked argument in VOS and the agentless transitive should be considered to be the subject. As illustrated by (12) below, PRO cannot occur in place of the ABS-marked argument in VOS and agentless transitive constructions.¹⁶ Sentences (12a) and (12b) are ungrammatical because PRO occurs as an object.¹⁷

- (12) a. <Agentless Transitive>
 *'Oku loto 'a e tamasi'i [ke 'ave PRO ki ai].
 PRS want ABS DEF boy to take to there
 'The boy wants to be taken there.'
- b. <VOS>
 *'Oku loto 'a e tamasi'i [ke 'ave PRO 'e he
 PRS want ABS DEF boy to take ERG DEF
 tangata ki ai].
 man to there
 'The boy wants to be taken there by the man.'

To conclude, the control data also show that the ABS-marked argument in VOS and the agentless transitive is not the subject. In other words, these constructions should not be regarded as passive.

3.3.4 *mo* coordination

Finally, let us consider yet another phenomenon that exclusively applies to subjects. When two clauses are conjoined by the conjunction *mo*, the second clause may, and generally does, contain a gap. The gap is taken to be a result of argument deletion under identity with an argument of the first clause. However, the deletion does not apply freely, but is restricted to subjects.¹⁸ That is, an argument may be deleted only if it is the subject

¹⁶ Ergative arguments can be PRO as illustrated below.

'Oku loto 'a e tamasi'i [ke 'ave PRO 'a e tangata ki ai]
 PRS want ABS DEF boy to take ABS DEF man to there
 'The boy wants to take the man there.'

¹⁷ However, note that both (12a) and (12b) are grammatical if we take PRO to be controlled by the matrix subject *e tamasi'i* 'the boy'. The embedded object is assumed to be phonetically null, an instance of argument-drop. Accordingly, (12a) means 'The boy wants to take (someone) there', and (12b), 'The boy wants the man to take (someone) there'.

¹⁸ This may sound obvious to the reader. However, it should be noted that Tongan demonstrates an ergative pattern with another conjunction *pea*: an argument of the second clause may be deleted only if it bears the same case (either ERG or ABS) as the coreferential argument in the first clause. Hence, the subject-only constraint on *mo* coordination is a useful diagnostic for subjecthood in Tongan.

of the second clause and is coreferential with the subject of the first clause. In other words, the two clauses conjoined by *mo* must have the common subject. See (13) below, in which \emptyset stands for the gap.

- (13) a. **Na'e tangi 'a Hina_i mo taa'i 'e Mele \emptyset _i.*
 PST cry ABS Hina and hit ERG Mele
 'Hina cried and Mele hit (her).'
- b. *Na'e tangi 'a Hina_i mo taa'i \emptyset _i 'a Mele.*
 PST cry ABS Hina and hit ABS Mele
 'Hina cried and (she) hit Mele.'

Sentence (13a) is ungrammatical since the gap in the second clause is the object. Similarly, sentences in (14) illustrate that the gap must be coreferential with the subject of the first clause; the subject of the second clause cannot be deleted if it is coreferential with the object of the first clause.

- (14) a. *Na'e taa'i 'e Hina_i 'a Mele_j mo kata \emptyset _{i/*j}.*
 PST hit ERG Hina ABS Mele and laugh
 'Hina hit Mele and (Hina/*Mele) laughed.'
- b. *Na'e taa'i 'e Sione_i 'a Pila_j mo 'akahi \emptyset _{i/*j} 'a Taniela.*
 PST hit ERG Sione ABS Pila and kick ABS Taniela
 'Sione hit Pila and (Sione/*Pila) kicked Taniela.'

The subject-only constraint on *mo* coordination can be used as a diagnostic for subjecthood. If the ABS-marked NP in a VOS and/or an agentless transitive construction is the subject, then, it should be able to be part of *mo* coordination, either as the gap or as the antecedent. Example (15) shows that *mo* coordination is not permissible with such an NP.

- (15) a. <Agentless Transitive>
Na'e taa'i 'a Mele mo tangi \emptyset .
 PST hit ABS Mele and cry
 'Mele was hit and (*Mele/the one who hit Mele) cried.'
- b. <VOS>
Na'e taa'i 'a Mele 'e Hina mo tangi \emptyset .
 PST hit ABS Mele ERG Hina and cry
 'Mele was hit by Hina and (*Mele/Hina) cried.'

Both (15a) and (15b) are grammatical, with the gap obligatorily taken to be coreferential with the agent of the first verb, i.e. the implied subject in (15a) and *Hina* in (15b). Similarly, the ABS-NP of the agentless transitive and VOS constructions cannot occur as a gap coreferential with the intransitive subject of the first clause, as shown in (16) below. This suggests that the ABS-NP is not the subject of the relevant constructions.

- (16) a. <Agentless Transitive>
 * *Na'e tangi 'a Mele mo taa'i \emptyset .*
 PST cry ABS Mele and hit
 Intended meaning: 'Mele cried and was hit.'

b. <VOS>

- * *Na'e tangi 'a Mele mo taa'i Ø 'e Hina.*
 PST cry ABS Mele and hit ERG Hina
 Intended meaning: 'Mele cried and was hit by Hina.'

To summarise, none of the four syntactic phenomena in which subjects are distinguished from objects treat the ABS-marked NP of VOS and the agentless transitive as a subject. Consequently, we may conclude that these constructions are not passive as defined in §2.1 above. In the remaining sections, an alternative analysis of VOS and the agentless transitive in Tongan will be proposed.

4 Agentless transitive as *pro*-drop

Let us discuss the agentless transitive first. I propose that the agentless transitive contains a phonetically null subject. Theoretically, there are two possibilities as to what this phonetically null subject can be: (a) *pro* and (b) topic variable in the sense of Huang (1984, 1989). It will be shown below that both possibilities are available in Tongan.

4.1 *Pro*-drop in Tongan

In Tongan, a pronominal argument can be omitted under two conditions: (a) it is third person singular and (b) its referent is identifiable by virtue of context. Thus, we frequently encounter subjectless sentences as well as objectless sentences in Tongan. It should be noted that the subject may be omitted in intransitive as well as transitive constructions. Furthermore, a subjectless transitive sentence does not necessarily require the passive interpretation. Consider (17) and (18) below.

- (17) A: *'Oku 'i fē 'a Mele?*
 PRS in where ABS Mele
 'Where is Mele?'

B: *Na'e 'ave 'a Sione ki he ako.*
 PST take ABS Sione to DEF school
 '(She) took Sione to school.'

- (18) A: *Ko e hā e me'a na'e fai 'e Pita?*
 PRED DEF what DEF thing PST do ERG Pita
 'What did Pita do?'

B: *Na'e kai 'a e ika.*
 PST eat ABS DEF fish
 '(He) ate fish.'

Let us assume that the phonologically absent subject in (17) and (18) is syntactically present, because otherwise such sentences would be ruled out due to the Extended Projection Principle (EPP), which requires that a sentence have a subject (Chomsky 1981).

It is known that empty pronouns in Chinese and Japanese observe a similar condition: their identity must be retrievable by virtue of context. Huang (1984, 1989) proposes that the Chinese-type empty pronouns are not *pro*, but should be treated as a variable bound by

topic. I take Huang's approach to be applicable to Tongan and assume the null pronouns in cases like (17) and (18) to be topic variables.

Another kind of phonetically null subjects is found with unaccusative verbs that take a sentential complement.

- (19) a. *'Oku tonu [ke 'alu 'a Sione].*
 PRS correct to go ABS Sione
 'It is good for Sione to go.'
- b. *'Oku tapu [ke ifi 'a e tamaiki ako].*
 PRS prohibited to smoke ABS DEF children school
 'It is prohibited for the students to smoke.'

These sentences lack an overt subject. Let us assume that there is a phonetically null expletive in the subject position of these constructions (cf. Levin and Massam (1986) for a similar analysis of the relevant Niuean data).

4.2 Proposal: topic variable and *pro*_{arb}

I propose that the agentless transitive involves the topic variable: when the subject is third person singular and represents the known information, it is realised as a topic variable rather than a pronoun. This conforms to Churchward's (1953) observation that the agentless transitive is used when the object is emphasised. In a context where the subject represents known information and the object, the new information, the former is reduced to the topic variable. Consequently, only the object surfaces as an overt element. I propose that an agentless transitive construction such as (4b) above has the representation illustrated in (20), in which *e* stands for the topic variable.

- (20) *Na'e 'ave e 'a e tamasi'i ki he fale mahaki.*
 PST take ABS DEF boy to DEF house sick
 '(He) took the boy to the hospital.'

Note that given the topic-variable analysis, a null pronoun in Tongan must not have arbitrary reference. However, this is not entirely true. While the agentless transitive always implies the existence of agent, the speaker need not know the identity of the agent in question. Thus, in an appropriate context, an agentless transitive may have a generic subject similar to English *they*.¹⁹ Consider the following examples.

- (21) a. *Na'e holoki 'a e fale lotu.*
 PST demolish ABS DEF house pray
 'The church was demolished.'

¹⁹ This point is made particularly clear by one informant, who claims that the use of the agentless transitive in a context like that shown below is unacceptable.

Na'e 'ave 'a e tamasi'i ki he fale mahaki.
 PST take ABS DEF boy to DEF house sick
Ko Sione na'a ne 'ave ia ki ai.
 PRED Sione PST 3.s. take 3.s. to there
 'The boy was taken to the hospital. It is Sione who took him there.'

- b. 'E fetongi 'a e maka.
 FUT replace ABS DEF battery
 'The battery will be replaced.'

In order to account for these examples, it is necessary to assume another type of null pronoun. I propose that sentences such as (21a, b) contain the generic *pro*, not the topic variable. The generic *pro* must have arbitrary reference, in a way analogous to PRO with arbitrary reference.

To summarise, I propose that there are three kinds of null pronouns in Tongan: (a) topic variable, (b) expletive, and (c) *pro* with arbitrary reference, *pro*_{arb}. Agentless transitive is indeed a transitive construction with a null subject. The null subject may either be a topic variable or *pro*_{arb}. Accordingly, an agentless transitive may have either of the following two constructions.

- (22) a. TNS V topic-variable ABS-NP
 b. TNS V *pro*_{arb} ABS-NP

The (22a) construction is used when the identity of the agent has been established in the previous conversation. The (22b) construction is used when the existence of the agent is implied but its identity is not known to the speaker. The presence of the phonetically null subject (i.e., either the topic variable or *pro*_{arb}) is supported by the *mo* coordination test. Both the topic variable and *pro*_{arb} may serve as the argument coreferential with the gap in the second clause. See (23) below: *e* stands for the topic variable, and \emptyset , the gap.

- (23) a. Na'e 'ave e 'a e tamasi'i ki h fale mahaki mo foki \emptyset .
 PST take ABS DEF boy to DEF house sick and return
 '(He) took the boy to the hospital and returned.'
- b. Na'e holoki *pro*_{arb} 'a e fale lotu mo langa \emptyset 'a
 PST demolish ABS DEF house pray and build ASB
e fale mahaki.
 DEF house sick
 '(They) demolished the church and built a hospital.'

5 VOS construction as scrambling

As for VOS construction, I propose that it is derived from the unmarked order VSO by scrambling. So far, we have restricted our attention to the transitive constructions and the alternation between VSO and VOS. However, such an alternation is not limited to transitive constructions. As illustrated below, the order in which NPs and PPs appear in intransitive constructions (24) and middle constructions (25) is also flexible although the default order is one in which the subject NP precedes the PP.

- (24) a. Na'e 'alu 'a Sione ki Tonga.
 PST go ABS Sione to Tonga
 'Sione went to Tonga.'

- b. *Na'e 'alu ki Tonga 'a Sione.*
 PST go to Tonga ABS Sione
 'Sione went to Tonga.'
- (25) a. *Na'e tokoni 'a Sione ki he faiako.*
 PST help ABS Sione to DEF teacher
 'Sione helped the teacher.'
- b. *Na'e tokoni ki he faiako 'a Sione.*
 PST help to DEF teacher ABS Sione
 'Sione helped the teacher.'

Note that in (24) and (25) there is no semantic difference between the (a)-sentences and (b)-sentences. Such freedom regarding the constituent order is found in certain other languages such as Japanese. The phenomenon is referred to as scrambling in the literature. Given the fact that the word order alternation is not restricted to the one between VOS and VSO in Tongan, we may consider Tongan to be a language that permits scrambling. Hence, it can be argued that VOS is an instance of scrambling.

Furthermore, there is evidence to suggest that native speakers never interpret VOS constructions as passive. The idea that VOS sentences are interpreted as passive derives from Churchward's (1953:67) comment on the difference between VOS and VSO: 'The difference in emphasis is much the same as it is in English between 'David killed Goliath' and 'Goliath was killed by David'. However, Churchward also emphasises the point that the two sentences do not differ from each other in terms of voice. His claim can be summarised as follows: (a) English uses passive when the emphasis is on the object rather than the subject, and (b) Tongan, lacking passive, expresses the emphasis on the object by placing the object before the subject. Churchward uses a passive in translating the Tongan VOS sentence in order to illustrate the subtle difference between VOS and VSO in Tongan. The difference is the focus of the sentence, which can be more accurately compared to the pitch accent in English. In English the element in focus carries the pitch accent. Thus, the more accurate translation of Churchward's example cited above would be '**David** killed Goliath' and 'David killed **Goliath**', respectively, where the elements carrying the pitch accent are in bold type.

Native speakers consider that VOS and VSO are essentially the same semantically when the two sentences are presented in isolation without any context. However, a significant distinction arises when certain context is provided. When asked to select the appropriate answer(s) to a *wh*-question, Tongan speakers clearly distinguish VOS from VSO. Compare (26) and (27) below.

- (26) Q: *Ko hai na'e fili 'e Sione?*
 PRED who PST choose ERG Sione
 'Who did Sione choose?'
- A1: *Na'e fili 'a Pila 'e Sione.*
 PST choose ABS Pila ERG Sione
 'Sione chose **Pila**.'
- A2: **Na'e fili 'e Sione 'a Pila.*
 PST choose ERG Sione ABS Pila
 '**Sione** chose Pila.'

- (27) Q: *Ko hai na'a ne fili 'a Pila?*
 PRED who PST 3.s. choose ABS Pila
 'Who chose Pila?'
- A1: **Na'e fili 'a Pila 'e Sione.*
 PST choose ABS Pila ERG Sione
 'Sione chose **Pila**.'
- A2: *Na'e fili 'e Sione 'a Pila.*
 PST choose ERG Sione ABS Pila
 '**Sione** chose Pila.'

When the *wh*-word is the object, the answer must be VOS. Similarly, when the *wh*-word is the subject, the answer must be VSO.²⁰ VOS in this case does indicate the focus on the object.

Thus, it is reasonable to argue that scrambling is triggered by the feature [+focus] in Tongan. The focused element is fronted so as to appear immediately following the verb. When the object bears the feature [+focus], it will appear in this position, thus preceding the subject, yielding the order VOS.

6 Concluding remarks

This study re-examines the two constructions in Tongan, agentless and VOS transitive, which have been claimed to have passive meaning. Empirical evidence supports the current proposal that they do not involve passivisation as a syntactic operation, but have to do with *pro*-drop and scrambling, respectively. First, the ABS-marked NPs in these constructions are not treated as subjects in various syntactic phenomena such as the use of possessive pronouns and clitic pronouns, control and *mo* coordination. Second, I have argued that the agentless transitive contains a phonetically null subject. Tongan permits two types of null arguments, topic variables and the arbitrary *pro*. The passive interpretation associated with the agentless transitive arises when the subject is the arbitrary *pro*, i.e. unspecified agent. Finally, it has been shown that it is misleading to say that VOS is associated with passive meaning. Rather, the contrast between VOS and VSO is that the focus is on the object in the former. The situation can be captured more appropriately if we regard VOS as a result of scrambling triggered by focus. When the object has the feature [+focus], it will be fronted and appear in the position preceding the subject.

The current study raises some important questions concerning the study of passives in Polynesian languages, particularly those belonging to the Tongic and Samoic-Outlier subgroups. As mentioned earlier, in contrast to the Eastern Polynesian languages these languages are said to lack passive constructions with *-Cia* suffixed verbs. It should be noted that the alternation between VSO and VOS is also freely permitted in other Polynesian languages such as Samoan, East Uvean, and East Futunan. These languages are similar to

²⁰ A more natural, preferred option is to use the nominal construction such as below.

Ko Sione.
 PRED Sione
 'It is Sione.'

Tongan in that (a) case marking is ergative and (b) *-Cia* does not function as a productive passive morpheme.²¹ Native speakers of these languages also do not seem to be aware of any particular semantic differences between VOS and VSO. In other words, VOS and VSO are used interchangeably. This observation leads us to the following questions: (a) whether VOS constructions in these languages are also associated with focus; and more generally, (b) whether there is a correlation between the lack of passive and VOS construction. In fact, the more fundamental question would be whether these languages actually do not have passives, and, if not, what alternatives they have to compensate for the lack of passives. In order to answer these questions, we need to specify the functions of passive (syntactic, semantic, or pragmatic) and examine how each of such functions is realised in the languages which apparently lack passives. The present discussion on the passive-like constructions in Tongan is intended to serve as a starting point of the more extensive research on languages without passives.

References

- Biggs, Bruce, 1969, *Let's learn Maori*. Wellington: A.H. and A.W. Reed.
- Chomsky, Noam, 1981, *Lectures on government and binding*. Berlin: Mouton de Gruyter.
- Chung, Sandra, 1978, *Case marking and grammatical relations in Polynesian*. Austin: University of Texas Press.
- Churchward, C. Maxwell, 1953, *Tongan grammar*. Oxford: Oxford University Press.
- 1959, *Tongan dictionary*. Nuku'alofa: Tongan Government Printing Press.
- Clark, Ross, 1976, *Aspects of Proto-Polynesian syntax*. Auckland: Linguistic Society of New Zealand.
- Cook, Kenneth, 1988, A cognitive analysis of grammatical relations, case, and transitivity in Samoan. PhD dissertation, University of California, San Diego.
- 1996, The *Cia* suffix as a passive marker in Samoan. *Oceanic Linguistics* 35:57–76.
- Dixon, R.M.W., 1979, Ergativity. *Language* 55:59–138.
- Gibson, Jeanne D. and Stanley Starosta, 1990, Ergativity East and West. In P. Baldi, ed. *Linguistic change and reconstruction methodology*, 195–210. Berlin: Mouton de Gruyter.
- Huang, C.-T. James, 1984, On the distribution and reference of empty pronouns. *Linguistic Inquiry* 15: 531–574.
- Huang, C.-T. James, 1989, *Pro-drop in Chinese: a generalized control theory*. In Osvaldo Jaeggli and Kenneth J. Safir, eds *The null subject parameter*, 185–214. Dordrecht: Kluwer Academic.
- Levin, Juliette and Diane Massam, 1986, Classification of Niuean verbs: notes on case. In Paul Geraghty, Lois Carrington and S.A. Wurm, eds *FOCAL I: Papers from the Fourth International Conference on Austronesian Linguistics*, 231–244. Canberra: Pacific Linguistics.

²¹ Cook (1996) argues that certain Samoan clause types that contain *-Cia* suffixed verbs can be analysed as impersonal passives.

- Lynch, John, 1972, Passives and statives in Tongan. *Journal of the Polynesian Society* 81:5–18.
- Otsuka, Yuko, 2000, Ergativity in Tongan. Doctoral dissertation, Oxford University.
- Pawley, Andrew, 2001, Proto Polynesian *-CIA. In Joel Bradshaw and Kenneth L. Rehg, eds *Issues in Austronesian morphology*, 193–216. Canberra: Pacific Linguistics.
- Tchekhoff, Claude, 1973, Some verbal patterns in Tongan. *Journal of the Polynesian Society* 82:281–292.

6 *Grammatical relations in Puyuma*

STACY FANG-CHING TENG

1 Introduction

Payne (1997:129), points out that grammatical relations, such as ‘subject’ or ‘object’, can be defined as ‘relations between arguments and predicates in a level of linguistic structure that is independent of semantic and pragmatic influences’. He also suggests that the formal properties that most directly identify grammatical relations are case marking, participant reference marking on verbs, and constituent order. Among the three properties, I find the first two, the case marking and participant reference marking (or the voice system), play important roles in showing grammatical relations in Puyuma.¹

The paper is organised as follows: in §2, the voice system will be examined, and actor/non-actor voice asymmetries will be investigated. Section 3 deals with the case marking of nouns and pronouns. In §4, I will discuss semantic and syntactic aspects of transitivity. A discussion of the notion of ‘ergativity’ as it relates to Puyuma follows in §5.

2 Voice

Like many other Formosan or Philippine languages, in Puyuma four different voices can be distinguished,² namely, actor voice, patient voice, locative voice, and instrumental/beneficiary voice, as exemplified from (1a) to (1d).

¹ I am grateful to Wayan Arka, John Bowden, Malcolm Ross and Elizabeth Zeitoun for their comments on drafts of this paper. The responsibility for its contents is mine.

² Different labels and categorisations are often given by different linguists in the literature. For example, Cauquelin (1991) names them actor focus, object focus, referent focus, and instrument focus; while Tan (1997) and Huang (2000a) merge Cauquelin’s object focus and referent focus under the label patient focus and distinguish only three different voices, which are actor focus, patient focus, and instrumental/beneficiary focus. Teng (1997) makes a distinction only between actor voice and non-actor focus.

- (1) a. *diru=ku*³
 <AV>bathe-1SG.NOM⁴
 'I washed.'⁵
- b. *ku=dirus-aw na gung*
 1SG.GEN-bathe-PV DF.NOM ox
 'I washed the ox.'
- c. *ku=dirus-ay na enay na bias*
 1SG.GEN-bathe-LV DF.NOM water DF.NOM hot
 'I washed in hot water.'
- d. *ku=dirus-anay na enay kan aliwaki*
 1SG.GEN-bathe-IV DF.NOM water SG.OBL Aliwaki
 'I washed Aliwaki with the water.' (Cauquelin 1991:43–44)

In each sentence, there is only one nominative argument. The voice marker in each example reflects the role of the nominative argument, but the correlation between semantic role and a specific voice marking is not as regular as in the above examples. Taking the instrumental/beneficiary voice as an instance, the nominative arguments in the following examples are not restricted to instrument or beneficiary; but also theme (2a) and patient (2b), and the semantic role in (2c) is hard to define because it is a lexicalised phrase.⁶

- (2) a. *ku=tuLud-anay na sarekuDan kana temumuwan*
 1SG.GEN-pass-IV DF.NOM stick DF.OBL offspring
 'I passed the stick to the offspring.'
- b. *ku=remerem-anay na kinsas*
 1SG.GEN-press.down-IV DF.NOM police
 'I pressed the police down.'
- c. *tu=Tukul-anay tu=Dakur*
 3.GEN-lug-IV his.NOM-back
 'He hunched his back.'

Thus, we need to note that the labels given to those voice affixes are for the sake of convenience; they do not necessarily reflect the semantic role of the nominative argument in a given sentence.⁷

³ The capital letters *D*, *L*, *T* in Puyuma represent retroflex sounds, and the letter *e* schwa [ə], and the apostrophe sign the glottal stop [ʔ]

⁴ Abbreviations used in this paper include: 1, 2, 3 first, second, third person; ASP aspect; AV actor voice; BV beneficiary voice; *Ca-* *Ca-*reduplication; CAUS causative; COP copula; DF definite; EXC exclusive; GEN genitive; ID indefinite; IRR irrealis; IV instrumental voice; LOC locative, LV locative voice; NEG negator; NEU neutral; NMSR nominaliser; NOM nominative; OBL oblique; PL plural; PROG progressive; PROJ projective; PV patient voice; REAL realis; RED reduplicate; SG singular; TOP topic; - indicates a prefix or suffix; <> indicates an infix; = indicates a clitic.

⁵ The free translations in the examples are from Cauquelin, but the interlinear glosses are mine.

⁶ Many linguists (Kess 1976:178; Li 1995:665; Huang 1995:38, among others) have pointed out that for many Philippine or Formosan languages it is almost impossible to have a one-to-one correlation between the voice affixes and the semantic roles they denote.

⁷ Although there seems to be a tendency that these affixes actually indicate the degree of affectedness of the nominative argument.

Not all Puyuma verbs demonstrate four voice alternations as shown in (1a) to (1d). For example, for the verb *kutang* ‘to spear’ in (3), there is no locative voice form **kutang-ay*, and no instrumental voice form **kutang-anay*. For the verb *pulang* ‘to help’ in (4), there is no patient voice form **pulang-aw*, and no instrumental voice form **pulang-anay*.

- (3) a. *aDi=ta la karuwa kutang*
 NEG-1PL.INC.NOM ASP can <AV>spear
 ‘We couldn’t spear.’
- b. *ta=kutang-aw la na Lutang*
 1PL.INC.GEN-spear-PV ASP DF.NOM monkey
 ‘We’ve speared the monkey(s).’
- (4) a. *karuwa=ku la pulang i ruma’ i sabak*
 can-1SG.NOM ASP help.AV LOC house LOC inside
 ‘I can then help my family.’
- b. *ku=pulang-ay i nanali b<en>ase*
 1SG.GEN-help-LV NOM my.mother <AV>wash.clothes
 ‘I helped my mother wash the clothes.’

From the above examples, we find that for a verb to have how many and which voice forms is to an extent unpredictable; first, not every verb has all the four alternations, and second, even when we know the semantic relationship between the verb and the nominative argument, which undergoer voice affix to take is not always predictable. Sometimes, certain voices are missing because of the semantics of this given verb.

Puyuma sentences display some asymmetries between the actor voice sentences and the non-actor voice sentences. First, in sentences denoting realis mood,⁸ while there is a one-to-one correlation between the non-actor voice markers and the voice they denote, the relation between the actor voice markers and actor voice is many-to-one.⁹ For example, in (1a) and (4a), the actor voice is marked by ** and \emptyset respectively.¹⁰

Second, consider the following templates for actor voice verbs and non-actor voice verbs:

- (5) a. V-pro[NOM]
daway=ta
 AV-make-1PL.INC.NOM
 ‘We made (something).’

⁸ Actor voice markers are missing in irrealis and imperative construction.

⁹ This was also demonstrated in Zeitoun et al. (1996) and Tan (1997).

¹⁰ There are at least five different AV markers, including *ma-*, **, *mu-*, *mi-* and \emptyset . Huang (2000b) has pointed out that in Maryrinax Atayal the selection of a particular marker for a verb to some extent lexically determined; some verbs can take more than one AV marker, marking different degrees of dynamicity/transitivity. The same phenomenon is also found in Puyuma. For example:

- a. *ma-bu’ut na lawlaw*
 AV-stop DF.NOM light
 ‘The light went out.’
- b. *b<en>u’ut Da apuy*
 <AV>stop ID.OBL fire
 ‘He stopped a fire.’

- b. pro[GEN]=V=pro[NOM]
tu-ka-aw-ku
 3.GEN-tell-PV-1SG.NOM
 ‘S/he told me.’

For actor voice verbs, there is only one enclitic pronoun.¹¹ However, for non-actor voices, as well as the subject pronoun, there is always a genitive pronoun, the non-subject actor, prefixed to the verb. All full nouns are optional, but the bound pronouns are obligatory. For example:

- (6) a. *aDi tu=na'u-i (na walak) (kan tinataw)*
 NEG 3.GEN-look-LV NOM child OBL his/her/their.mother
 ‘The mother didn’t look after the children.’
- b. **aDi na'u-i na walak kan tinataw*
 NEG look-LV NOM child OBL his/her/their.mother

3 Case marking

Puyuma makes a three-way case distinction: nominative, marking the grammatical subject, genitive, marking the non-subject actor, and oblique, marking the other arguments. While independent nouns depend on the noun phrase markers preceding them to assign a case role, the form of a pronoun shows its case role. Case marking of nouns and pronouns is discussed in §3.1 and §3.2 respectively.

3.1 Case marking of nouns

Three classes of nouns are distinguished: proper nouns, common nouns, and location nouns. Each is preceded by different sets of noun phrase markers. Table 1 presents the noun phrase markers in Puyuma.

Table 1: The noun phrase markers in Puyuma

Types of nouns	Common nouns		Proper nouns		Location nouns
	indefinite	definite	singular	plural	
Cases					
Nominative	<i>a</i>	<i>na/ina</i>	<i>i</i>	<i>na</i>	<i>i</i>
Oblique	<i>Da</i>	<i>kana</i>	<i>kan</i>	<i>kana</i>	

The nominative case marks the grammatical subject in a sentence, while the oblique case marks non-subject arguments. The non-subject arguments include the non-subject

¹¹ There is no third person nominative bound pronoun. So, when a verb is not suffixed with a pronoun, we know that the nominative argument is third person.

agent (which has a genitive bound pronoun cross-reference with the full noun),¹² and the oblique arguments.

Common nouns are subcategorised into two classes in terms of whether the referent is definite or indefinite. For example:

- (7) a. *Dua me-nau-a a Tau i ...*
 come.AV AV-see-PROJ ID.NOM people TOP
 ‘People come to see ...’
- b. *paragan=ta Da ruma*
 build.AV-1PL.NOM ID.OBL house
 ‘We built houses.’
- c. *tu=aLak-aw na barasa*
 3.GEN-take-PV DF.NOM rock
 ‘They took the rock.’
- d. *mu-Tereb kurenang kana baLi*
 AV-fall follow.AV DF.OBL wind
 ‘It fell down with the wind.’

Proper nouns consist of personal names and kinship terms. They are further subcategorised in terms of number, as indicated by sentences (8a) to (8d).

- (8) a. *tu=paDek-aw i temutaw*
 3.GEN-carry.on.back-PV SG.NOM his.grandparent
 ‘He carried his grandmother on his back.’
- b. *tu=bes-besbes-ay kan ma’iDang kakawalan*
 3.GEN-RED-massage-LV SG.OBL old Kakawalan
 ‘The old man Kakawalan kept massaging him.’
- c. *tu=pu-kiping-ay na namali kay baeli*
 3.GEN-CAUS-clothes-LV PL.NOM my.father and my.brother
 ‘They have my father and brother wear the traditional clothes.’
- d. *tu=pulabus-ay tengerD kana barubaru*
 3.GEN-almost-LV <AV>kill PL.OBL Barubaru
 ‘They were almost killed by Barubaru people.’

¹² Although both *buwang* and *walak* are marked by oblique case in the following sentences, *walak* has a genitive bound pronoun cross-reference with it, but *buwang* doesn’t. Thus, the non-subject agent, even though marked oblique, still behaves differently from the oblique arguments.

- a. *tu=laseD-aw kana buwang i temutaw*
 3.GEN-hide-PV DF.OBL hole SG.NOM his/her/their.grandparent
 ‘He hid his grandmother in the hole.’
- b. *tu=paDek-aw i temutaw kana walak*
 3.GEN-carry.on.back-PV SG.NOM his/her/their.grandparent DF.OBL child
 ‘The child carried his grandmother on back.’

Location nouns refer to places (9a), directions (9b), and locative relations¹³ (9c).

- (9) a. *m-uka=ku i taihok*
 AF-go-1SG.NOM LOC Taipei
 ‘I went to Taipei.’
- b. *ma-kiteng i timuL*
 AV-small LOC south
 ‘It is small in the south.’
- c. *ku=atel-anay na paysu i nguwayan kantaw*
 1SG.GEN-throw-IV DF.NOM money LOC front 3.OBL
 ‘I threw the money in front of him.’

There is only one locative noun phrase marker in Puyuma, and location nouns will usually co-occur with it, whether being a subject (10a) or not (10b).

- (10) a. *ku=seLap-ay i sabak*
 1SG.GEN-sweep-LV LOC inside
 ‘I swept the inside.’
- b. *sagar=ku i baLangaw*
 like.AV-1SG.NOM LOC Taitung
 ‘I like Taitung.’

In addition to place names and relator nouns, some common nouns which have locative connotations, such as *ruma* ‘house’, *kaLi* ‘river’, *LangiT* ‘sky’, *ine* ‘sea’, *Dekal* ‘village’ can sometimes be marked by *i*. For instance:

- (11) a. *mi-riwanes na LangiT*
 AV-rainbow DF.NOM sky
 ‘The sky has a rainbow.’
- b. *ulaya a riwanes i LangiT*
 exist ID.NOM rainbow LOC sky
 ‘There is a rainbow in the sky.’

Both sentences are grammatical, and the voice marking and case marking suggest that *LangiT* in (11a) is the core argument, while in (11b) it is more like an adjunct.

When a genitive pronoun is preceded by *i*, it refers to their place.

- (12) *ta=ateD-anay i kantaw*
 1.INC.GEN-send-BV LOC their.place
 ‘We sent them back to their place.’

¹³ Locative relations are often encoded by ‘relator nouns’ in Puyuma. A relator noun is a noun which does not name an entity but indicates the spatial relation between two entities. For example, in English, ‘front’ in the phrase ‘in front of’, and ‘back’ in the phrase ‘in back of’ are relator nouns. Relator nouns in Puyuma are *LikuDan* ‘back, after’, *nguwayan* ‘front, before’, *isaT* ‘above, up’, and so on.

3.2 Case marking of pronouns

There are both bound and free pronouns in Puyuma. The bound pronouns are summarised in Table 2.

Table 2: The bound pronouns in Puyuma

Number/Person		Singular			Plural			
		1 st person	2 nd person	3 rd person	1 st (INC)	1 st (EXC)	2 nd person	3 rd person
Nom	subject	-ku	-yu	∅	-ta	-mi	-mu	∅
	possessor of subject	ku-	nu-	tu-	ta-	niam-	mu-	tu-
Gen		ku-; ti-	nu-	tu-	ta-	mi-	mu-	tu-

There are both genitive and nominative bound pronouns. Nominative bound pronouns denote the grammatical subject and are usually cliticised to the first element¹⁴ in a sentence, as exemplified in (13a). Genitive bound pronouns can denote a non-subject actor, as shown in (13b), or the possessor of a noun when this noun is the grammatical subject in the sentence, as in (13c).

- (13) a. *sa-sanga'*=*ta* *Da* *dedereran* *i*,
 Ca<AV>produce-1PL.INC.NOM ID.OBL spear TOP
 'When we were making spears,'
- b. *ta=tusuk-aw* *kana* *dedereran*
 1PL.INC.GEN-pierce DF.OBL spear
 'We pierced them (the monkey) with the spears.'
- c. *an ma-kiteng ta=Takuban* *i*, *sayma*
 if AV-small our.NOM-Takuban TOP few
 'If our Takuban is small, few (bamboos are needed).'

Free pronouns occur in neutral, possessive, and oblique forms, which are summarised in Table 3.

The neutral free pronouns usually appear in the topic position or in cleft sentences, as shown in the sentences below:

- (14) a. *taita* *i*, *ka<a>Du=ta* *i* *taihok*
 1INC.NEU TOP <PROG>live-1PL.INC.NOM LOC Taipei
 'As for us, we are living in Taipei.'
- b. *amau taytaw na sa-senay*
 COP 3SG.NEU NOM Ca<AV>-sing
 'It is him who was singing.'

¹⁴ The first element is not restricted to a verbal element; it can be a negator or a nominal predicate.

Table 3: The Free Pronouns in Puyuma

Number/Person Cases	Singular			Plural			
	1 st person	2 nd person	3 rd person	1 st (incl.)	1 st (excl.)	2 nd person	3 rd person
Neutral	<i>kuiku</i>	<i>yuyu</i>	<i>taytaw</i>	<i>taita</i>	<i>mimi</i>	<i>muimu</i>	\emptyset
Possessive	possessor of subject	<i>nanku</i>	<i>nantaw nantu</i>	<i>nanta</i>	<i>naniam</i>	<i>nanemu</i>	<i>nantaw nantu</i>
		possessor of non-subject	<i>kanku kananku</i>	<i>kantu kanantu</i>	<i>kanta kananta</i>	<i>kaniam kananiam</i>	<i>kanemu kananemu</i>
Oblique	indefinite		<i>Daku Dananku</i>	<i>Danu Dananu</i>	<i>Data Dananta</i>	<i>Daniam Dananiam</i>	<i>Danemu Dananemu</i>
		<i>kanku</i>	<i>kanu</i>	<i>kantaw</i>	<i>kanta</i>	<i>kaniam</i>	<i>kanemu</i>

The neutral pronouns are also used as the reply for the interrogative sentences starting with *i manay* ‘who’.

- (15) a. *i manay na Da-Dua i baLangaw*
 NOM who NOM Ca-come LOC Taitung
 ‘Who is coming to Taitung.’
- b. *kuiku*
 1SG.NEU
 ‘Me.’

We might suspect from the above examples that the neutral pronouns are actually nominative pronouns. However, the third person pronoun *taytaw* can appear in declarative sentences beginning with verbal predicates, and the participant it manifests can be the nominative argument, as in (16a) and (16b), or the genitive argument, as in (16c). In those cases, the free pronoun is optional, and its function is basically emphatic.

- (16) a. *ku=babuLas-ay Da kabung taytaw*
 1SG.GEN-lend-LV ID.OBL hat 3.NEU
 ‘I lent him a hat.’
- b. *sa’eru-’eru misasa la taytaw*
 laugh.AV-RED one ASP 3.NEU
 ‘She was laughing alone.’
- c. *tu=pa-a-arum-ay nu-kiruan taytaw*
 3.GEN-CAUS-PROG-dry-LV your.NOM-clothes 3.NEU
 ‘He is drying your clothes.’

The possessive free pronouns refer to possessors of nouns. If the noun is the grammatical subject, then it is preceded by one set of pronouns, which assign it the nominative case, as shown in (17a); if the noun is not the grammatical subject, then another set of possessive pronouns assign it the oblique case, as in (17b).

- (17) a. *tu=reTa-anay nantu basak kana ma’iDangan*
 3.GEN-put.down-BV their.NOM bag DF.OBL old.people
 ‘The elders put down their bags.’
- b. *tu=betbet-anay la kantu basak*
 3.GEN-bind-IV ASP their.OBL bag
 ‘They bound their bags.’

It should be noticed that the possessors of non-subjects are further subcategorised into two classes in terms of the definiteness of the possessed nouns. For example:

- (18) a. *aDi ka-la-laDam kananku in-aLak-an*
 NEG ka-PROG-know my.DF.OBL REAL-take-NMSR
tu=ngaLad kan Luba’ib
 his.NOM-name SG.OBL Luba’ib
 ‘He would not know (the story about) my taking Luba’ib’s name.’

- (21) a. *tu=bau-bau-i-aw=ku m-uka i takesi-an*
 3.GEN-RED-push-PV-1SG.NOM AV-go LOC study-NMSR
 ‘She urged me to go to school.’
- b. *tu=gingaging-aw (na Takuban) mu-Tereb*
 3.GEN-shake-PV DF.NOM meeting.place AV-fall
 ‘It (the wind) shook (the meeting place) down.’

Arguments other than the nominative arguments and genitive arguments cannot be controllers. For example in (22a) the oblique argument *bekaLan* ‘new rice’ cannot be the controller of the verb *m-u-ami*; in (22b) the controller is the agent *tu-*, not the oblique argument *LangeTi*. Similarly, in (22c), the controller is *-ku* not *paDekan*.

- (22) a. *aDi=ta ka-kasu Da _____ bekaL-an m-u-ami*
 NEG-1PL.NOM Ca<AV>bring ID.OBL new-NMSR AV-go-north
 ‘We are not bringing new (rice) to the north.’
- b. *tu=pa-laDam-aw Da _____ LangeTi pa-karun*
 3.GEN-CAUS-know-PV ID.OBL stick CAUS-work
 ‘They used the stick to teach them to work.’
- c. *ma-tara-paDek=ku Da _____ paDekan mu-languy*
 AV-take-carry.on.back-1SG.NOM ID.OBL backpack AV-swim
 ‘I swam with a backpack on my back.’

We have demonstrated that both nominative and genitive arguments can be manifested as enclitics on verbs, and only they can be controllers. So they are core arguments. And while each non-actor voice sentence has two core arguments, actor voice sentences have only one core argument. This suggests that actor voice sentences are intransitive.

From the semantic perspective, Hopper and Thompson (1980:251–253) suggest transitivity should be viewed as a continuum rather than a binary property. They propose a number of parameters of transitivity, as stated in (23).

	High	Low
Participants	2 or more participants, A and O	1 participant
Kinesis	action	non-action
Aspect	telic	atelic
Punctuality	punctual	non-punctual
Volitionality	volitional	non-volitional
Affirmation	affirmative	negative
Mode	realis	irrealis
Agency	A high in potency	A low in potency
Affectedness of O	O totally affected	O not affected
Individuation of O	O highly individuated	O non-individuated

Hopper and Thompson go on to add that the referents of nouns with the properties, such as being proper, human/animate, concrete, singular, count, and referential/definite (as

opposed to common, inanimate, abstract, plural, mass, and non-referential), are more highly individuated.

In Puyuma texts, in independent clauses, if the undergoer is definite, then, it will be chosen to be the subject, and the sentence will be manifested as non-actor voice. For example:

- (24) a. *puka=ku Da apuT*
 put.AV-1SG.NOM ID.OBL flower
 ‘I put some flowers.’
- b. *ku=puka-ay na apuT Da pakeLing*
 1SG.GEN-put-LV DF.NOM flower ID.OBL hook
 ‘I added some hooks to the wreath.’

Thus, from semantic perspective, non-actor voice sentences are also higher in transitivity.

5 Ergativity¹⁶

Following Dixon (1994), ergativity in this paper is viewed as a linguistic feature which treats the only argument in an intransitive clause (S, hereafter) and the undergoer in a transitive clause (O, hereafter) in the same way. Ergativity is in opposition to accusativity, which refers to a system in which S is treated the same as the actor in a transitive clause (A, hereafter.) Thus, the two systems can be illustrated by the diagram below:

Accusative	Ergative
A	A
S	S
O	O

Let us examine the marking of arguments in Puyuma sentences. Sentence (25a) is an example of intransitive sentence, and sentence (25b) is transitive.

- (25) a. *aru ki<a>naTay i temutaw*
 will <IRR>die SG.NOM his.grandparent
 ‘His grandmother is going to die.’
- b. *tu=kiumal-ay i temutaw*
 3.GEN-ask-LV SG.NOM his.grandparent
 ‘He asked his grandmother.’

The only argument in the intransitive sentence (25a) and the O argument in (25b) bear the same marking, while the A argument in (25b) is marked differently. This means Puyuma is syntactically ergative.

Croft (2001:155) proposes a Subject Construction Hierarchy, which defines ‘an implicational scale such that for any construction on the scale, if the construction patterns

¹⁶ Ergativity in this paper does not refer to ergative case markings; rather we discuss ergative alignment.

ergatively, then all the constructions to the right of it on the scale also pattern ergatively; if the construction patterns accusatively, then all the constructions to the left of it on the scale also pattern accusatively’.

- (26) The Subject Construction Hierarchy
coordination < purposive < relativisation < verb agreement < case marking

We have demonstrated that Puyuma patterns ergatively in simple constructions, and since Puyuma displays no verb agreement, we move on to investigate relativisation.

It has been demonstrated in Ross and Teng (2003) that relative clauses in Puyuma pattern accusatively. Let us start from the role the domain noun plays in the matrix clause. It is found that NP_{mtx} can be actor or non-actor, being the grammatical subject or not. In other words, there is no constraint on what role NP_{mtx} plays, semantically or syntactically. On the other hand, two different strategies are utilised according to whether the NP_{rel} is an actor or not. If the NP_{rel} plays the role of actor (either S or A in Dixon’s terms), then the RC is manifested as a clause; on the other hand, if the NP_{rel} is an undergoer (O in Dixon’s terms), then the RC is nominalised. Examples are given in (27) and (28). This suggests that A and S pattern together, while O uses another construction. So relative clauses in Puyuma pattern accusatively.

- (27) a. NP_{rel} as S
kiyuma=ta Da mialup {Da mi-a-kelep kaDini}
ask.AV-1PL.NOM OBL god OBL AV-ASP-reside here
‘We asked gods who reside here.’
- b. NP_{rel} as A
iDi {na aDi kiberay kan tayban Da bini}
this.NOM DF.NOM NEG get.give SG.OBL Tayban ID.OBL seed
‘This (person), who didn’t get seeds from Tayban ...’
- (28) a. NP_{rel} as O (Patient)
ala amuna saDu {tu=T<in>ekeL(-an)¹⁷} na asi
maybe because a.lot 3NOM.GEN-<REAL>drink(-NMSR) DF.NOM milk
‘Maybe because the milk he drank is a lot.’
- b. NP_{rel} as O (Instrument)
tu=laseD-aw i TaLu-TaLun {na in-abak-an
3.GEN-hide-PV LOC Red-grass NOM REAL-contain-NMSR
kana walak} na paDekan
OBL child NOM basket
‘She hid the basket which was used to contain the child.’
- c. NP_{rel} as O (Location)
{nantu p<in>uatel-an kana tangtang}
3NOM.GEN drop<REAL>drop-NMSR OBL box

¹⁷ Although the infix <in> serves mainly as a modal marker denoting the realis modality, it nevertheless occurs only in the nominalised construction in Puyuma. In those cases, the marker -an can be omitted.

na dare i, mi-ngaLad la D matang ka
 NOM soil TOP AV-name ASP OBL Matang say<AV>say
 ‘The soil (place) in which the box was dropped, people said it gets the name Matang.’

According to the Subject Construction Hierarchy, this implies that all constructions to the left of relativisation on the scale should also pattern accusatively. This implication needs further investigation. However, we do find that in complex constructions, such as serial verb constructions, it is the agent which is the grammatical subject, even if the undergoer is definite. For example:

- (29) a. *kurudung=ku mi-eDeng kana tutwi*
 lean.against-1SG.NOM AV-sleep DF.OBL puppy
 ‘I leant against the puppy to sleep.’
- b. *mu-laseD i Taru-Tarun pa-su’su’ kantu walak*
 AV-hide LOC RED-grass CAUS-breast her.OBL child
 ‘She hid in the field to breast-feed her child.’

To sum up, Puyuma simple independent clauses are morphosyntactically ergative. In complex constructions, such as serial verb constructions and relative clauses, it is syntactically accusative.

References

- Cauquelin, Josiane, 1991, The Puyuma language. *Journal of the Royal Institute of Linguistics and Anthropology* 147:7–60.
- Croft, William, 2001, *Radical construction grammar: syntactic theory in typological perspective*. Oxford: Oxford University Press.
- Dixon, R.M.W., 1994, *Ergativity*. Cambridge: Cambridge University Press.
- Dryer, Matthew S., 1997, Are grammatical relations universal? In Joan Bybee, John Haiman, Sandra A. Thompson, eds *Essays on language function and language type*, 115–143. Amsterdam: John Benjamins Publishing Company.
- Hopper, Paul J. and Sandra Thompson, 1980, Transitivity in grammar and discourse. *Language* 56:251–299.
- Huang, Lillian M., 1995, *A study of Mayrinax syntax*. Taipei: The Crane Publishing Company.
- 2000a, *A reference grammar of Puyuma*. (In Chinese). Taipei: Yuan-liou Publishing Company.
- 2000b, Verb classification in Maryrinax Atayal. *Oceanic Linguistics* 39:364–390.
- Kess, Joseph F., 1976, Reconsidering the notion of focus in the description of Tagalog. In Nguyen Dan Liem, ed. *South-east Asian linguistics* vol.2, 173–186. Canberra: Pacific Linguistics.
- Kroeger, Paul, 1993, *Phrase structure and grammatical relations in Tagalog*. Stanford: CSLI.

- Li, Paul Jen-kuei, 1995, Formosan vs. non-Formosan features in some Austronesian languages in Taiwan. In Paul Jen-kuei Li et al., eds *Austronesian Studies Relating to Taiwan*, 651–681. Taipei: Academia Sinica.
- Liao, Hsiu-chuan, 2004, Transitivity and ergativity in Formosan and Philippine languages. PhD dissertation, University of Hawai'i, Honolulu.
- Payne, Thomas E., 1997, *Describing morphosyntax: a guide for field linguists*. Cambridge: Cambridge University Press.
- Ross, M.D., 2002, History and transitivity of western Austronesian voice. In F. Wouk and M.D. Ross, eds *The history and typology of Western Austronesian voice systems*, 17–62. Canberra: Pacific Linguistics.
- Ross, M.D. and Stacy Fang-ching Teng, 2003, The notion of “verb” in Puyuma: a Construction Grammar approach. Paper presented in the Symposium on the Notion of ‘Verb’ in Formosan Languages, Academia Sinica, Taipei.
- Starosta, Stanley, 1997, Formosan clause structure: transitivity, ergativity, and case marking. In Chiu-yu Tseng, ed. *Chinese languages and linguistics vol.4: Typological studies of languages in China*, 125–154. Taipei: Academia Sinica.
- 1999, Transitivity, ergativity and the best analysis of Atayal case marking. In Elizabeth Zeitoun and Paul Jen-kuei Li, eds *Selected papers from the Eighth International Conference on Austronesian linguistics*, 371–392. Taipei: Academia Sinica.
- Tan, Cindy R., 1997, A study of Puyuma simple sentences. MA thesis. Graduate Institute of English, Taipei: National Taiwan Normal University.
- Teng, Stacy F., 1997, Complex sentences in Puyuma. MA thesis. Graduate Institute of English, Taipei: National Taiwan Normal University.
- Zeitoun, Elizabeth, L. Huang, M. Yeh, A. Chang and J. Wu, 1996, The temporal, aspectual, and modal systems of some Formosan languages. *Oceanic Linguistics* 35:21–56.

7 *Voice, ergativity and transitivity in Tagalog and other Philippine languages: a typological perspective*

MASUMI KATAGIRI

1 Introduction

In recent studies of Philippine languages, there have been a number of analyses that treat these languages as ergative in type (e.g. Payne 1982; Cooreman, Fox & Givón 1984; Gerds 1988; De Guzman 1988; Blake 1990; Mithun 1994). This paper re-examines and discusses some of the issues surrounding the ergative hypothesis from a typological perspective, and proposes an alternative view that will lead to a unified account of the Philippine-type voice system.¹

In Philippine languages, each major clause has one NP in ‘primary relation’, which in Tagalog is marked by prepositional *ang* (or *si* with singular personal names, *sina* with plural personal names). Henceforth I label such markers ‘ANG’. The semantic role of the ANG-marked NP is cross-referenced on the verb by an affix.² In transitive clauses in which an agentive nominal A and a patient nominal P are present, then, there are at least two possible sentences: those in which A is marked by ANG, and those in which P is marked by ANG. When other NPs with different semantic roles, such as Beneficiary, Direction or

¹ This paper is a revised version of ‘Typological positions of Philippine-type languages and their voice systems’, presented at the 9th International Conference on Austronesian Linguistics, Canberra, January 8–11, 2002. The work contained here was partly supported by a grant from the Japan Society for the Promotion of Science for the Grant-in-Aid for Encouragement of Young Scientists ‘Typological studies on voice and ergativity in Philippine-type languages’ (No. 12710283).

² The status of ANG-marked NPs has long been under debate in Philippine linguistics — specifically, whether they are subject or topic — but I simply refer to them as ANG-marked NPs or ANG-NPs, since the issue is beyond the scope of this paper. Sama, a language spoken in southern parts of the Philippines, does not have a marker corresponding to *ang*, but verbal affixes tell us which nominal in a clause is in primary relation.

Recipient, Locative, and Instrumental are also present, each one of them can basically be marked by ANG, too. Since it is now generally agreed that such alternations are those of voice, I refer to the types of clauses and verbal forms as Actor Voice (AV) clauses and forms, Patient Voice (PV) clauses and forms, and so forth.³

In the ergative analysis of Philippine languages, PV clauses are treated as basic transitive clauses, while AV clauses are analysed as their detransitivised versions. Since the sole argument of intransitive sentences (S) is also marked by ANG, the ANG-marked NP is labelled as absolutive, the PV clauses are construed as ergative (active), and the AV clauses as antipassive. The following Tagalog sentences from De Guzman (1988:323–324) illustrate the point.⁴ (Henceforth, unless otherwise indicated, quoted examples are from Tagalog and their glosses and translations are original).

- (1) a. *Nagsalita ang babae.*
 spoke S woman
 ‘The woman spoke.’
- b. *Lulutuin ng babae ang manok.*
 will.cook A woman P chicken
 erg abs
 ‘The chicken will be cooked by the woman.’
 ‘The woman will cook the chicken.’
- c. *Magluluto ang babae ng manok.*
 will.cook A woman P chicken
 abs gen
 ‘The woman will cook (a) chicken.’

From the ergative viewpoint, the PV sentence (1b) is treated as a basic transitive and the AV sentence (1c) is considered as an antipassivised version of (1b). Most of the arguments for regarding PV sentences as basic originate from the strong patient-preference or patient-orientedness observed in Tagalog and other Philippine languages in general — that is, preference for the PV construction over the AV construction in transitive sentences. Cooreman, Fox and Givón (1984:17), for example, describe the PV construction in Tagalog as being ‘by far the commonest’, based on their statistical data.

Although patient-preference or patient-orientedness is an earmark for an ergative system, one cannot (and should not) classify an entire grammar as ergative on the basis of one characteristic alone, since most of the languages have both accusative and ergative characteristics within their grammars. In particular, no matter how predominant a certain clause type may be over another, there should be some morphosyntactic properties that

³ The alternations are traditionally called ‘actor focus’ or ‘actor topic’, ‘goal focus’ or ‘goal topic’, and so forth, but these terms, especially ‘focus’ and ‘topic’, do not reflect their general use.

⁴ Abbreviations: A agentive argument of a transitive clause; ABS/abs absolutive; ACC accusative; ADV adverbial; ANG *ang* for common nouns, *si* for singular personal names and *sina* for plural personal names in Tagalog, and their equivalents in other Philippine languages; AOR aorist; AP antipassive; AV actor voice; cho chômeur; Cont contemplated; DAT dative; def definite; ERG/erg ergative; GEN/gen genitive; Imp imperfective; Inf infinitive; INSTR instrumental; LK linker; LOC locative; NOM nominative; OBL/obl oblique; P patient argument of a transitive clause; Perf perfective; PL plural; PRES present; PV patient voice; S subject (sole argument) of an intransitive clause; SG singular; spec specific; 1 first person; 2 second person, or direct object relation; 3 third person.

indicate its unmarkedness. Moreover, even if the unmarkedness of a certain clause type is evidenced morphologically, this does not automatically provide support for the claim that its counterpart is derived from it. The latter may be a clause type governed by a different principle.

In this paper, I show that a careful study of morphological and syntactic properties of AV verbal forms and clauses in comparison with their PV counterparts indicates that the AV construction is not derived from the PV construction, and hence, not an antipassive construction. Rather, on the assumption that the Philippine-type voice system represents a case of split system, it is shown that Tagalog may be taken to represent a very early stage of split, arising through the grammaticalisation of a fluid system.

2 Actor voice construction as antipassive

Based on studies of a wide variety of languages, Dixon (1994:146) puts forward criteria by which a process or a construction may be recognised as antipassive:

- (2) a. the construction/process applies to an underlying transitive clause and forms a derived intransitive;
- b. the underlying A NP becomes the S of the antipassive;
- c. the underlying O [=P] NP acquires a peripheral function, being marked by a non-core case, preposition, etc.; this NP can be omitted, although there is always the option of including it;
- d. there is some explicit formal marking of an antipassive construction (generally, by a verbal affix or else by a periphrastic element in the verb phrase although it could be marked elsewhere in the clause).

To identify a certain clause or construction as ‘antipassive’, one should carefully see if it exhibits all or most of the properties shared by prototypical antipassive constructions in other languages. To call a certain construction antipassive on the basis of only one or two properties is quite dangerous because that property may have no direct bearing on the issue in question (see below).

In most of the studies which claim or assume Philippine languages as ergative, there have been extensive discussions of why PV sentences should be analysed as basic transitive and thus ergative — based mainly on patient-orientedness as I mentioned above — but, with a few exceptions, most of the scholars are then *forced* to classify AV sentences as antipassive without presenting convincing evidence. In the following subsections, I examine Philippine AV sentences in terms of the criteria above and argue against the antipassive analysis for such sentences. I will first examine the morphological markedness of AV verbs ((2d) above) for reasons that will be clear from the following discussion.

2.1 Morphological markedness of AV verbs

As Dixon (1994) suggests, I assume criterion (2d), i.e. explicit formal marking on the verb, should always be maintained for a certain sentence to be identified with antipassive, given that it also satisfies some or all the other criteria. This is a plausible assumption

because we can find a number of constructions in accusative languages that satisfy all the criteria above except (2d). For example, such English intransitive sentences as ‘She has already eaten’, ‘She is cooking’, ‘She drinks’, ‘Speed kills’, etc. are not generally considered to be antipassive.⁵ Although these sentences are syntactically intransitive (satisfying (2a)), the agent is an intransitive subject (satisfying (2b)), and the patient is omitted (satisfying (2c)), there is no explicit formal marking on the verbs or anywhere else in the clauses that is distinct from their transitive counterparts, i.e. ‘She has already eaten a meal’, ‘She is cooking a meal’, ‘She drinks alcohol’, ‘Speed kills people’, etc.

In the same vein, Turkish and some Indic languages have another type of construction that satisfies (2a), (2b) and (2c), but not (2d). In Turkish, a typical accusative language, an inanimate and indefinite direct object does not receive an accusative case, as shown in (3b) below:

(3) Turkish

- a. *Ben kitab-ı al-dı-m.*
I book-ACC buy-PAST-1SG
‘I bought the book.’
- b. *Ben kitap al-dı-m.*
I book buy-PAST-1SG
‘I bought a book.’ (I did book-buying.)
- c. *Kitab-ı, ben al-dı-m.*
book-ACC I buy-PAST-1SG
‘The book, I bought.’
- d. * *Kitap, ben al-dı-m.*
book I buy-PAST-1SG
(For: ‘A book, I bought.’)

It is well attested that the caseless object in (3b) *kitap* ‘book’ is morphologically incorporated into the verb and loses its termhood as direct object, as indicated by inapplicability of scrambling (3d), and by other tests such as adverb insertion, relativisation, etc. (Kuribayashi 1989). And incorporation in Turkish (and in other languages for that matter) is a valency-decreasing device which derives intransitive sentences when it is applied to transitive clauses. Thus, sentences such as (3b) satisfy all the criteria in (2) except (2d), i.e. explicit formal marking on the verb, but they are not construed as antipassive.

Moreover, the incorporation phenomenon is widely observed also in ergative-type languages, and if it exists, it occurs independently of the antipassive construction. The following sentences from Chukchee illustrate the point: (4a) is an antipassive sentence, in which the verb is overtly marked for antipassive, and (4b) is a sentence in which the object is incorporated into the verb.

⁵ But see Heath (1976).

(4) Chukchee (Kozinsky, et al. 1988:667)

- a. *atlʔa ine-nni-gʔi məčək-w-a.*
 mother+ABS AP-sew-3SG+AOR shirt-INSTR
 ‘The mother sewed the shirt.’
- b. *atlʔa məčək-wə-nni-gʔi.*
 mother+ABS shirt-sew-3SG+AOR
 ‘The mother sewed the shirt.’

Hence, I assume that explicit formal marking on verbs that indicates they are derived from their transitive counterparts is the most essential property in identifying a certain construction as antipassive, given that it also satisfies other criteria. In the case of Philippine languages, then, one must show that AV forms are morphologically more complex than and derived from their PV counterparts in order to analyse AV sentences as antipassive. De Guzman (1979, 1992) is the first, to my knowledge, that drew attention to the issue. In her attempt to claim that PV forms function as basic transitives, she shows that certain classes of AV verbs in Tagalog, specifically those marked by an AV affix *mag-*, are more complex than their PV forms.

Tagalog verbs inflect for voice and aspect/mood, and major action verbs take either *-um-* or *mag-* for AV, and *-in*, *i-* or *-an* for PV.⁶ Although there are some correlations between the semantic nature of a verb and a form of affixes it takes for each voice, it is largely unpredictable which particular voice affix a given verb takes. For example, *luto* ‘cook’ takes either *-in* or *i-* for PV (*lutuin/iluto*), but it takes *mag-* rather than *-um-* for AV (*magluto*, not **lumuto*). For the present discussion, I will simply refer to the verb classes as UM-verbs, MAG-verbs, IN-verbs, I-verbs, and AN-verbs. The perfective aspect characterises an event as completed, the imperfective as not completed but begun, and the contemplated as not begun (Schachter & Otanes 1972:66). Reduplication of the first syllable of the root indicates that the action is not completed. The following (5) shows paradigms for verbs *sulat* ‘write’, *bigay* ‘give’, and *tulong* ‘help’. Voice affixes are in boldface in the following paradigms, with $-\emptyset$ indicating that the form is unmarked for voice (see below).

(5)	Infinitive	Realis		Irrealis
		Perfective	Imperfective	Contemplated
		(+begun +completed)	(+begun -completed)	(-begun -completed)
<i>sulat</i> (AV)	<i>s-um-ulat</i>	<i>s-um-ulat</i>	<i>s-um-u-sulat</i>	<i>su-sulat-\emptyset</i>
<i>bigay</i> (AV)	<i>mag-bigay</i>	<i>nag-bigay</i>	<i>nag-bi-bigay</i>	<i>mag-bi-bigay</i>
<i>sulat</i> (PV)	<i>sulat-in</i>	<i>s-in-ulat-\emptyset</i>	<i>s-in-u-sulat-\emptyset</i>	<i>su-sulat-in</i>
<i>bigay</i> (PV)	<i>i-bigay</i>	<i>i-b-in-igay</i>	<i>i-b-in-i-bigay</i>	<i>i-bi-bigay</i>
<i>tulong</i> (PV)	<i>tulong-an</i>	<i>t-in-ulung-an</i>	<i>t-in-u-tulong-an</i>	<i>tu-tulong-an</i>

The points to be discussed in this subsection are: the morphological structure of an AV affix *mag-*, morphological unmarkedness of PV verbs which take a PV affix *-in*, and occurrence of unaffixed PV verb.

⁶ The affix *-an* is often classified as Locative or Directional voice affix.

2.1.1 Morphological structure of MAG- verbs

Although an AV affix *mag-* has been treated as one unit in the standard analysis, De Guzman (1979, 1992) analyses it as an AV affix *-um-* plus a gerundive-forming affix *pag-*. Her arguments are based on the facts that the initial consonant of the affix *mag-* is phonologically close to and a viable alternative form of an AV affix *-um-*, and that the affix *pag-* appears as it is in other non-PV. For example, De Guzman (1992) argues that the root *luto* ‘cook’ inflects for voice as in (6) (in infinitive forms). Each voice affix is boldfaced:

- | | | | | |
|-----|-----------------------------|---------------------------------------|--------------------------|---------------------------|
| (6) | Actor Voice | Patient Voice | Beneficiary Voice | Locative Voice |
| | <i>mag(<um+pag)-luto</i> | <i>lutu-in / i-luto</i> | <i>i-pag-luto</i> | <i>pag-lutu-an</i> |

Thus, PV forms are shown to be simpler compared with the other voice forms in that the PV affix attaches directly to the root while in the other voice forms a voice affix is attached to gerundive stems. The above analysis of the internal structure of MAG- verbs is also supported by Blake (1990) and Himmelmann (1991). De Guzman (1992) provides similar arguments for non-volitive verbs such as *kita* ‘see’, which inflects for voice as PV *makita* (<ma+kita) versus AV *makakita* (<ma+ka-kita). Based on these verb forms in which Patient Voice is shown to be simpler than other voice forms in their morphological structure, De Guzman concludes that PV verbs are unmarked voice forms.

Although the above argument — if it is assumed to be correct — shows that AV verbs of *mag-* class (or other non-PV forms for that matter) are morphologically more complex than their corresponding PV forms, it has no direct bearing on the claim that AV forms are antipassive, as implied by the ergative hypothesis. In general, antipassive forms in ergative languages are not only morphologically more complex than their active (ergative) counterparts, they are *derived* from their active counterparts, just like direct passive verbs in accusative languages are derived from their active counterparts, and its derivation is explicitly marked by verbal affixes or periphrastic elements in the verb phrase. For example, Dyirbal antipassive forms are derived from their active base by affixing *ŋa-y* to their corresponding active (ergative) forms, Yidin^y by affixing *-ji-n*.

- (7) Yidin^y (Dixon 1994:59–60)
- | | | | | | |
|----|--|--------------------|-----------------|-----------------------------|------------------------------------|
| a. | <i>Waguja-ŋgu</i> | <i>jugi-∅</i> | <i>gunda-l.</i> | [Ergative=Basic transitive] | |
| | man-ERG | tree-ABS | cut-PRES | | |
| | ‘The man is cutting a tree.’ | | | | |
| b. | <i>Wagu:ja-∅</i> | <i>gunda-:ji-ŋ</i> | <i>jugi-:l.</i> | <i>(galba:n-da).</i> | [Antipassive=derived intransitive] |
| | man-ABS | cut-AP-Pres | tree-LOC | axe-INST | |
| | ‘The man is cutting a tree (with an axe).’ | | | | |
| c. | (De Guzman 1998) | | | | |
| | <i>Wagu:ja-∅</i> | <i>gunda-l.</i> | | | |
| | man-ABS | cut-PRES | | | |
| | ‘The man is cutting.’ | | | | |

Notice that the structure of the verb in the antipassive sentence (7b) reflects that it is not only more complex than, but also derived from, its ergative counterpart in (7a) *gunda* ‘cut’, with tense inflection being marked at the outmost. It is also important to note that although antipassive sentences like (7b) are syntactically intransitive, they should be clearly

distinguished from non-derived intransitive sentences like (7c), in which a patient nominal is simply omitted, and there is no marking on the verb distinct from that of (7a).

This is not the case in Tagalog as the paradigm for *luto* ‘cook’ in (6) illustrates. If the AV form were an antipassive form derived from its corresponding PV form, we would expect such forms as **maglutuin* or **mag-iluto* for infinitive AV, which never occur in Tagalog verbs. Also, verbal forms in AV sentences with a patient nominal and those without one are not morphologically distinct, as shown in (8).

- (8) a. *Nagluluto ang babae ng manok.*
 AV.Imp+cook ANG woman GEN chicken
 ‘The woman is cooking a chicken.’
- b. *Nagluluto ang babae.*
 AV.Imp+cook ANG woman
 ‘The woman is cooking.’

In fact, most of the ergative proposals for Tagalog and other Philippine languages assume that AV sentences are intransitive, and thus antipassive. However, being intransitive and being antipassive are essentially independent properties of a given construction, as mentioned earlier with regard to English and Turkish examples, and whether the construction in question is intransitive should be evidenced independently. I will discuss the issue separately in §2.2.

2.1.2 Morphological unmarkedness of IN- verbs

The second point concerns morphological unmarkedness observed in IN- verbs in the paradigms (5). As pointed out by some linguists (cf. Blake 1990; Himmelmann 1991; Kroeger 1993a; De Guzman 1998), IN- verbs are unmarked for PV voice in realis (perfective and imperfective) contexts, while UM- verbs are unmarked for AV voice in irrealis (contemplated) contexts. In particular, much attention has been paid to the morphological unmarkedness of realis IN- verbs, which leads to the claim that the PV construction is basic transitive (ergative), from which AV forms are derived (antipassive). Traditionally, the infix *-in-* is taken to be a marker for Patient Voice in realis context, in the same manner that the infix *-um-* marks Actor Voice. But it seems to be generally agreed now that the infix *-in-* is an aspectual marker rather than a voice marker because it is common to all the non-AV forms in realis mood (Blake 1990). Consider, for example, the following paradigm for *bili* ‘buy’, with the infix *-in-* being underlined and voice affixes boldfaced:

(9)	Infinitive	Realis		Irrealis
		Perfective	Imperfective	Contemplated
Actor Voice	<i>bumili</i>	<i>bumili</i>	<i>bumibili</i>	<i>bibili</i>
Patient Voice	<i>bilhin</i>	<i>binili</i>	<i>binibili</i>	<i>bibilhin</i>
Locative Voice	<i>bilhan</i>	<i>binilhan</i>	<i>binibilhan</i>	<i>bibilhan</i>
Beneficiary Voice	<i>ibili</i>	<i>ibinili</i>	<i>ibinibili</i>	<i>ibibili</i>
Instrumental Voice	<i>ipambili</i>	<i>ipinambili</i>	<i>ipinambibili</i>	<i>ipambibili</i>
Causal Voice	<i>ikabili</i>	<i>ikinabili</i>	<i>ikinabibili</i>	<i>ikabibili</i>

It is clear from (9) that the infix *-in-* is common to all the non-AV forms in realis mood. Kroeger (1993a:16, fn.7) notes that an ‘OV [=PV] suffix *-in* and perfective infix *-in-* are etymologically distinct, the OV [=PV] marker *-in* < PAN [=Proto Austronesian] **-ən*, the aspect marker *-in-* < PAN **-in-*, though they have the same shape in Tagalog due to a merger of PAN **ə* and **i* as /i/. Thus, it is plausible to assume on these etymological grounds that the Tagalog infix *-in-* is a realis (perfective and imperfective) marker, while the suffix *-in* is a PV marker, and that the voice affix is absent in realis forms of IN- verbs.

Notice that PV suffixes in some other languages take shapes distinct from perfective or past affixes, which provides clearer support for the above assumption. For example, the corresponding PV suffix is *-en* in Ilokano and *-on* in Cebuano, while the marker for past tense or perfective aspect is *-in-* in Ilokano and *gi-* in Cebuano. In these languages, too, the PV suffix is absent in past tense or perfective aspect. For example, consider the following paradigm for an Ilokano root *bisita* ‘visit’:

- (10) Ilokano
- | | | | | |
|-----------------|-----------------|--------------------|-------------------|-------------------|
| <i>bisitaen</i> | <i>binisita</i> | <i>binisbisita</i> | <i>bisbistaen</i> | <i>bisitaento</i> |
| ‘to visit’ (PV) | ‘visited’ | ‘was visiting’ | ‘is visiting’ | ‘will visit’ |

One may wonder why, then, the infix *-in-* does not occur in AV forms in realis mood. Kroeger (1993a:15) notes that the perfective infix *-in-* is realised as a mutation to an identical form beginning with /n/ in AV forms whose initial prefix begins with /m/ in the infinitive form. That is, the initial consonant of *nagbigay* (‘gave’ AV perfective) above is taken to be a mutation form of the infix *-in-*. For UM-verbs, on the other hand, there is no phonological trace of the infix within Tagalog, which itself is an interesting fact, but we can detect its trace in the corresponding AV forms in Ilokano:

- (11) Ilokano
- | | | | | | |
|--------------|----------------|-----------------|-------------------|-------------------|------------------|
| <i>panaw</i> | <i>pumanaw</i> | <i>pimmanaw</i> | <i>pimampanaw</i> | <i>pumampanaw</i> | <i>pumanawto</i> |
| (root) | ‘to leave’ | ‘left’ | ‘was leaving’ | ‘is leaving’ | ‘will leave’ |

It is phonologically plausible to assume that the infixes *-imm-* (past) and *-im-* (past progressive) are merged forms of a past tense marker /in/ plus AV affix /um/. Thus, the above paradigm of an Ilokano verb seems to provide indirect evidence for the claim that the infix *-in-* in Tagalog also marks (realis) aspect rather than patient voice in general.

The fact that verbs which take *-in* for PV in Tagalog (and the corresponding verb classes in other Philippine languages) are unmarked for voice in realis mood leads some linguists to claim that PV forms are basic transitive forms from which AV verbs are derived (cf. Blake 1990; De Guzman 1998). One of the problems with this line of argument is that it overgeneralises this morphological argument to the entire grammar. Firstly, notice such a statement as ‘PV forms are morphologically unmarked for voice’ can be true only for one class of verbs, i.e. IN- verbs. The other PV verb classes, i.e. I- verbs and AN- verbs are always marked for voice. It does not provide enough evidence to the claim that PV forms are unmarked and basic on the basis of one class of verbs.

Secondly, IN- verbs are unmarked for voice only in realis mood (i.e. perfective and imperfective aspects). In irrealis mood (contemplated aspect), they are marked for PV, while UM- verbs in turn are unmarked, e.g. *susulat-ø* (AV) versus *susulat-in* (PV) ‘will write’. With the line of argument above, then, we would be forced to make the dubious claim that in irrealis mood AV forms are morphologically unmarked and thus are the basic forms from which PV forms are derived, while in realis mood it is the other way around.

Thirdly, since the voice affix of IN- verbs in realis mood is zero, it does not provide sound evidence for a claim that the corresponding AV verbs are derived. That is, we have no morphological clue to determine whether the latter are derived from PV forms or not.

Thus the morphological unmarkedness of one class of PV verbs in a limited aspectual context does not give sound evidence for the claim that PV forms are basic transitive (ergative) and AV forms are derived (antipassive). It is critical to note here, though, that this is not to deny that PV sentences with IN- verbs are the most typical transitive clauses. In particular, some of the problems mentioned above are obviously related to transitivity in the sense of Hopper and Thompson (1980), which I clarify in §3. The main problem of concern here is the treatment of AV sentences as antipassive.

2.1.3 Unaffixed PV verbs

Similar problems arise with arguments relating to the occurrence of unaffixed PV verbs. As pointed out by Cena (1977:14–15), Tagalog has a class of verbs that require neither voice nor aspect affixes, and those ‘bare’ verbs function as PV verbs, as illustrated by the following examples (glosses and translations are mine):

(12) (Cena 1977:14–15)

- a. *Hawak/Hinahawakan ni John ang libro.*
hold/PV.Imp+hold GEN John ANG book
‘John holds the book.’
- b. **Hawak si John ng libro.*
(For: ‘John holds a book.’)
- c. *Humahawak si John ng libro.*
AV.Imp+hold ANG John GEN book
‘John holds a book.’

A verb that can occur without being marked for aspect and voice can function only as PV, as in (12a), while an AV verb must be inflected for voice and aspect, as in (12c). The occurrence of unaffixed verbs which can only function as PV verbs leads some linguists to claim that PV verbs are unmarked and basic, and that AV verbs are derived from them. However, only a handful of verbs can occur unaffixed, and they can function as PV verbs only in realis contexts. Other such verbs included in Cena’s (1977) and De Guzman’s (1992) lists are *kuha* ‘get’, *dala* ‘carry’, *taban* ‘hold’, *akay* ‘lead’, *pasan* ‘carry on one’s shoulders’, *kipkip* ‘carry under one’s arm’, *sakop* ‘conquer’, and *ari* ‘own, possess’. There seem to be slight semantic similarities among these verbs, such as possession or conveyance, but it is not clear yet why these verbs do not require voice and aspect affixes. More interestingly, Himmelmann (1999) points out that these unaffixed verbs are used only in realis context, and that in irrealis mood it is AV verbs that can appear unaffixed. For example:

(13) (Bloomfield 1917:221 quoted by Himmelmann 1999)

- Hampas*[=*Humampas*] *na kayo, mga bata, sa mga langgam.*
blow [AV.Inf+blow] now 2PL PL child LOC PL ant
‘Whip at the ants, boys.’

Overall, then, these facts about morphological complexity and unaffixed roots do not support the antipassive analysis of AV sentences since PV verbs are morphologically unmarked in certain contexts, but AV verbs are unmarked in other contexts. As mentioned earlier, all the problems in the ergative analysis discussed in this section boil down to the fact that AV verbs are treated as antipassive forms derived from PV verbs. As we have seen in this section, AV verbs in Tagalog are in no way derived from PV verbs. In other words, the PV/AV alternation is not that of active/antipassive.

2.2 AV sentences as intransitive

It is clear from the above discussion that while a certain class of AV verbs, i.e. MAG-verbs, can be analysed as being morphologically more complex than their corresponding PV verbs in certain contexts, AV forms including MAG- verbs are in no way derived from their corresponding PV forms. On the other hand, past works have shown that AV sentences in Tagalog and other Philippine languages encode events that are less transitive than their corresponding PV sentences in terms of individuation (cf. Hopper & Thompson 1980). The critical point is how productively such transitivity is reflected in the grammar. In this section, I discuss the syntactic arguments surrounding the issue, i.e. whether AV sentences are grammatically intransitive.

It is not at all straightforward whether a given sentence in Tagalog is grammatically intransitive. In line with Palmer (1994:181), I assume that a sentence is ‘grammatically intransitive’ if P is either absent or demoted to oblique status. In other words, even if a given sentence is ‘semantically transitive’, i.e. two participants are present with A performing an action on P, it is ‘grammatically intransitive’ if P is omitted or demoted to an oblique. Incidentally, ‘semantically intransitive’ sentences, i.e. those without P, are always ‘grammatically intransitive’.

As observed in many languages, accusative or ergative, P is readily omitted for some action verbs such as ‘cook’ and ‘eat’ when the referent of P is generic or understood without any previous context. In ergative languages, since one of the functions of the antipassive construction is to defocus P, such deletion is much more productive, not restricted to such verbs. In the case of Tagalog, however, P of AV sentences is not readily deleted unless its referent is present in the previous context. The following sentences without any previous context in which the referent of P is mentioned are either unacceptable or elliptical:

- (14) a. ?? *Nag-dadala sila.*
 AV.Imp+carry they
 ‘They are carrying (something).’
- b. ?? *B-um-ibili sila.*
 AV.Imp+buy they
 ‘They are buying (something).’
- c. *Namimili (< Nang+bibili) sila (ng pagkain).*
 AV.Imp+buy they GEN food
 ‘They are buying/shopping (food).’

Although P can be omitted in sentences like (14c), the verbs that can take the AV affix *mang-* are restricted to those which typically denote actions that are habitual or recursive, or those performed toward plural referents (cf. Oue 1994).

It is more difficult to determine whether demotion of P to an oblique has occurred in the AV construction. Kroeger (1993a, b) argues against such an analysis on the basis of two tests, namely, adjunct fronting and participial adjunct constructions. Kroeger shows that while adverbial phrases marked by an oblique case *sa* can be fronted, the patient nominal of AV sentences that is marked by a genitive case never undergoes fronting. Compare (15a) and (15b):

- (15) *Nagluluto ang babae ng manok sa kusina.*
 AV.Imp+cook ANG woman GEN chicken OBL kitchen
 ‘The woman is cooking a chicken in the kitchen.’
- a. *Sa kusina nagluluto ang babae ng manok.*
 ‘In the kitchen, the woman is cooking a chicken.’
- b. **Ng manok nagluluto ang babae sa kusina.*
 (For: ‘A chicken, the woman is cooking in the kitchen.’)

Ungrammaticality of (15b) suggests that the genitive-marked patient nominal *manok* ‘chicken’ is not an oblique.

This is further supported by evidence from what Kroeger calls participial adjunct constructions. It is shown that Terms may control the gap in a participial *nang* ‘while, when’ clause. The following examples illustrate that genitive-patients can control a gap of such clauses, and therefore terms.

- (16) (Kroeger 1993a:47)
- Nanghuli ng magnanakaw ang polis [nang pumapasok
 AV.Perf+catch GEN thief NOM police ADV AV.Imp+enter
 ø sa bangko].
 OBL bank*
- ‘The police caught a/the thief when entering the bank.’

According to Kroeger, the sentence (16) is ambiguous; either the agent *polis* ‘police’ or the patient *magnanakaw* ‘thief’ can be interpreted as controlling the gap of the participial clause. In contrast, the following example (17) is unambiguous. Only the agent *Juan* ‘Juan’ can be interpreted as the controller, ‘because the dative argument *hari* ‘king’ is an oblique, and hence not a possible controller’ (Kroeger 1993a:48).

- (17) (Kroeger 1993a:48)
- Bumisita si Juan sa hari nang nagiisa ø.*
 AV.Perf+visit NOM Juan DAT king ADV AV.Imp+one
 ‘Juan visited the king alone.’ (Juan is alone)

Thus, Kroeger (1993a:48) concludes ‘the data from adjunct fronting and participial adjuncts provide clear evidence for the termhood of genitive patients’.

At first glance, it appears clear from the above arguments that the patient nominal of an AV clause is not demoted to an oblique, and hence the AV clause is not grammatically intransitive. However, it also follows that a dative-marked patient nominal is demoted to

an oblique since it cannot control the gap of a participial clause as in (17). Also, it is accessible to fronting, as in the following example:

- (18) *Sa kaibigan tumatawag ang babae.*
 OBL friend AV.Imp+call ANG woman
 ‘(To) the/a friend, the woman is calling.’

In fact, a patient nominal of an AV clause is marked either by genitive *ng*, or by oblique/dative *sa*, depending on the verb. Some verbs only subcategorise for a *ng*-patient while other verbs subcategorise for a *sa*-patient, and still others can take either one with differences in specificity: *ng* for non-specific patients, and *sa* for specific patients.

- (19) *Nagluto ang babae ng/*sa manok.*
 AV.Perf+cook ANG woman GEN/OBL chicken
 ‘The woman cooked a/*the chicken.’

- (20) (Cena 1995:29)
*Tumulong si Ben sa/*ng bata.*
 AV.Perf+help ANG Ben OBL/GEN child
 ‘Ben helped the child.’

- (21) (De Guzman 1999)
- a. *Magbabasa ang bata ng libro.*
 will.read-AV 2,abs child 2-cho, obl book
 [-def/-spec]
 ‘The child will read a book.’
- b. *Babasa ang bata sa libro.*
 will.read-AV 2,abs child 2-cho, obl book
 [+spec]
 ‘The child will read of/from the book.’

Since *sa*-marked patients behave in the same manner as ‘pure’ obliques such as locative and temporal adjuncts in terms of fronting and control of a participial clause, one might assume that *sa*-marked patients have been demoted to oblique status, while *ng*-marked patients have not (cf. Cena 1995). However, a serious problem arises when a verb allows either *sa* patients or *ng* patients with differences in specificity, as in (21a) and (21b). On the assumption that *sa* patients are demoted while *ng* patients are not, the situation would be unique from a cross-linguistic point of view: non-specific (*ng*-marked) patients are not demoted to obliques while specific (*sa*-marked) patients are. The universal tendency is the reverse: patients that are low in referentiality are more likely to be put into periphrastic status, such as obliques.

De Guzman (1999) takes a different approach. In her analysis of Tagalog and other Philippine languages in RG framework, both *ng* patients and *sa* patients are demoted, specifically to 2-chômeur, but they are distinguished in terms of syntactic operativity: *ng* patients are syntactically inoperative while *sa* patients are operative. This argument can take account of fronting, but not of control of participial clauses since *sa* patients cannot control the gap of a participial clause, as in (17).

Overall, then, there is no convincing evidence that might indicate that a patient nominal of AV sentences is demoted to an oblique. We have seen that the case is similar with omission of patient nominals: although they are not omitted in general, there are cases where they can be deleted, especially with an AV affix *mang-*. It follows that it is not empirically possible to determine whether AV sentences in Tagalog are grammatically intransitive.

As pointed out by some linguists, on the other hand, AV sentences in Philippine languages in general are less transitive than their PV sentences in terms of transitivity parameters (Hopper & Thompson 1980). For example, the patient of AV sentences is less individuated than that of PV sentences in that the former is typically non-specific and/or indefinite while the latter is mostly definite and specific. This implies that Tagalog has not formally regularised such different degrees of transitivity into the grammar. In other words, Tagalog may be at a transitional stage of grammaticalisation of transitivity, and hence, it cannot be classified straightforwardly in terms of accusative/ergative parameter. In the next section, I discuss this possibility and propose a view in which the voice system in Tagalog is better treated as a split system.

3 Split voice system

In the preceding sections, we have seen that there is no morphosyntactic evidence that might indicate that AV verbs and clauses are antipassive. It follows, then, that the AV construction in Tagalog is not a construction derived from the PV construction but better taken to be a system governed by a different principle. In this section, I demonstrate a view in which the Philippine voice system is a split system with different degrees of split in different languages.

With regard to the typology of Philippine languages, Shibatani (1988:113) concludes ‘in their overall characteristics, they are neither accusative nor ergative’. Taking a step forward toward a more positive characterisation, Shibatani (1999) claims that the voice system of Philippine languages represents a fluid voice system. According to him, in so far as both AV forms and PV forms involve some morphemes, ‘there is no basic diathesis in Philippine languages; all constructions are derived’ (Shibatani 1999). The voice system is fluid if there is no basic diathesis, and the voice alternation between AV and PV is not conditioned by those factors which condition split, such as semantic nature of NPs, tense/aspect/mood, etc. This implies, then, that the distinction between AV and PV is made pragmatically, rather than grammatically, directly reflecting its context of use.

This characterisation of the Philippine voice system as fluid may largely hold for Cebuano, on which Shibatani’s argument is mostly based. This is reflected in the almost equal text frequency of AV and PV clauses in Cebuano for semantically transitive clauses. For example, Shibatani (1988) reports his Cebuano corpus contained 52% AV clauses versus 46% PV clauses, while Payne (1994), whose survey is restricted to perfective sentences, reports 41% AV versus 59% PV.

However, this characterisation of the Philippine voice system as fluid does not suffice to explain the remarkable patient-orientedness observed in Tagalog and other Philippine languages in general. This patient-orientedness is partially reflected in the text frequency, e.g. 24% AV versus 76% PV in Tagalog, according to Cooreman, Fox and Givón (1984). Moreover, as we have seen in the preceding sections, there is some morphological

evidence to show that certain PV verbs are unmarked for voice in certain moods in Tagalog. In fact, the distinction between AV and PV in Tagalog seems to be more grammatically motivated compared with Cebuano, exhibiting more of a split character than a fluid.

Dixon (1994) surveys the kinds of factors that condition splits, and summarises them as follows:

- (22) Factors that condition splits:
- a. Semantic nature of the core NPs
 - b. Semantic nature of the main verbs
 - c. Tense or aspect or mood of the clause
 - d. Grammatical status of a clause, i.e. whether it is main or subordinate, etc.

It is shown, for example, that if a language exhibits a split conditioned by tense/aspect/mood, an ergative system is likely to be employed in past tense or perfective aspect, while an accusative system is found in future tense, imperfective aspect, negative polarity, or imperative or hortative moods. Moreover, Dixon (1994) argues that some languages that show a split system do not just operate with one conditioning factor, but involve a combination of two or more conditioning factors. If the voice system of Philippine languages is a split conditioned by one or more than one of the factors in (22), it is expected that PV sentences are manifestations of an ergative system, while AV sentences are those of an accusative system.

The factors that condition splits are apparently related to transitivity of a clause. Hopper and Thompson (1980) propose the parameters of transitivity as given in the following table:

Table 1: Parameters of transitivity
(Hopper & Thompson 1980)

	Parameter	High	Low
A.	Participants	2 or more central	1 participant
B.	Kinesis	kinetic	static
C.	Aspect	telic	atelic
D.	Punctuality	punctual	non-punctual
E.	Volitionality	volitional	non-volitional
F.	Affirmation	affirmative	negative
G.	Mode	realis	irrealis
H.	Agency	A high in potency	A low in potency
I.	Affectedness of P	P totally affected	P not affected
J.	Individuation of P	P highly individuated	P not individuated

In the following sections, I discuss the situation of Tagalog in terms of each of these factors, in comparison with some other Philippine languages, and argue that Tagalog is at an early stage of split conditioned by the semantic nature of NPs (22a) and by tense/aspect/mood (22c).

3.1 Split conditioned by the semantic nature of NPs

In the case of Philippine languages, splits conditioned by the semantic nature of NPs are observed in individuation of P (Parameter J) and affectedness of P (Parameter I). The individuation of P is concerned specifically with definiteness/referentiality of P.

It is well known that whenever a patient is definite, there is a strong preference to choose the patient as the ANG-NP, resulting in PV sentences in Tagalog and other Philippine languages in general (cf. Schachter & Otanes 1972; Foley & Van Valin 1984:239). However, it is also widely acknowledged that there are differences among different languages as to how strongly the above rule applies. For example, this rule applies more strongly to Tagalog than to Cebuano:

- (23) a. *Nagluto ang babae ng/*sa manok.* [= (19)]
 AV.Perf+cook ANG woman a/*the chicken
 ‘The woman cooked a/*the chicken.’
- b. *Niluto ng babae ang manok.*
 PV.Perf+cook GEN woman ANG chicken
 ‘The woman cooked the chicken.’
- (24) Cebuano⁷
- a. *Miluto ang babaye ug/sa manok.*
 AV.Perf+cook ANG woman a/the chicken
 ‘The cooked a/the chicken.’
- b. *Giluto sa babaye ang manok.*
 PV.Perf+cook GEN woman ANG chicken
 ‘The woman cooked the chicken.’

In Cebuano, the patient of AV sentences can take either *ug* or *sa*, depending on its definiteness, i.e. *ug* for an indefinite patient and *sa* for a definite one, as in (24a). In contrast, the distinction between *ng* and *sa* in Tagalog is more of a grammatical one determined by subcategorisation of verbs rather than a semantic one such as definiteness. Thus, the definite patient must be an ANG phrase in PV sentences like (23b).

However, the so-called ‘definite object constraint’ (Schachter & Otanes 1972) is not an absolute constraint even in Tagalog. A definite patient can occur in AV sentences, as in the following sentence:

- (25) (Bloomfield 1917:48 quoted by Himmelmann 1991:39)
- Nagdadala sila ng sarile nilang banda ng musika.*
 AV.Imp+bring 3PL GEN own 3PL+LK band GEN music
 ‘They bring their own band.’

In general, then, PV clauses are high in transitivity since the definite or highly individuated patient is likely to be expressed in PV clauses, while an indefinite patient is likely to be expressed in AV clauses. However, as seen in sentences like (25), it is not an absolute condition but a tendency, which is stronger than in Cebuano. Hence, we may say

⁷ I wish to thank Cora Lalobis from Cebu City, Cebu for providing me with the Cebuano data that appear in this paper.

that Tagalog is more ergative than Cebuano in terms of definiteness of P, but it is not at all an absolute condition to determine the split.

With regard to the degrees of ergativity, Kapampangan, a language spoken in central Luzon, seems to be more ergative than Tagalog in that it has developed cross-referencing pronouns that operate clearly in an ergative manner. Consider the following examples:

- (26) Kapampangan
- a. *Minta ya i Rosa king Maynila.*
 AV+go 3SG ANG Rosa OBL Manila
 ‘Rosa went to Manila.’
- b. *Linutu ne[<na+ya]-ng Rosa ing sagin.*
 PV+cook [3SG+3SG]-LK Rosa ANG.3SG banana
 ‘Rosa cooked the banana.’
- b’. *Linutu no[<na+la]-ng Rosa reng sagin.*
 PV+cook [3SG+3PL]-LK Rosa ANG.3PL banana
 ‘Rosa cooked the bananas.’
- c. *Miglutu ya-ng sagin i Rosa.*
 AV+cook 3SG-LK banana ANG Rosa
 ‘Rosa cooked bananas.’

In intransitive sentences like (26a), the third person singular enclitic pronoun *ya* cross-references the sole argument S, *Rosa*. In PV sentences like (26b) and (26b’), two enclitic pronouns occur in portmanteau forms, and it is clear from (26b’) that the third person singular enclitic pronoun *na* cross-references the A, and the third person plural pronoun *la* cross-references the P. Hence, *ya* in (26b) cross-references the P, *sagin* ‘banana’, rather than the A, *Rosa*. Thus, the enclitic *ya* is taken to be an absolutive pronoun for the third person singular nouns that cross-references S and P in exclusion of A.

Even in Kapampangan, however, there are some cases in which an absolutive pronoun does not occur. For example, Mithun (1994) points out that mass nouns and abstract nouns are not cross-referenced by any pronoun.

- (27) Kapampangan (Mithun 1994:253)
- Tatangapan ku (∅) ing amun mu.*
 accepting 1ERG (ABS) ABS challenge your
 ‘I accept your challenge.’

Since mass or abstract nouns are low in individuation, it is also taken to be a case of split in terms of the individuation of P.

The second point regarding the semantic nature of NPs is concerned with the affectedness of P. When a main verb is high in transitivity in terms of affectedness of P, the PV construction is preferred or sometimes obligatory in Tagalog:

- (28) a. *Pinatay ni Juan ang aso.*
 PV.Perf+kill GEN Juan ANG dog
 ‘Juan killed the/a dog.’

- b. * *Pumatay si Juan ng aso.*
 AV.Perf+kill ANG Juan GEN dog
 (For: 'Juan killed a dog.')
- c. *Sino ang pumatay ng aso?*
 who ANG AV.Perf+kill GEN dog
 'Who killed a dog?'
- (29) a. *Sinira ng bata ang mesa.*
 PV.Perf+break GEN child ANG table
 'The child broke the/a table.'
- b. *? *Sumira ang bata ng mesa.*
 AV.Perf+break ANG child GEN table
 (For: 'The child broke a table.')
- c. *Sino ang sumira ng mesa?*
 who ANG AV.Perf+break GEN table
 'Who broke a table?'
- (30) a. *Kinagat ng aso ang anak ni Juan.*
 PV.Perf+bite GEN dog ANG child GEN Juan
 'The dog bit Juan's child.'
- b. ?? *Kumagat ang aso sa anak ni Juan.*
 AV.Perf+bite ANG dog OBL child GEN Juan
 (For: 'The dog bit Juan's child.')
- (Cena 1995:7 (Gloss and translation are mine))
- c. *Ano ang kumagat sa anak ni Juan?*
 what ANG AV.Perf+bite OBL child GEN Juan
 'What bit Juan's child?'

With verbs like *patay* 'kill', *sira* 'break', *kagat* 'bite', etc., whose patient is highly affected, the AV construction is usually avoided in Tagalog unless such syntactic pressures like *wh*-question, clefting or relativisation force AV forms to occur ((28c), (29c), and (30c)). Low acceptability of AV sentences ((28b), (29b), and (30b)) suggests that PV sentences are higher in transitivity in terms of affectedness of P.

On the other hand, Cebuano allows both AV and PV in such cases. For example, compare Tagalog sentences (29a, b) with their Cebuano equivalents (31a, b):

- (31) Cebuano
- a. *Gipatay ni Juan ang ero.*
 PV.Perf+kill GEN Juan ANG dog
 'Juan killed the dog.'
- b. *Mipatay si Juan ug/sa ero.*
 AV.Perf+kill ANG Juan a/the dog
 'Juan killed a/the dog.'

Again, we can say that Tagalog is more ergative than Cebuano in that PV forms are strongly preferred with verbs of high transitivity such as *patay* 'kill' in Tagalog, while AV forms as well as PV are allowed in Cebuano.

If a verb does allow either AV or PV, the distinction signals different degrees of affectedness of P. Consider the following sentences (glosses and translations are mine):

- (32) (Cena 1977:6–7)
- a. *Binaril ni Juan si Fred.*
 PV.Perf+shoot GEN Juan ANG Fred
 ‘Juan shot Fred.’
- b. *Bumaril si Juan kay Fred.*
 AV.Perf+shoot ANG Juan OBL Fred
 ‘Juan shot at Fred.’

The PV construction in (32a) encodes an event in which the patient is wholly affected, while the AV construction in (32b) encodes one in which the patient is either not directly affected or just partially affected.

Overall, then, definiteness and affectedness of P are factors that may condition splits between the AV construction and the PV construction in Tagalog. The important point, however, is that these factors are not determinative, i.e. even when the P is definite, and wholly affected, we can find cases in which the AV construction is used. Hence, we can say that Tagalog is still at an early stage of split, where grammaticalisation of transitivity, i.e. its regularisation in the grammar, is not at all complete compared with other languages like Kapampangan.

3.2 Split conditioned by tense/aspect/mood

We have seen in §2 that IN- verbs are morphologically unmarked for voice in realis mood. This morphological unmarkedness is also partially reflected in syntax. As mentioned above, with verbs of high transitivity like *patay* ‘kill’, the AV construction is usually avoided in perfective aspect. However, in imperfective or contemplated aspect, acceptability of the AV construction increases:

- (33) a. * *Pumatay si Juan ng aso.* [= (28b)]
 AV.Perf+kill ANG Juan GEN dog
 (For: ‘Juan killed a dog.’)
- b. *Nagpapatay si Juan ng aso.*
 AV.Imp+kill ANG Juan GEN dog
 ‘Juan kills a dog.’ (‘John butchers dogs.’)
- c. *Papatay si Juan ng aso.*
 AV.Cont+kill ANG Juan GEN dog
 ‘Juan will kill a dog.’

Dixon (1994) argues that if there is a split conditioned by tense/aspect/mood, an ergative system is likely to be employed in past tense or perfective aspect, while an accusative system is found in future tense, imperfective aspect, negative polarity, or imperative or hortative moods. The above examples illustrate the point although the split in Tagalog is not clear-cut but remains a tendency.

Also recall from §2 that UM- verbs are unmarked for voice in irrealis mood. Moreover, unaffixed ‘bare’ verbs can function as AV verbs in irrealis, especially in imperative or hortative mood. In the following examples, bare forms of verbs function as AV verbs:

- (34) (Bloomfield 1917:221 quoted by Himmelmann 1999)
Hampas na kayo, mga bata, sa mga langgam. [= (13)]
 blow now 2PL PL child LOC PL ant
 ‘Whip at the ants, boys.’
- (35) (Himmelmann 1999)
Umuwi na tayo, Daddy! Uwi na tayo!
 AV+return.home now 1PL Daddy return.home now 1PL
 ‘Let’s go home Daddy! Let’s go home!’

In other words, the PV forms are more likely to be unmarked in realis mood, while the AV forms are more likely to be unmarked in irrealis mood. This is expected if we assume that the PV construction represents an ergative system and the AV construction an accusative system.

The split characteristic is further observed in the case of so-called pseudo-verbs. Some adjectives of liking/disliking, necessity, and ability can function as verbs, such as *gusto* ‘like, want’, *ayaw* ‘don’t like’, *ibig* ‘like, want’, *nais* ‘like, want’, *kailangan* ‘need’, etc. (cf. Schachter & Otnes 1972:261ff.). But in their bare forms, they do not usually function as AV verbs. They occur either in the so-called ‘non-topic’ construction, in which no ANG-marked NP occurs, or the PV construction, in which a patient (theme) is ANG-marked. For example:

- (36) a. *Gusto/Kailangan ng bata ng libro.* [Non-topic]
 like/need GEN child GEN book
 ‘The child likes/needs a book.’
- b. *Gusto/Kailangan ng bata ang libro.* [PV]
 like/need GEN child ANG book
 ‘The child likes/needs the book.’
- c. * *Gusto/Kailangan ang bata ng libro.* [AV]
 like/need ANG child GEN book
 (For: ‘The child likes/needs a book.’)

Thus, these pseudo-verbs are highly patient-oriented in that if one of the NPs is marked by ANG, it is the patient in Tagalog.

On the other hand, *ayaw* ‘don’t like’, the negative counterpart of *gusto* ‘like’, allows AV as well as non-topic and PV constructions, as in (37c) (Schachter & Otnes 1972:265):

- (37) a. *Ayaw ni Juan ng mansanas.* [Non-topic]
 don’t.like GEN Juan GEN apple
 ‘Juan doesn’t like an apple.’
- b. *Ayaw ni Juan ang mansanas.* [PV]
 don’t.like GEN Juan ANG apple
 ‘Juan doesn’t like the apple.’

- c. *Ayaw ng mansanas si Juan.* [AV]
 don't.like GEN apple ANG Juan
 'Juan doesn't like an apple.'

Interestingly, the corresponding Cebuano form does not allow PV. Compare the Tagalog sentences in (37) with their corresponding Cebuano sentences in (38):

- (38) Cebuano
- a. *Dili nako ug nangka.* [Non-topic]
 don't.like 1SG.GEN GEN jackfruit
 'I don't like jackfruit.'
- b. * *Dili nako ang nangka.* [PV]
 don't.like 1SG.GEN ANG jackfruit
 (For: 'I don't like the jackfruit.')
- c. *Dili ako ug nangka.* [AV]
 don't.like 1SG.ANG GEN jackfruit
 'I don't like jackfruit.'

Thus, even with pseudo-verbs that usually exhibit strong patient-orientedness, negation opts for the AV construction. This is expected if we assume that the PV construction represents an ergative construction while the AV construction an accusative construction. The situation accords with a general tendency: an ergative system is more likely to be found in clauses that describe some definite result while an accusative system is more likely to be employed when the clause refers to something that has not happened.

In sum, the PV construction is more likely to be employed in realis mood or perfective aspect, while the AV construction is more likely to be found in irrealis mood, imperfective aspect, imperative and hortative moods, and in negative polarity. This is in accordance with the general tendency of split observed in other languages in which an ergative system is more likely to be found in the former contexts while an accusative system is more likely to be employed in the latter. An important difference of Tagalog from those languages with clear splits is that the factors that condition the splits are not at all definite factors in the choice of particular voice.

4 Conclusion

This paper started off with a discussion of the nature of the actor-voice construction in Tagalog and other Philippine languages. Although the ergative analysis assumes that the actor-voice construction is an antipassive, I have shown that such an interpretation is not morphologically and syntactically plausible. I have presented another interpretation of the AV construction as accusative, and suggested that Tagalog seems to be at an early stage of split voice system in that factors that condition splits in other languages are not determining factors.

In the discussion, some significant differences among different Philippine languages have also been mentioned. It has been shown that Kapampangan is more ergative than Tagalog in that the former has developed a cross-referencing pronominal system that operates clearly in an ergative manner, and that Cebuano is less ergative than Tagalog in that it often allows the AV construction where Tagalog does not. Similar differences in the

voice system are also observed in other Austronesian languages. For example, Polynesian languages, a subgroup of the Austronesian family, also exhibit differences in ergativity. Dixon (1994) points out that while such Austronesian languages as Tongan and Samoan are ergative, some other languages like Maori are accusative with a passive construction being extensively used. It is also reported that Formosan languages in general have Philippine-type voice systems with differences in ergativity/accusativity. I believe that careful studies of a wider variety of Austronesian languages will enable a unified account of the Philippine-type voice system, which in turn will give an explanation of the motivation behind the rise of ergative characteristics observed widely in Austronesian languages.

References

- Blake, Barry J., 1990, *Relational grammar*. London: Routledge.
- Bloomfield, Leonard, 1917, *Tagalog texts with grammatical analysis*. 3 vols. Urbana IL: University of Illinois.
- Cena, Resty M., 1977, Patient primacy in Tagalog. Paper read at the LSA Winter Meeting, Chicago, Illinois.
- 1995, Surviving without relations. *Philippine Journal of Linguistics* 26:1–31.
- Cooreman, Ann, Barbara Fox and Talmy Givón, 1984, The discourse definition of ergativity. *Studies in language* 8:1–34.
- De Guzman, Videia P., 1979, Morphological evidence for primacy of patient as subject in Tagalog. Paper read at the LSA Annual Meeting, Los Angeles, California.
- 1988, Ergative analysis for Philippine languages: an analysis. In McGinn, ed. 1988:323–345.
- 1992, Morphological evidence for primacy of patient as subject in Tagalog. In M.D. Ross, ed. *Papers in Austronesian linguistics*, No. 2, 87–96. Canberra: Pacific Linguistics.
- 1998, The ergative analysis: a different view of structure. Lecture in the series *Ergativity in Philippine languages*. University of the Philippines, Diliman, 10th December.
- 1999, On the *ng*-agent and the *ng*-patient in Tagalog. Lecture in the series *Ergativity in Philippine languages*. University of the Philippines, Diliman, 18th March.
- De Wolf, Charles M., 1988, Voice in Austronesian languages of Philippine type: passive, ergative, or neither? In Shibatani, ed. 1988:143–93.
- Dixon, R.M.W., 1994, *Ergativity*. Cambridge: Cambridge University Press.
- Foley, William A. and Robert D. Van Valin, 1984, *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.
- Gerds, Donna B., 1988, Antipassives and causatives in Ilokano: evidence for an ergative analysis. In McGinn, ed. 1988:295–321.
- Heath, Jeff, 1976, Antipassivization: a functional typology. *Proceedings of the Second Annual Meeting of the Berkeley Linguistics Society*, 202–211.
- Himmelmann, Nikolaus P., 1991, *The Philippine challenge to Universal Grammar*. Arbeitspapier Nr. 15. Köln: Institut für Sprachwissenschaft.

- 1999, Lexical categories and voice in Tagalog. Manuscript.
- Hopper, Paul and Sandra Thompson, 1980, Transitivity in grammar and discourse. *Language* 56:251–299.
- Kozinsky, Isaac Š., Vladimir P. Nedjalkov and Maria S. Polinskaya, 1988, Antipassive in Chukchee: oblique object, object incorporation, zero object. In Shibatani, ed 1988:651–706.
- Kroeger, Paul R., 1993a, *Phrase structure and grammatical relations in Tagalog*. Stanford: CSLI Publications.
- 1993b, Another look at subjecthood in Tagalog. *Philippine Journal of Linguistics* 24:1–15.
- Kuribayashi, Yuu, 1989, Accusative marking and noun–verb constructions in Turkish. *Gengo Kenkyuu* (Journal of the Linguistic Society of Japan) 95:94–119.
- McGinn, Richard, ed., 1988, *Studies in Austronesian Linguistics*. Athens, Ohio: Center for International Studies.
- Mithun, Marianne, 1994, The implications of ergativity for a Philippine voice system. In Barbara Fox and Paul J. Hopper, eds *Voice: form and function*, 247–277. Amsterdam/Philadelphia: John Benjamins.
- Oue, Masanao, 1994, *Firipino-go Bumpoo Nyuumon* (Introduction to the Filipino Grammar). Tokyo: Hakusuisha.
- Palmer, Frank R., 1994, *Grammatical roles and relations*. Cambridge: Cambridge University Press.
- Payne, Thomas E., 1982, Role and reference related subject properties and ergativity in Yup'ik Eskimo and Tagalog. *Studies in Language* 6:75–106.
- 1994, The pragmatics of voice in a Philippine language: actor-focus and goal-focus in Cebuano narrative. In Talmy Givón, ed. *Voice and inversion*, 317–364. Amsterdam/Philadelphia: John Benjamins.
- Schachter, Paul and Fe Otones, 1972, *Tagalog reference grammar*. Berkeley: University of California Press.
- Shibatani, Masayoshi, 1988, Voice in Philippine languages. In Shibatani, ed. 1988:85–142.
- 1991, Grammaticization of topic into subject. In Elizabeth C. Traugott and Bernd Heine, eds *Approaches to grammaticalization*, vol. 1, 93–133. Amsterdam/Philadelphia: John Benjamins.
- 1999, Case, voice and language type. Lecture at University of the Philippines, Diliman, 8th June.
- Shibatani, Masayoshi, ed., 1988, *Passive and voice*. Amsterdam/Philadelphia: John Benjamins.

8

Aspects of pragmatic focus in Tagalog

DANIEL KAUFMAN

1 Background

This paper sets out to accomplish the following three goals:

- i. to show that Tagalog possesses regular syntactic expressions of the universal pragmatic relations of focus and topic;
- ii. to identify the role of prosody in the pragmatic component of Tagalog;
- iii. to account for a portion of what was previously considered to be semantically vacuous scrambling in Tagalog based on the first two observations.

Because of its typologically interesting nature, the morphosyntax of Philippine case/voice alternations has overshadowed work on the pragmatics of these languages.¹ Consequently, work on Philippine languages has often conflated syntactic and pragmatic categories, a problem which is reflected by the use of pragmatic terms such as ‘topic’ and ‘focus’ to refer to what should properly be considered grammatical relations. Assumptions about the pragmatics and discourse role of the Tagalog *ang* phrase (which has been variously termed ‘focus’, ‘topic’, and ‘subject’ among others) have too often gone unexamined.

As a repercussion of ignoring discourse context we are left with the appearance that scrambling is applied freely as an ‘optional rule’. But upon taking a closer look at the interactions between the pragmatic and syntactic components of the grammar we find that much of this variation is in fact conditioned by information structure. Within the Chomskyan model, information structure has been subsumed into the syntax to a large extent by positing functional projections for pragmatically salient positions. The presence of a Topic Phrase and Focus Phrase in the left periphery of the sentence has been proposed in numerous places. These serve as potential landing sites for movement from case

¹ Tagalog pragmatics is the primary focus of Naylor (1975) and is given a brief treatment by Schachter and Otones (1972) and Kroeger (1993:Ch. 3).

positions. Pragmatically salient functional projections are now widely accepted in generative work on languages that are described as being ‘discourse configurational’ (e.g. Hungarian, Catalan, Somali) as well as those which are not generally characterised as such (e.g. Italian, English).

For the purposes of this paper, I abstract away from the full implications of an X-bar theoretic analysis in order to give attention to the basic pragmatic and syntactic properties of topic and focus in Tagalog. I also show that a prosodically based analysis of the postverbal field as that of Zubizarretta (1998) may be profitably applied to Tagalog to account for some of what was previously dismissed as free word order variation.

1.1 Definition of terms

I rely here primarily on Lambrecht’s (1994) framework for dealing with topic and focus. Lambrecht defines topicality as the following:

A referent is interpreted as the topic of a proposition if in a given discourse the proposition is construed as being about the referent, i.e. as expressing information which is relevant to and which increases the addressee’s knowledge of this referent. Following Reinhart (1982), we may say that the relation ‘topic-of’ expresses the pragmatic relation of aboutness which holds between a referent and a proposition with respect to a particular discourse. The term ‘pragmatic relation’ should be understood as meaning ‘relation construed within particular discourse contexts’. Topic is a pragmatically construed sentence relation. (Lambrecht 1994:127)

The focus of a sentence on the other hand is conceived of as a relation between the pragmatic presupposition and assertion of a sentence:

PRAGMATIC PRESUPPOSITION: The set of propositions lexico-grammatically evoked in a sentence which the speaker assumes the hearer already knows or is ready to take for granted at the time the sentence is uttered.

PRAGMATIC ASSERTION: The proposition expressed by a sentence which the hearer is expected to know or take for granted as a result of hearing the sentence uttered.

FOCUS: The semantic component of a pragmatically structured proposition whereby the assertion differs from the presupposition. (Lambrecht 1994: 213)

Lambrecht (1994) exemplifies these notions with the sentences in (1) and (2). Without entering into the particulars of this construction and the mapping between intonation and information structure, we see how these pragmatic categories follow basic intuitions closely.

- (1) *My car broke DOWN.*
 Presupposition: ‘speaker’s car is a topic for comment x’
 Assertion: ‘x = broke down’
 Focus: ‘broke down’
- (2) *My CAR broke down.*
 a. Presupposition: ‘speaker’s x broke down’
 Assertion: ‘x = car’
 Focus: ‘car’

- b. Presupposition: \emptyset
 Assertion: ‘speaker’s car broke down’
 Focus: ‘speaker’s car broke down’

The intonation pattern in (1) correlates strictly with the given information structure. (1) is appropriate only when the speaker’s car is a topic or possible contrastive topic. On the other hand, because of the nature of this type of predicate in English, the intonation indicated in (2) may correlate to one of two different information structures. In (2a) there is narrow focus on the argument, making this sentence appropriate in answering the question, ‘What broke down?’. This same intonation pattern may also be used to signal wide focus as represented in (2b) where no presupposition is evoked. In this case the focus and assertion are identical.

1.2 Methodology

All natural languages allow for multiple ways of saying the same thing with essentially the same words. Daneš (1966) of the Prague School, coined the term ALLOSENTENCES (on analogy with the allophone) to refer to the various surface manifestations of a single logical proposition. Elements such as word order, sentential stress and passivisation may all contribute to creating the different allosentences of a single proposition.²

Here we undertake a systematic comparison of allosentences in Tagalog, concentrating particularly on position and intonation. By examining the interaction between discourse conditions, intonation and word order, we can better understand the role of topicalisation, focalisation and scrambling in Tagalog. In comparing allosentences, I employ the question–answer pair method that has become popular in the literature since Jackendoff (1972) where it was introduced as a formal diagnostic for focus. Here the focus of a given sentence is revealed by examining what questions that sentence may answer felicitously. The informative part of the answer — the part of the answer which corresponds to the information being sought by the question — is taken to be the focus. Observe the simple dialogue in (3).

- (3) A: What did John take to the picnic?
 B: John took potato salad to the picnic.

The underlined portion of B’s response (which would be given prosodic prominence) is the focus based on the fact that it provides the answer elicited by a preceding question.³ The unacceptability of (4a) and (4b) within the same context shows that these sentences cannot contain the same information structure as the answer in (3).

² This was reflected in earlier Generative work where scrambling, passivisation, etc. were seen as rules that applied to an underlying structure which corresponded more directly to semantic content.

³ I abstract away from the elliptic nature of natural discourse in order to concentrate on the differences which exist between formally different but semantically similar full sentences. Thus, although it might be argued that a more natural answer to A’s question above might simply be ‘potato salad’, we must remember that this is in part due to the absence of potential ambiguity in such a simple context. Within actual discourse, full sentences are often employed to relate the focus to particular referents. The question–answer test represents only a truncated version of discourse and is thus an ‘extreme case’ of how discourse relations such as topic and focus function within a larger context.

- (4) A: What did John take to the picnic?
 a. B: # The picnic is where John took the potato salad.
 b. # John took the potato salad to the picnic.

Employing this and other diagnostics we find that Tagalog does possess well defined, pragmatically significant positions within the phrase structure. Furthermore, we find that Tagalog employs syntactic methods commonly where English employs phonological means.

2 Allosentences in Tagalog

Tagalog, like many Austronesian languages, possesses a predicate-initial word order. The predicate may be of any category and since Tagalog lacks a copula there is little syntactic distinction between verbal and non-verbal predicates. The variants we are concerned with here involve occupying the predicate position with different elements from the sentence, inversion with the particle *ay*, and fronting of the oblique phrase. We can see the basic paradigm of Tagalog allosentences under investigation in (5-9).⁴

- (5) *Lumangoy ang bátá sa Bulakan.*
 AV.COM:swim NOM child OBL Bulacan
 ‘The child swam in Bulacan.’
- (6) *Bátá ang lumangoy sa Bulakan.*
 child NOM AV.COM:swim OBL Bulacan
 ‘The one who swam in Bulacan is a child.’
- (7) *Ang bátá ay lumangoy sa Bulakan.*
 NOM child ay AV.COM:swim OBL Bulacan
 ‘The child, swam in Bulacan.’
- (8) *Sa Bulakan ay lumangoy ang bátà.*
 OBL Bulacan ay AV.COM:swim NOM child
 ‘In Bulacan, the child swam.’
- (9) *Sa Bulakan lumangoy ang bátà.*
 OBL Bulacan AV.COM:swim NOM child
 ‘It was in Bulacan that the child swam.’

Sentences (5)–(9) may all truthfully describe the same situation but all differ in regard to topic and focus (information structure) and the novelty of the participants (definiteness). Sentences such as those in (6) have been considered to be clefts by several authors (Kroeger 1993; Aldridge, 2002) but for our purposes it is sufficient to treat them simply as nominal predicate sentences without concern for their derivation. Sentences (7) and (8)

⁴ Abbreviations used: ADJ adjectival; AV actor voice; CAU causative; COM completed aspect; CV conveyance voice; EXT existential; GEN genitive (ergative) case marker; GER gerund; IMP imperative; INC incomplete aspect; INF infinitive; IRR irrealis; LNK linker; LV locative voice; NEG negative; NOM nominative (absolutive) case marker; NONV non-volitional; OBL oblique/locative case marker; P personal (+human); PV patient voice; Q question marker; REP reported speech; STA stative. Pronouns are nominative unless glossed otherwise.

exemplify *ay* inversion of the subject, and an oblique, respectively. *Ay* inversion involves preposing an *ang* phrase or oblique phrase to the preverbal position followed by the particle *ay*. Sentence (9) exemplifies a different kind of fronting operation which may only apply to obliques and adjuncts. It is characterised by the fact that it forms a unitary prosodic phrase with the rest of the sentence and that clitics follow the fronted element directly, in contrast to *ay* inversion (cf. Kroeger 1998a). The difference in clitic placement is seen in (10) (clitics are underlined).

- (10) a. *Sa Bulakan silá lumangoy.*
 OBL Bulacan 3.PL AV.COM:swim
 ‘It was in Bulacan that they swam.’
- b. *Sa Bulakan ay lumangoy silá.*
 OBL Bulacan ay AV.COM:swim 3.PL
 ‘In Bulacan, they swam.’

It will be shown here through the use of the question–answer test and through several syntactic diagnostics that XP in the construction [XP [*ang* YP]] occupies a focus position while DPS and PPs preceding *ay* in sentences such as (7) and (8) are topics. Similar conclusions have been reached by Schachter and Otones (1972), Naylor (1975) and Kroeger (1993). What will be further developed here is the extent to which these pragmatic relations appear to follow universal syntactic configurations for foci and topics.

3 Identifying the focus and topic positions in Tagalog

The information structure of the allosentences discussed in the previous section was already hinted at by the translations. These intuitions can be systematically verified through the question–answer test. In (11) we see three allosentences expressing the same proposition in answer to a question which elicits VP focus.

- (11) A: *Ano ang ginawá ninyo sa Bulakan?*
 what NOM OV.COM:do 2.PL.GEN OBL Bulacan
 ‘What did you do in Bulacan?’
- a. B: *Nag-píknik kami sa Bulakan.*
 AV.COM-picnic 1.PLEX OBL Bulacan
 ‘We picnicked in Bulacan.’
- b. *Sa Bulakan ay nag-píknik kami.*
 OBL Bulacan ay AV.COM-picnic 1.PLEX
 ‘In Bulacan, we picnicked.’
- c. # *Sa Bulakan kami nag-píknik.*
 OBL Bulacan 1.PLEX AV.COM-picnic
 ‘It was in Bulacan that we picnicked.’

Several important points may be gleaned from native speaker judgements concerning (11a–c). Because Bulacan is mentioned in the question it is part of the background and cannot felicitously be presented as the focus. The unacceptability of (11c) therefore suggests that the position occupied by the oblique phrase is indeed a focus position. The

acceptability of (11b), on the other hand, suggests that the oblique occupies a topic position and that topicalisation of a referent in the background is optional, not effecting felicity judgements. These generalisations are supported by the judgments found in (12) where the question elicits focus on the oblique phrase.

- (12) A: *Saan kayo nag-piknik?*
 where 2.PL AV.COM-picnic
 ‘Where did you picnic?’
- a. B: #? *Nag-piknik kami sa Bulacan.*
 AV.COM-picnic 1.PL.EX OBL Bulacan
 ‘We picnicked in Bulacan.’
- b. # *Sa Bulacan ay nag-piknik kami.*
 OBL Bulacan ay AV.COM-picnic 1.PL.EX
 ‘In Bulacan, we picnicked.’
- c. *Sa Bulacan kami nag-piknik.*
 OBL Bulacan 1.PL.EX AV.COM-picnic
 ‘It was in Bulacan that we picnicked.’

(12b) shows that, as it was infelicitous to present an element from the background as a focus, e.g. (11c), it is also infelicitous to present elicited information as a topic. The elicited focus, *sa Bulacan*, must appear either in the focus position as in (12c) or more marginally in postverbal position with the verb as predicate. The fact that (12a) is not fully accepted by all speakers as an appropriate response is noteworthy. This suggests that the verb must be interpreted as part of the focus in this position. That is, (12a) can serve to answer the questions, ‘What did you do in Bulacan?’, ‘What did you do?’ or ‘What happened?’ but cannot answer a question where the verb is part of the background. If this turns out to be the dominant interpretation then we may conclude that the first position of the clause proper is a focus position regardless of the lexical category which fills it.

3.1 Focus sensitive operators

Typically, the class of words considered to be focus sensitive includes the adverbs ‘only’, ‘also’, ‘even’, and negation (Rooth, 1996; König, 1991). This class of words is predicted to only take scope over the focus of a sentence. We can therefore predict that focus sensitive elements will be ungrammatical if syntactically forced to associate with a topic as presupposed information should not be available for modification. These predictions are borne out by the scope and grammaticality of negation as shown in (13).

- (13) a. *Hindi sa Bulacan kami nag-piknik.*
 NEG OBL Bulacan 1.PL.EX AV.COM-picnic
 ‘It wasn’t in Bulacan where we had a picnic.’
- b. *Hindi kami nag-piknik sa Bulacan.*
 NEG 1.PL.EX AV.COM-picnic OBL Bulacan
 ‘We didn’t picnic in Bulacan.’

- c. **Hindí sa Bulakan ay nag-píknic kami.*
 NEG OBL Bulacan ay AV.COM-picnic 1.PL.EX
 (cf. * 'Not in Bulacan, we picnicked.')

In (13a) we find the oblique focus construction with the clause-initial negator *hindí*.⁵ The only reading available in (13a) is the one in which negation takes narrow scope over the oblique phrase and the proposition 'we picnicked' is presupposed. This contrasts with (13b), where negation may take scope over the entire sentence.⁶

If we try to negate the phrase fronted with the particle *ay* the result is ungrammatical as seen in (13c). The ungrammaticality of (13c) further supports the analysis of this type of fronting as a topicalisation. A pure phrase-structural approach to these facts could simply state that this results from a universal clausal architecture where the order TopP > NegP > FocP is generated in the left periphery. This, however, is no more than a stipulated template. The semantic basis for these facts is that negation takes scope from its surface position. Because negation cannot scope over a topic it must therefore follow it as in (14).

- (14) *Sa Bulakan ay hindí kami nag-píknic.*
 OBL Bulacan ay NEG 1.PL.EX AV.COM-picnic
 'In Bulacan, we didn't picnic.'

The facts are similar for quantificational adverbs which are also considered to be members of the universal set of focus-sensitive items. The Tagalog quantificational adverb *lang* 'only', can be seen to follow the same pattern of scope and grammaticality as that of negation.

- (15) a. *Sa simbahan lang ako nag-bi-bigay ng pera.*
 OBL church only 1.SG AV-INC-give GEN money
 'It's only to the church that I give money.'
- b. *Sa simbahan ay nag-bi-bigay lang ako ng pera.*
 OBL church ay AV-INC-give only 1.SG GEN money
 'To the church, I only give money.' (OR: 'In church, all I do is give money.')
- c. **Sa simbahan lan ay nag-bi-bigay ako ng pera.*
 OBL church only ay AV-INC-give 1.SG GEN money

(15a) shows that *lang* may associate only with the oblique phrase, and cannot take wider scope with the meaning, 'the only thing I do is give money to the church'. Again, this supports the idea that the oblique in clause-initial position is focused. In (15b) we see that the focus-sensitive adverb *lang* may not associate with the PP fronted with *ay* but rather must associate with the VP or a constituent thereof. We can obtain a unified explanation for why the second position clitic *lang* cannot follow, and negation may not precede, a preposed topic. Adverbial clitics such as *lang* modify the prosodic hosts to their left or

⁵ The negator *hindí* is a host for clitic placement thus the pronominal clitic *kami* appears preverbally in (13b)

⁶ This has implications for the issue of symmetry between lexical categories since we find no presupposition, 'x happened in Bulacan' in (13b) when the verb is in the predicate position. Based on (12) and (13) we may conclude that when the verb is in predicate position it must be part of the focus (for many speakers) but does not render the rest of the sentence presuppositional as does the fronted oblique phrase. This suggests that the element filling the predicate position must be part of the focus but that only fronted obliques are in a dedicated focus phrase.

larger constituents which contain that host. We see then that the same scopal basis for the ungrammaticality of (13c) applies also to (15c). Interestingly, it is only when the verb is in predicate position that the adverbial clitic may take wide scope over the entire proposition as in (16).

- (16) *Nag-bí-bigay lang ako ng péra sa simbáhan.*
 AV-INC-give only 1.SG GEN money OBL church
 ‘I only give money to the church.’ (All that happens is I give money to church.)

This again suggests that when the verb is in this position the sentence is not split into a focus and a presupposition.

3.2 Cooccurrence of focus and *wh*- elements

Wh- questions are also seen to interact differently with topics and foci. Both from the semantics of *wh*- questions and their cross-linguistic behaviour, we expect to find a parallel between *wh*- elements and foci. In Tagalog we find an exact parallel between the position of the *wh*- words, *ano* ‘what’, *alin* ‘which’ and *sino* ‘who’ and the position of the NP in the cleft-like construction.

- (17) *Sino ang lumangoy sa ílog?*
 who NOM AV.COM:swim OBL river
 ‘Who swam in the river?’ (OR: ‘Who was the one who swam in the river?’)

Similarly we find that the position of the *wh*- words *saan* ‘where’, *kailan* ‘when’ and *paano* ‘how’ occupy the same position as the fronted oblique. This set of *wh*-words also acts as a host for clitics as was observed with fronted obliques. The two constructions are compared in (18) and (19).

- (18) a. *Saan ka pumunta?*
 where 2.SG AV.COM:go
 ‘Where did you go?’
 b. *Sa Manila ka ba pumunta?*
 OBL Manila 2.SG Q AV.COM:go
 ‘Did you go to Manila?’
- (19) a. *Kailan ka nag-túrò?*
 when 2.SG AV.COM-teach
 ‘When did you teach?’
 b. *Kahápon ka ba nag-túrò?*
 yesterday 2.SG Q AV.COM-teach
 ‘Was it yesterday you taught?’
 (cf. *Nagtúro ka kahápon?* – ‘Did you teach yesterday?’)

These fronted elements are in complementary distribution. Thus, an oblique cannot co-occur in the initial focus position with an oblique *wh*- element.

- (20) * *Saan ka kahápon pumunta?*
 where 2.SG yesterday AV.COM:go

- (21) * *Kailan ka sa Manila pumunta?*
 when 2.SG OBL Manila AV.COM:go

Again, from a purely phrase-structural standpoint, the argument could be made that the focus phrase happens to be unique within the clause. Although this position is taken by Rizzi (1997), he also suggests a semantic basis for such a template. The ungrammaticality of multiple focus phrases may be a result of the bifurcating nature of focus. As we observed above, the presence of focus renders the rest of the sentence into a presupposition. Given that these two pragmatic relations are mutually exclusive it is unclear how this operation may apply recursively and still yield an interpretable output. In the presence of two focus phrases, what is designated as focal by the first, will be rendered as presuppositional by the second.⁷

This brings us back to the issue of categorial symmetry. If there is no inherent pragmatic difference between verbal and nominal predicate sentences we expect (22) to be felicitous.

- (22) ? *Kailan ba babae ang magiging presidente?*
 when Q woman NOM IRR.LV:become president
 ‘When is it a woman that will become president?’

We find though that (22) is not fully acceptable. There is a strong tendency to follow *wh-* words of the type in (22) with a verb and not a noun. This appears to be one of the few areas in Tagalog where lexical categories are syntactically relevant.

Stronger judgments are found for the following example. In (23a) a pronoun is in the predicate position of a cleft-like construction while in (23b) the *wh-* element is in the position reserved for focused obliques. (Recall that the pronoun appears in second position in [23b] because of its clitic status.)

- (23) a. *Siya ang sikat.*
 3.SG NOM famous
 ‘S/he is the famous one.’
 b. *Saan siya sikat?*
 here 3.SG famous
 ‘Where is s/he famous?’

Crucially, we find that these two structures may not co-occur as in (24) and (25).

- (24) * *Saan siya ang sikat?*
 where 3.SG NOM famous
 (For, ‘Where is s/he the famous one?’)

⁷ This particular property is one of those which distinguish syntactic focus from prosodic focus. Cross-linguistically there appears to be no restriction on sentences containing multiple elements which are prosodically focused. This is the distinction which Rooth (1996) refers to as ‘strong’ versus ‘weak’ focus. ‘Strong focus’ is realised through syntactic means and usually contains an exhaustive listing effect and the property of turning the non-focal portion of the proposition into a presupposition. The difference can be seen in comparing a cleft (syntactic focus) as in (i) with prosodic focus as in (ii).

- (i) #Either it was John who broke the window or the window isn’t broken.
 Presupposition: *x broke the window*
 (ii) Either JOHN broke the window or the window isn’t broken.
 No presupposition

- (25) *? *Saan babae ang mas ma-tangkad?*
 where woman NOM more ADJ-height
 (For, 'Where is it that women are taller?')

A comprehensive treatment of the ungrammaticality behind (24) and (25) must await further research.⁸ For our purposes, it is sufficient to note that a parallel exists between the fronted oblique and the cleft-like construction on the one hand and their *wh*- counterparts on the other hand.

3.3 Relative positions of topic and focus

Interrogatives may not occur fronted with *ay* as in (26) suggesting that foci in general are banned from occurring before *ay*.

- (26) a. **Saan ay pumunta ka?*
 where *ay* AV.COM:go 2.SG
 b. **Ano ay ginawa mo?*
 what *ay* PV.COM:do 2.SG.GEN

Taking fronting with *ay* to be a topicalisation, the facts in (26) are accounted for by the pragmatic incompatibility of a focus in a topic position. This is true for English as it is for Tagalog. (27a) contains a topicalised PP in a declarative sentence while (27b) displays the ungrammaticality of having a *wh*- phrase (a focus by definition) in the position of a topic.

- (27) a. To George, they gave a tortoise.
 b. *To whom, they gave a tortoise?

We do find topics co-occurring with foci in sentences such as (28a) in which the topic precedes a focused oblique. The reverse order, focus>topic, as in (28b) is not allowed in Tagalog as is also the more general case.⁹

- (28) a. *Ang isdá ay sa túbig na-bú-búhay.*
 NOM fish *ay* OBL water STA-INC-live
 'Fish live in the water.'
 b. **Sa túbig ang isdá ay na-bú-búhay.*
 OBL water NOM fish *ay* STA-INC-live

⁸ See Foley (1998) and Kroeger (1998b) for other issues involved in lexical categoriality in Tagalog.

⁹ Compare, for instance, Hebrew where topics are often marked by coindexed resumptive pronouns:

- (i) [*Dagim*]_{Top} [*ba-maym*]_{Foc} *hem hayyim.*
 fish in-water 3.PL-M live-PL-M
 'Fish, they live in the water.'
 (ii) **[Ba-maym]_{Foc} [dagim]_{Top} hem hayyim*
 in-water fish 3.PL-M live-PL-M

3.4 Recursivity

Recursivity is widely accepted to be a property of topics as opposed to ‘strong’ (see fn.7) focus in several languages. Multiple topics are also allowed in Tagalog, although this possibility is tempered by the well-known Austronesian constraint on object extraction. Multiple topic constructions as in (29) and (30) necessarily contain adjuncts and obliques besides the *ang* phrase.

- (29) *Ngayon ay siya ay na-tá-tákot.*
 now ay 3.SG ay STA-INC-fear
 ‘Now he is scared.’
- (30) *Ngayon ay itong dalawang bátá ay palibhásá*
 now ay this-LNK two-LNK child ay because
ay intersado ...
ay interested
 ‘Now, these two children, on account of being interested ...’
 (Taken from Wolff forthcoming; text 20)

As shown by (31), recursion of focus is ungrammatical, as is cooccurrence of a focused phrase with a *wh*- element (cf. §3.2).

- (31) **Sa Bulacan kahapon sila pumunta.*
 OBL Bulacan yesterday 3.PL AV.PRF;go

3.5 Dislocation

Left-dislocation constructions are also widely held to be compatible with topics in contrast to foci. Left-dislocation is marked by a preposed argument co-indexed with a resumptive pronoun in base position. Indeed we find that in Tagalog, resumptive pronouns do not refer back to foci but may be co-indexed with topics. Example (32) shows a question–answer pair in which the Agent, *Dódong*, has topic status while (33) is a minimally differing question–answer pair in which the Agent has focus status. The status of B’s response containing left-dislocation is regarded as more felicitous in the context of (32) than it is in the context of (33).

- (32) A: *Ano ang ginagawa ni Dódong ngayon?*
 what NOM PV.IMP:do P.GEN Dodong now
 ‘What’s Dodong doing now?’
- B: *Si Dódong, nag-ááral siya sa UP.*
 P.NOM Dodong AV.INC-study 3.SG OBL UP
 ‘Dodong, he studies at UP.’
- (33) A: *Sino ang nag-ááral sa UP?*
 who NOM AV.INC-study OBL UP
 ‘Who studies at UP?’

B: #*Si Dódong, nag-ááral siya sa UP.*
 P.NOM Dodong AV.INC-study 3.SG OBL UP
 ‘Dodong, he studies at UP.’

In addition, a preposed focus cannot be coindexed with a resumptive pronoun as in (34).

(34) **Si Dódong ang nag-ááral siya sa UP.*
 P.NOM Dodong NOM AV.INC-study 3.SG OBL UP

The findings above are summarised in Table 1. Preposed topic refers to the constituent preposed with the particle *ay* or with comma intonation while the preposed focus I take to be the first phrase in the basic clause (following any preposed topics).

The results show clearly that topic and focus have distinct syntactic properties in Tagalog and furthermore that these properties align well with larger cross-linguistic patterns. We can now move onwards to goals (ii) and (iii) as outlined in the introduction, namely describing the role of prosody in focusing, and accounting for some of the free word order in Tagalog by recourse to the prosodic component.

Table 1: Preposed topic versus preposed focus in Tagalog

	PREPOSED TOPIC	PREPOSED FOCUS
SINGLE PROSODIC PHRASE	*	✓
ASSOCIATION WITH <i>only</i> , NEG	*	✓
CO-OCCURRENCE WITH <i>WH-</i>	✓	*
RECURSIVITY	✓	*
LEFT-MOST POSITION	✓	*
LICENSES RESUMPTIVE PRONOUNS	✓	*

4 The role of prosody in focusing

If Tagalog possesses a syntactic focus position, as has been shown, when would prosody be necessary to mark pragmatic relations? It will be shown in this section that prosody functions somewhat as a last resort to mark the focus when the syntax is unable to do so. The cases where syntax is unable to mark focus are the following:

- i. double focus (as used for contrastive purposes);
- ii. situations where the ‘strong’ effects of syntactic focus are pragmatically inappropriate;
- iii. focus within non-predicational domains (i.e. NPs, PPs, relative clauses etc.).

4.1 Double focus

Double focus typically occurs in a situation where two separate constituents of a previous statement are contrasted. As mentioned earlier, one of the characteristic effects of syntactic focus is to turn the non-focus portion of the sentence into a presupposition. Thus, we can see that the syntactic position discussed earlier is incompatible with more than one focus leaving double focus to be realised prosodically rather than syntactically. We see in

(35) that A's question contains two overt arguments, *pérà* 'money' and *gúrò* 'teacher'. In a response which contrasts one of these arguments, the optimal answer focuses the contrasted constituent via the syntactic focus position. In the case of an oblique this would involve fronting as in (35a) whereas in the case of a contrasted NP as in (35b) this would involve fronting the NP in the cleft-like structure. However, when both the NP and the PP are contrasted the syntactic position is no longer available for either phrase. Rather, the focus must be realised prosodically as indicated in (35c). Any attempt to focus one constituent syntactically while focusing the other prosodically as in (35d) and (35e) results in ungrammaticality since the portion not included in the syntactic focus will always be interpreted as a presupposition.

- (35) A: *Dápat ba-ng mag-bigay ng pérà sa mga gúrò?*
 should Q-LNK AV.INF-give GEN money OBL PL teacher
 'Should one give money to the teachers?'
- a. B: *Hindè sa mga bátá ka na lang mag-bigay ng pérà.*
 neg OBL pl child 2.SG only AV.INF-give GEN money
 'No, just give money to the KIDS.'
- b. *Hindè, regalo na lang ang i-bigay mo sa mga*
 NEG gift only NOM CV.IMP-give 2.SG.GEN OBL PL
gúrò.
 teacher
 'No, just give GIFTS to the teacher.'
- c. *Hindè, mag-bigay ka na lang ng kéndi sa mga bátà.*
 NEG AV.IMP-give 2.SG only GEN candy OBL PL child
 'No, just give CANDY to the KIDS.'
- d. **Hindè, sa mga bátá ka na lang mag-bigay ng kéndi.*
 NEG OBL PL child 2.SG only AV.INF-give GEN candy
- e. **Hindè, kendi na lang ang i-bigay mo sa mga bátà.*
 NEG candy only NOM CV.IMP-give 2.SG.GEN OBL PL child

4.2 Avoiding exhaustivity

The second case where syntactic focus is inadequate is that in which the semantic effects of 'strong' focus are inappropriate within a given context. The effect referred to here in particular is exhaustive listing, which is universally associated with clefts. Exhaustive listing is the phenomenon by which, for a sentence such as *It's John who broke a window*, we obtain the reading, 'there is a unique *x* such that *x* broke a window and *x* is John'. In other words, 'John' is understood to be an exhaustive list of possible entities who broke a window within the discourse context. In the cleft-like structure, the exhaustive listing effect can be derived compositionally by virtue of the fact that the headless relative is within the domain of the definite determiner *ang*.¹⁰ Thus, the determiner properties of

¹⁰ Definite descriptions are understood to trigger uniqueness readings. Cf. Kadmon (2001) for a general overview of these issues.

ang yield the presupposition, ‘There is a unique *x* such that *x* broke a plate’ in (36). The assertion is, ‘*x* = Kengkoy’.

- (36) *Si Kéngkoy ang naka-básag ng pinggan.*
 P.NOM Kengkoy NOM AV.COM.NONV-break GEN plate
 ‘Kengkoy was the one who broke the plate (accidentally).’

Now we can see how the exhaustive listing effect might be inappropriate in a discourse such as (37).

- (37) A: *May binili ka ba?*
 EXT PV.COM:buy 2.SG Q
 ‘Did you buy anything?’
- a. B: *?Radyo ang binili ko.*
 radio NOM PV.COM:buy 1.SG.GEN
 ‘Yes, a radio is what I bought.’
- b. *Bumili ako ng radyo.*
 AV.COM:buy 1.SG GEN radio
 ‘Yes, I bought a radio.’

The sentence in (37a) can be considered felicitous through the pragmatic process known as ‘accommodation’ (Lewis 1979). Here a presupposition is accepted (i.e. accommodated) by an interlocutor as felicitous where in fact no such presupposition existed in the previous discourse. By virtue of the semantics of definite descriptions, (37a) presupposes that there was something that B bought although no such presupposition exists in A’s question. As a result, (37a) is felt to require an extra step on A’s part, and this is manifested by slightly reduced discourse felicity. In cases such as these, where focus is elicited on a constituent but where no uniqueness or exhaustivity is appropriate, prosodic focus comes into play as in (37b). This generalisation is equally true for definite foci as in (38).

- (38) A: *Meron ka ba-ng na-kilála sa New York?*
 EXT 2.SG Q-LNK PV.COM-know OBL New York
 ‘Did you meet anyone in New York?’
- B: *Na-kilála ko doon ang pangúlo.*
 PV.COM-know 1.SG.GEN there NOM president
 ‘I met the president in New York.’

The third case mentioned above in which preposing is not an option is the case of focus within non-predicational domains. The difference between predicational and non-predicational domains is essentially that between a sentence and an NP. In Tagalog, this difference manifests itself morphosyntactically in the following way: a predication is distinguished by the presence of two XPs in apposition as in (39) while an NP has its sub-constituents connected by the nasal linker as in (40).¹¹

¹¹ The linker has two allophonic variants: the velar nasal occurs following a word with a final vowel, /n/, or glottal stop; while /na/ occurs after words ending in anything else.

- (39) a. *Ma-súngit an gúro.*
ADJ-grouchy NOM teacher
'The teacher is grouchy.'
- b. *Gúro ang ma-súngit.*
teacher NOM ADJ-grouchy
'The grouchy one is a teacher.'
- (40) a. *(ang) gúro-ng ma-súngit*
NOM teacher-LNK ADJ-grouchy
'(the) grouchy teacher'
- b. *(ang) masúngit na gúro*
nom ADJ-grouchy LNK teacher
'(the) grouchy teacher'
- c. *ang nása Manílà ko-ng kapatid*
NOM LOC Manila 1.SG.GEN-LNK sibling
'my sibling in Manila'
- d. *ang ma-talino-ng kapatid ko-ng nása Manílà.*
NOM ADJ-intelligence-LNK sibling 1.SG.GEN-LNK LOC Manila
'my smart sibling in Manila'

The pragmatically salient clause-initial positions discussed earlier are predictably absent within the smaller domain of the NP. Despite the fact that NPs permit a very wide range of word-order permutations, there is no syntactically determined focus position. The topic position is also absent as shown in (41).¹²

- (41) a. *gúro-ng (*ay) ma-súngit*
teacher-LNK ay ADJ-grouchy
- b. *ma-súngit (*ay) na (*ay) gúro*
ADJ-grouchy ay LNK ay teacher

¹² However, topicalisation and focalisation of obliques within a larger nominal domain may be found in more formal language as in (i) (from traditional poetry) and (ii), respectively.

(i) *Pagbati ang sa inyo ay akin-g hatid.*
greetings NOM OBL 2.PL ay 1.SG.GEN-LNK escort
'My delivery to you is greetings.'

(ii) *ang sa akin lang mag-ma-mahal*
NOM OBL 1SG only AV-IRR:love
'the (one) who will love only ME'

Interestingly, we find that these positions are only licensed by verbs. Focused obliques may not be fronted with gerunds as shown in (iii), but rather must remain in their postverbal position as in (iv).

(iii) * *ang sa akin lang pagma-mahal*
NOM OBL 1SG only GER:love

(iv) *ang pagma-mahal sa akin lang*
NOM GER:love OBL 1SG only
'the love of only ME'

We find then, as in the previous two cases reviewed, that the prosody is the only available means of indicating focus when focus is elicited on a single constituent within a larger referring expression.

(42) A: *Sino ba? Iyon-g ma-bait na babáe?*
 who Q that-LNK ADJ-nice LNK woman
 ‘Who was it? That nice woman?’

B: *Hindí iyon, iyon-g babáe-ng ma-súngit!*
 NEG that that-LNK woman-LNK ADJ-grouchy
 ‘Not that one, the grouchy woman!’

The question which leads us to the next section is, ‘In the absence of pragmatically salient syntactic positions, what are the determinants of word order in referring expressions?’. To answer this we must first examine some basic features of Tagalog prosody.

5 Basic Tagalog prosody and its consequences

Although a truly adequate description of Tagalog prosody has yet to be produced, basic points have been described by Schachter and Otones (1972). These, in addition to preliminary acoustic studies I have undertaken, prove to be sufficient for accounting for the word-order variation discussed here. Tagalog possesses two prosodic characteristics relevant here:

- i. Focus is marked with a high tone and is linked to a stressed syllable within the focus.
- ii. There is no regular defocalisation strategy in Tagalog such that postfocal or anaphoric material is phonologically reduced (via intonation, duration or intensity).

The first characteristic is part of a larger cross-linguistic phenomenon of iconic marking and is certainly not unique to Tagalog.¹³ The second characteristic, however, sets Tagalog apart from English and is ultimately responsible for the word order variation under discussion.

It has been shown by recent work in the field of sentence prosody and the phonology–syntax interface that languages may be gainfully classified into roughly two groups according to whether or not they possess a process of phonological defocalisation (cf. Ladd 1996; Zubizarreta 1998).¹⁴ Defocalisation refers to a phonological reduction effecting

¹³ Focus spreading — that is, the ability of one pitch movement to signal focus on a larger constituent — cannot be discussed here for lack of space. This issue must be addressed in a more comprehensive account of focus in Tagalog.

¹⁴ Main contributors to the development of this grouping include Zubizarreta (1998), who formalises the prosodic difference between these two language groups as the interactions between a set of rules; Ladd (1996:175–179), who gives examples of cross-linguistic variation in defocalisation (referred to by him as deaccenting); as well as Contreras (1976) and Vallduvi (1991) who both offer evidence for prosody–syntax interactions in Spanish and Catalan respectively similar to what I claim here for Tagalog.

background information which follows the focus portion of the sentence. English, and perhaps Germanic languages in general, allow for defocalisation as in the exchange in (43).

(43) A: I had the time of my life in the jungles of Uruguay in the summer of seventy-eight.

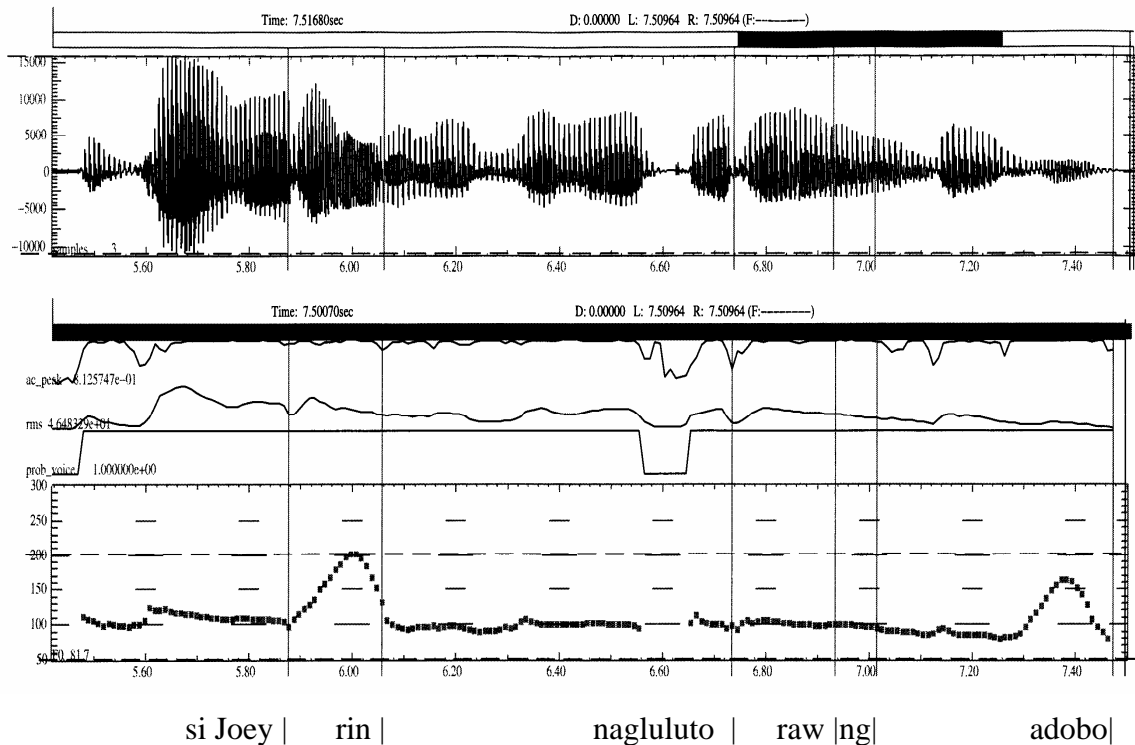
B: That's funny, I was stuck in a coal mine *in Uruguay in the summer of 1978*.

The repeated information in B's response to A, marked by italics, is felicitous in English so long as it is phonologically reduced. This defocalisation is manifested most notably by a flattening of the pitch contour and a reduction in the overall duration. In A's statement, assuming a pronunciation which consisted of a unitary intonational phrase, the nuclear stress (marked by the acute accent) would fall on the final accentable unit, 'eight'. In B's response, on the other hand, the italicised portion would be extrametrical and not be calculated in determining the placement of nuclear stress which would thus fall on the first vowel of the compound 'coal mine'. This is a phenomenon which distinguishes languages like English and German on the one hand from languages such as Spanish (Contreras 1976), European Portuguese (Cruz-Ferreira 1998) and Catalan (Vallduvi 1991) on the other hand. The first group utilises defocalisation while the latter group does not. Tagalog patterns very much with the latter group in not employing this strategy. We can see this in the pitch track below which comes from a recording of a native speaker uttering B's response in (44) as a reaction to A's statement.

(44) A: *Marámi raw ang nag-lulúto ng adóbo.*
 many REP NOM AV.COM-cook GEN adobo
 'Many people cooked adobo.'

B: *Si Joey rin nag-lulúto raw ng adóbo.*
 P.NOM Joey also AV.COM-cook REP GEN adobo
 'Joey is also cooking adobo (I heard).'

We do not find a flattening of the intonational contour which typically marks defocalisation. Instead, we see a secondary pitch movement on *adóbo* at the end of the intonational phrase regardless of the fact that it was mentioned in the immediately preceding context.



It is now important that we introduce the notion of ‘default stress’. In the default case, the most prominent stress in the sentence (i.e. the sentential or nuclear stress) coincides with the final phrasal stress of the larger intonational domain. This gives the general impression that the final phrase of a sentence is the most prominent.

If we conceive of these prosodic tendencies as constraints (Truckenbrodt 1995), we can easily imagine a situation in which a conflict arises in satisfying all of them at once. In the clearest case, this occurs when a focalised constituent, seeking to be marked intonationally, is found in a position which does not receive the default sentential stress (i.e. anywhere but the the final prosodic phrase). In this case, the inability to defocalise would cause a clash between the principle of giving prominence to the focused constituent and the default nuclear stress. In these cases, it is the syntax which rescues these sentences by placing the focalised constituents in the appropriate position in the sentence in order to receive the nuclear stress (i.e. the sentence-final position). Thus we may explain the following types of word-order alternations in the responses of (45) and (46).

(45) A: *Ano ang gusto mo-ng prutas na ma-ásim?*
 what NOM like 2.SG.GEN-LNK fruit LNK ADJ-sour
 ‘What fruit do you like that is sour?’

B: *Gusto ko ng ma-ásim na mangga.*
 want 1.SG.GEN GEN ADJ-sour LNK mango
 ‘I like sour mango.’

#? *Gusto ko ng mangga-ng ma-ásim.*
 want 1.SG.GEN GEN mango-LNK ADJ-sour

- (46) A: *Alin ang Amerikáno-ng dumating kanína?*
 which NOM American-LNK COM.AV:arrive earlier
 ‘Which American arrived earlier today?’
- a. B: *Iyo-ng Amerikáno-ng ma-tangkad.*
 that-LNK American-LNK ADJ-height
 ‘The tall American.’
- b. #? *Iyo-ng ma-tangkad na Amerikáno.*
 that-LNK ADJ-height LNK American

Extending this principle of prosodically motivated movement to the clausal level, we find that the order of phrases is also determined to an extent by the same considerations effecting NP-internal order. The order of the postverbal arguments in (45) and (46) is a function of their divergent focus marking. This is reflected concretely by the different readings obtained as a result of the interactions with the focus-sensitive word *rin* ‘also’ in (47) and (48).

- (47) *Bukod kay Ricky, ipiná-kilala ko rin*
 besides P.OBL Ricky CV.COM.CAU-know 1.SG.GEN also
kay Paolo si John.
 P.OBL Paolo P.NOM John
 ‘Besides Ricky, I also introduced John to Paolo.’
 Focus implication: I also introduced Ricky to Paolo.
- (48) *Bukod kay Ricky, ipiná-kilala ko rin*
 besides P.OBL Ricky CV.COM.CAU-know 1.SG.GEN also
si John kay Paolo.
 P.NOM John P.OBL Paolo
 ‘Besides Ricky, I also introduced John to Paolo.’
 Focus implication: I also introduced John to Ricky.

We are now in a position to explain previously unaccounted for cases of ‘scrambling’. The canonical phrasal order for Tagalog sentences with non-Actor Voice verbs and full NPs is widely accepted to be [VERB [GEN-P][NOM-P][OBL-P]]. Nevertheless, it has been noted that native speakers accept all permutations with little hesitation. Employing the methodology of eliciting speaker judgments in a pragmatic vacuum, this variation appears entirely unmotivated. Moreover, since we are dealing not with pragmatically marked syntactic positions but rather with the results of prosodic interactions with the syntax, speakers themselves do not sense any inherent focal distinctions between these minimal pairs. Equipped however with a basic understanding of Tagalog prosody we can approach this variation in a more informed manner. The ‘non-canonical’ order [VERB [NOM-P] [GEN-P]] turns out to be the preferred order provided that we can find a context in which GEN-P must be prosodically focused. Such a context is presented in (49). (Note that the use of the cleft-like construction to focus the Agent would be infelicitous here.)

- (49) A: *Kinaúsap ng báwat propeso ang mga*
 PV.COM:speak.with GEN each professor NOM PL

estudyánte, dí ba?
 student NEG Q
 ‘Each professor spoke with the students, right?’

- a. B: *Hindì. Hindí kinaúsap ang mga estudyánte*
 NEG NEG PV.COM:speak.with NOM PL student
ni Propesor Martínez.
 P.GEN Professor Martínez.
 ‘No. Professor Martínez didn’t speak with the students.’
- b. #? *Hindì. Hindí kinaúsap ni Propesor Martínez*
 NEG NEG PV.COM:speak.with P.GEN Professor Martínez.
ang mga estudyánte.
 NOM PL student

6 Conclusion

It has been shown here that while the left-periphery is often home to pragmatically salient syntactic positions — indeed this has been claimed to be a universal in verb-initial/predicate-initial languages — the right periphery also plays an important role in focalisation by virtue of being the natural position of sentential stress. In Tagalog, utilisation of the right periphery can be best characterised as a last resort for focus marking to be realised, the grammar realising this marking through the syntax when possible.

I would like to briefly address here the relationship between pragmatic focus and the Philippine voice/case system which has dominated the limelight of linguistic investigations into Tagalog and other languages. As stressed earlier, Tagalog case marking has no implicit pragmatic content. We must note however that the relationship between the voice morphology and pragmatic focus is not an arbitrary one either.

To recap, when exhaustive listing is appropriate, the voice morphology selects the thematic role of the focus which appears as the predicate in apposition to a headless relative clause. This is what has been referred to here as the cleft-like structure as shown in (50).

- (50) a. *Bátá ang lumálangoy.*
 child NOM AV.IMP:swim
 ‘It’s a child that is swimming.’
- b. *Si Juan ang lumálangoy.*
 P.NOM Juan NOM AV.IMP:swim
 ‘It’s Juan that is swimming.’
- c. *Dágat ang nilálanguyan ni Juan.*
 sea NOM LV.IMP:swim P.GEN Juan
 ‘It’s in the sea that Juan swims.’

Similarly, when the focus may be expressed as an oblique phrase, the oblique focus position may be used for the same effect as in (50). Oblique fronting may occur without regard to Voice marking.

- (51) *Sa dágat siya lumálangoy.*
 OBL sea 3.SG AV.IMP:swim
 ‘It’s in the sea that s/he swims.’

When exhaustive listing is not appropriate, an element of any type may be aligned with the right edge of the intonational phrase in order to receive the nuclear stress. This type of ‘prosodic movement’ also takes place without regard to voice marking. Therefore, we see that information structure interacts with voice only in one out of the three focusing strategies available in Tagalog. This is an indirect result of the Austronesian constraint against object extraction and should not be understood to be directly related to case marking. In Tagalog and similar Philippine languages, voice/case is determined by the definiteness of arguments and not information structure.

References

- Aldridge, Edith, 2002, Nominalization and wh-movement in Seediq and Tagalog. *Language and Linguistics* 3(2):393–426.
- Contreras, Heles, 1976, *A theory of word order with special reference to Spanish*. Amsterdam: North Holland Publishing Company.
- Cruz-Ferreira, Madalena, 1998, Intonation in European Portuguese. In Daniel Hirst and Albert Di Cristo, eds *Intonation systems*, 167–178. Cambridge University Press.
- Daneš, František, 1966, A three-level approach to syntax. In František Daneš et al. eds. *Travaux linguistiques de Prague* vol. I, 225–440. Tuscaloosa: University of Alabama Press.
- Foley, William, 1998, Symmetrical voice systems and precategoriality in Philippine languages. Manuscript available on <http://www.sultry.arts.usyd.edu.au/LFG98/austro/download/download.htm>
- Jackendoff, Ray, 1972, *Semantic interpretation in generative grammar*. Cambridge: MIT Press.
- Kadmon, Nirit, 2001, *Formal pragmatics*. Oxford: Blackwell Publishers.
- König, Ekkehard, 1991, *The meaning of focus particles: a comparative perspective*. London: Routledge.
- Kroeger, Paul, 1993, *Phrase structure and grammatical relations in Tagalog*. Stanford: CSLI Publications.
- 1998a, Tagalog clitics and clause structure. In Lourdes Bautista, ed. *Pagtanaw: Essays in honor of Teodoro A. Llamazon*, 53–72. Manila: Ateneo de Manila University Press.
- 1998b, Nouns and verbs in Tagalog: a reply to Foley. Manuscript available on <http://www.sultry.arts.usyd.edu.au/LFG98/austro/download/download.htm>.
- Ladd, Robert D., 1996, *Intonational phonology*. Cambridge University Press.
- Lambrecht, Knud, 1994, *Information structure and sentence form*. Cambridge University Press.
- Lewis, David, 1979, Scorekeeping in a language game. In R. Baurle, U. Egli, A. von Stechow, eds *Semantics from a different point of view*, 127–187. Berlin: Heidelberg.

- Naylor, Paz Buenaventura, 1975, Topic, focus and emphasis in the Tagalog verbal clause. *Oceanic Linguistics* 14:12–79.
- Reinhart, Tanya, 1982, Pragmatics and linguistics: an analysis of sentence topic. Indiana University Linguistics Club.
- Rizzi, Luigi, 1997, The fine structure of the left periphery. In Liliane Haegeman, ed. *Elements of grammar: handbook in generative syntax*, 281–337. Dordrecht: Kluwer Academic Publishers.
- Rooth, Mats, 1996, Focus. In Shalom Lappin, ed. *Handbook of contemporary semantics*, 271–297. Oxford: Blackwell.
- Schachter, Paul and Fe T. Otanes, 1972, *Tagalog reference grammar*. Berkeley: University of California Press.
- Truckenbrodt, Hubert, 1995, Phonological phrases: their relation to syntax, focus and prominence. PhD dissertation, MIT.
- Vallduví, Enric, 1991, The role of plasticity in the association of focus and prominence. *Proceedings of the Eastern States Conference on Linguistics (ESCOL)* 7:295–306.
- Wolff, John, forthcoming, *Tagalog texts*. Ithaca: Cornell University Press.
- Zubizarretta, Maria, 1998, *Prosody, focus and word order*. Cambridge: MIT Press.

9 *The richness of Madurese voice*

WILLIAM D. DAVIES

Perhaps the most conspicuous feature of western Austronesian languages is the voice system, the morphosyntactic system for identifying the semantic role of the most prominent argument in a given clause.¹ This argument has been referred to variously as subject, topic, and focus of the clause, what Starosta (1986) succinctly characterises as the ‘perpetual centre of the sentence’. This is not an issue that will concern us here; I will adopt the term ‘subject’ for convenience, a position I consider correct but will not argue for. The concern here is the richness of this system in Madurese, the fourth most widely spoken language in Indonesia with obvious similarities to Javanese, Indonesian, Balinese, Sundanese, and others. It is well known that Philippine languages such as Tagalog, Cebuano and Ilokano permit a variety of arguments of a clause to be selected as the subject, with concomitant verbal morphology. Examining the literature, it is less clear that languages such as Madurese and its closest relatives have this same richness. I argue here that the voice system of Madurese is every bit as rich as some other western Austronesian languages; by recognising that certain suffixes not universally considered part of the voice system crucially participate in encoding voice distinctions, this richness becomes apparent. While there are some proposals that these suffixes are part of the voice system (e.g. Naylor 1978) and their development has been traced from the Proto Austronesian voice system (Wolff 1996; Ross 2002), this view has not always been conspicuous in the literature. Here I propose an analysis that captures the similarities of the Madurese and Philippine systems, and present novel comparative evidence from connected discourse.

The paper is organised as follows. Section 1 briefly reviews the richness of the voice systems of Philippine languages and some other western Austronesian languages. Section 2 provides the basics of what is usually recognised as comprising the voice system of Madurese. Section 3 introduces the ‘extended’ voice system for Madurese. Section 4

¹ I would like to thank Wayan Arka and Stan Dubinsky for comments on preliminary drafts of this work. As much as I might like to place the blame on them or others, I accept responsibility for any errors or omissions. This work was supported in part by the National Science Foundation through grant SBR-9809044 to the University of Iowa. The Obermann Center for Advanced Studies provided an excellent atmosphere to work on this project as well as some of the necessary physical support.

discusses the use of voice in Madurese discourse, noting parallels with Philippine languages. Section 5 concludes, briefly describing some additional evidence from alternative object voice marking.

1 Voice in Philippine and western Austronesian languages

In Tagalog, a wide variety of clausal dependents may be selected as the perceptual centre of the clause, what Schachter (1976) refers to as the topic of the clause. As described in Schachter and Otnes 1972, the actor, object, directional, beneficiary, causee, location, instrument, measure phrase, and others may serve as topic, and in most cases the verb in the clause takes distinctive morphology. This is partially illustrated in (1), in which different arguments of the same predicate are most prominent (Schachter 1976:494–495).²

- (1) a. *Mag-salis ang babae ng bigas sa sako para sa bata.*
 AT-will.take.out T-woman G-rice D-sack B-child
 ‘The woman will take some rice out of a/the sack for a/the child.’
- b. *Aalisin ng babae ang bigas sa sako para sa bata.*
 GT.will.take.out A-woman T-rice T-sack B-child
 ‘A/The woman will take the rice out of a/the sack for a/the child.’
- c. *Aalisan ng babae ng bigas ang sako para sa bata.*
 DT.will.take.out A-woman G-rice T-sack B-child
 ‘A/The woman will take some rice out of the sack for a/the child.’
- d. *Ipag-salis ng babae ng bigas sa sako ang bata.*
 BT-will.take.out A-woman G-rice D-sack T-child
 ‘A/The woman will take some rice out of a/the sack for the child.’

The examples in (1a–d) show examples of what Schachter refers to as actor-topic (AT), goal-topic (GT), direction-topic (DT), and beneficiary-topic (BT) verbs, respectively. The voice marking is distinct in each case.

It is important to note one well-attested property of the Tagalog subject (Schachter’s ‘topic’)³ which it shares with other Austronesian languages; that is, it must be specific or definite in some sense. The subject may not be what Soemarmo (1970) characterises as [-anaphoric, -specific] in the system he develops. Specifically, the subject must be marked such that ‘the speaker assumes that [the] hearer knows the referent’ (1970:38). (In casual speech this is not adhered to in many western Austronesian languages.) Note that in all cases the *ang*-marked nominal is translated with the definite article in English. This specificity of reference is something that recurs in Austronesian languages and is shared by Madurese. There are other properties of subjects that these languages share. In many of them only subjects can be directly relativised, and in some, though not Tagalog, only the subject can be the locus of control. So, it is clear that the subject has special priority in these languages,

² The following abbreviations are used in the morphemic glosses: A actor; AV actor voice; B beneficiary; CS causative; D direction; DEF definite; DET determiner; G goal; GEN genitive; IRR irrealis; P preposition; PST past tense; RED reduplication; REL relative clause marker; OV object voice; T topic.

³ Some, including Kroeger (1993), use ‘subject’ to refer this argument in Tagalog.

Malagasy is one such language. Keenan (1976) identifies four morphologically distinct voices, which he labels active, goal, circumstantial, and intermediary; these are now generally referred to in the literature on Malagasy as actor topic (AT), theme topic (TT), circumstantial topic (CT), and the *a*-passive, respectively. Examples are given in (2) (from Paul 2000).

- (2) a. *Nanapaka ity hazo ity tamin'ny antsy i Sahondra.*
 PST.AT.cut this tree this PST.P.GEN.DET knife Sahondra
 'Sahondra cut this tree with the knife.'
- b. *Notapahin'i Sahondra tamin'ny antsy ity hazo ity.*
 PST.TT.cut.GEN.Sahondra PST.P.GEN.DET knife this tree this
 'Sahondra cut this tree with the knife.'
- c. *Nanapahan'i Sahondra ity hazo ity ny antsy.*
 PST.CT.cut.GEN.Sahondra this tree this det knife
 'Sahondra cut this tree with the knife.'
- d. *Nafatratra ny harona ny vary.*
 PST.A.stuff DET basket DET rice
 'The rice was stuffed into the basket.'

In (2a-d), the actor, theme, instrument, and material theme are subject, respectively. In each case, it is the subject that must be specific, can be relativised, or can be controlled.⁴ Other possible subjects are benefactee, location, time, purpose, manner, and others with verbs taking the circumstantial morphology (Keenan 1976).

The foregoing show that in Philippine languages and Malagasy the voice systems make a variety of arguments available to be the perceptual centre of a given clause.

2 Basic voice in Madurese

The literature on Indonesian, Javanese, and the others generally recognises two morphologically marked voices, in what have been referred to as 'Indonesian-type' systems (Wolff 1996; Ross 2002). The first is variously referred to as active, actor focus, actor voice, and others. The second has been referred to as passive, nonactor focus, object voice, and others.⁵ I will use actor voice (AV) and object voice (OV) in the discussion here.

2.1 Transitive predicates

In Madurese, actor voice occurs on syntactically transitive verbs and some intransitive verbs when the actor of the clause is the subject. The morphological manifestation of actor voice is either a nasal consonant or the prefix *a*-. The choice of the nasal or *a*- prefix appears to largely be a lexical idiosyncrasy. As is true of the actor voice prefix in Javanese, the nasal consonant assimilates to the place of articulation of the initial consonant of the verb root. Additionally, root-initial obstruents are deleted. Unlike

⁴ See, for example, Paul (2000) for evidence of these properties of the Malagasy subject.

⁵ For various characterisations of voice in various Indonesian-type languages, see Arka (1998), Bintoro (1980), Cumming (1986), Naylor (1978), Sneddon (1996), and Thomas (1980), among many others.

Javanese, in which the deletion occurs only with voiceless obstruents, all three series of consonants (voiceless, aspirated, and voiced) delete in the presence of the nasal actor voice prefix. The data in (3) illustrate this assimilation and additionally show that the base form of the morpheme *ng-* occurs with vowel-initial roots (as well as the liquids *l* and *r*).

(3)	Verb Root	Actor Voice
	<i>enom</i> 'drink'	<i>ngenom</i>
	<i>rosak</i> 'ruin'	<i>ngrosak</i>
	<i>baca</i> 'read'	<i>maca</i>
	<i>toles</i> 'write'	<i>noles</i>
	<i>kera</i> 'think'	<i>ngera</i>
	<i>bundhu</i> 'wrap'	<i>mundhu</i>
	<i>semprot</i> 'spray'	<i>nyemprot</i>

The prefix *a-* marks actor voice for a variety of syntactically transitive verbs. While Stevens (1968) reports that *a-* is used predominantly with roots with initial aspirated and voiced consonants, there seems to be a great deal of dialectal and individual variation. The roots in (4) exemplify some of those that generally take the *a-* form.

(4)	Verb Root	Actor Voice
	<i>berri</i> 'give'	<i>aberri</i>
	<i>temmo</i> 'meet'	<i>atemmo</i>
	<i>gabay</i> 'make'	<i>agabay</i>
	<i>jelling</i> 'look at'	<i>ajelling</i>
	<i>sassa</i> 'wash'	<i>asassa</i>

A number of roots admit either manifestation of actor voice, with some variation among speakers.

(5)	Verb Root	Actor Voice
	<i>kerem</i> 'send'	<i>ngerem</i> or <i>akerem</i>
	<i>buketagi</i> 'prove'	<i>mokteagi</i> or <i>abukteagi</i> ⁶
	<i>bukka</i> 'open'	<i>mokka</i> or <i>abukka</i>

The choice of *ng-* versus *a-* with transitive roots appears to be rather arbitrary for the most part; however, the eastern dialect shows some preference for *a-* and the western some preference for *ng-*. One class for which there appears to be a preference for the *a-* form across dialects is with verbs of saying. Thus one finds *abala* 'say', *akoto* 'whisper', *atanya* 'ask', *alapor* 'report', and others. Notably, however, *oca* 'speak' generally takes the *ng-* prefix, although *aoca* is attested.⁷

⁶ As a general rule, high vowels follow voiced and voiceless obstruents, but not nasals. Thus, there is a vowel alternation apparent in *mokteagi*~*abukteagi* and *mokka*'~*abukka*'. However, there is also speaker variation, and for some the high vowel of the root perseverates in the actor voice form, resulting in *mukteagi* and *mukka*'.

⁷ Predicates of this class are clearly semantically transitive but may appear to be syntactically intransitive since frequently there is no overt complement, as in (i):

(i) *Marlena a-koto' dha' Siti.*
 Marlena AV-whisper to Siti
 'Marlena whispered to Siti.'

2.2 Object voice and the bare-stem construction

Object voice in Madurese indicates that the actor of a transitive predicate has not been selected as the subject in the clause. Except for a few cases discussed in §5, object voice is invariably marked with the prefix *e-*.

Examples of the voice alternation for some transitive roots follow.

- (6) a. *Ali maca buku-na Siti.*
Ali AV.read book-DEF Siti
'Ali read Siti's book.'
- b. *Buku-na Siti e-baca Ali.*
book-DEF Siti OV-read Ali
'Ali read Siti's book.'
- (7) a. *Ba'eng a-temmo Bambang neng pasar?*
you AV-meet Bambang at market
'Did you meet Bambang at the market?'
- b. *Bambang e-temmo ba'eng neng pasar?*
Bambang OV-meet you at market
'Did you meet Bambang at the market?'
- (8) a. *Sengko' mokol Alwi.*
I AV.hit Alwi
'I hit Alwi.'
- b. *Alwi e-pokol sengko'.*
Alwi OV-hit I
'I hit Alwi.'

Finally, although ignored in most of the literature, many speakers of Madurese also accept a bare-stem form in which the theme or object is the most prominent argument of the clause.⁸ This is similar to the Indonesian construction that Sneddon (1996) refers to as 'passive type two' and Chung (1976b) characterises as 'object preposing'. In this construction, the object occurs clause-initially, and the agent in immediate preverbal position. Below are the bare-stem analogues of the clauses in (6–8).

- (6) c. *Buku-na Siti Ali baca.*
book-DEF Siti Ali read
'Ali read Siti's book.'
- (7) c. *Bambang ba'eng temmo neng pasar?*
Bambang you meet at market
'Did you meet Bambang at the market?'

However, it is indeed possible to include an overt object, in which case the *a-* prefix is still used:

- (ii) *Marlena a-koto' dha' Siti ja' Hasan badha e kamar.*
Marlena AV-whisper to Siti COMP Hasan exist at room
'Marlena whispered to Siti that Hasan was in the room.'

⁸ There are some speakers who do not accept this form when specifically asked. However, even a great many of these speakers use it in conversation.

- (8) c. *Alwi sengko' pokol.*
 Alwi I hit
 'I hit Alwi.'

It has been reported for Indonesian that bare-stem forms can occur with the agent preceding the verb and the patient following, with the agent retaining its function as the subject (for example Musgrave 2001, Voskuil 2000). One can also find clauses of the form actor–bare predicate–object, as in (9).

- (9) a. *Ali baca buku-na Siti.*
 Ali read book-DEF Siti
 'Ali read Siti's book.'
- b. *Sengko' pokol Alwi.*
 I hit Alwi
 'I hit Alwi.'

However, the sentences in (9) are not 'active', as has been claimed for Indonesian. The sentences in (9a,b) are simply analogues of (6c) and (8c) respectively, in which the subject has been postposed. There are two types of evidence for this. First, (9a,b) cannot be spoken with neutral intonation. There must be rising intonation on the predicate (*baca* or *pokol*) followed by a slight pause before the postposed patient subject (*Siti* or *Alwi*). Second, a similar restriction on indefinites is found in both the bare-stem forms. As is generally the case, indefinites are disallowed, or at least dispreferred, as subjects in the standard bare-stem form, as the ungrammaticality of (10a,b) show.

- (10) a. * *Buku Ali baca.*
 book Ali read
 (Ali read a book.)
- b. * *Oreng sengko' pokol.*
 man I hit
 (I hit someone.)

The same restriction applies to the patient when it follows the bare stem.

- (11) a. * *Ali baca buku.*
 Ali read book
 (Ali read a book.)
- b. * *Sengko' pokol oreng.*
 I hit man
 (I hit someone.)

This is not a restriction shared when actor voice morphology occurs on the predicate, as (12a, b) show.

- (12) a. *Ali maca buku.*
 Ali AV.read book
 'Ali read a/the book.'

- b. *Sengko' mokol oreng.*
 I AV.hit man
 'I hit someone.'

2.3 Intransitive predicates

As is true of other western Austronesian languages, actor voice morphology occurs with some intransitive roots as well. Although there are certainly exceptions, for the most part, actor voice morphology occurs only with 'active' roots, predicates which denote some degree of agentivity or volitionality on the part of the subject. And, for the most part, *a-* signals the actor voice with intransitives. Examples of both *a-* and *ng-*marked roots are given in (13).

(13)	Verb Root	Actor Voice
	<i>lako</i> 'work'	<i>alako</i>
	<i>berka</i> 'run'	<i>aberka</i> '
	<i>tare</i> 'dance'	<i>atare</i>
	<i>pate</i> 'die'	<i>mate</i> ⁹
	<i>abber</i> 'fly'	<i>ngabber</i>

A number of apparently agentive predicates, for example, *entar* 'go', *dhateng* 'come', *buru* 'run', take no voice morphology, as is true of non-agentive predicates such as *labu* 'fall', *badha* 'exist', and *raja* 'be big'. At first blush, this might appear to undermine the active/non-active distinction which seems to correlate with the presence or absence of actor voice morphology on intransitives. However, as Arka (1998) points out, in the main, those motion predicates which lack actor voice morphology are verbs of directed motion, e.g. 'come', 'go', rather than verbs of manner of motion, e.g. 'run', 'dance'. It has been suggested that directed motion predicates are unaccusative while manner of motion verbs are unergative (Hoekstra 1984; Levin & Hovav Rappaport 1995).¹⁰ Thus, roughly speaking, unergative predicates in Madurese take actor voice morphology, while unaccusative predicates do not, as has been argued for Balinese (Arka 1998) and Javanese (Davies 1991).¹¹

⁹ The inclusion of *mate* 'die/dead' may be viewed as controversial. Arka (1998) places this in his list of intransitives taking no overt voice morphology. However, in Madurese, *pate* is clearly the root, as seen in the object voice form of the causative 'kill', *e-pate-e* and the nominalised form *pate-na maleng rowa* 'that thief's death'.

¹⁰ The terms 'unergative' and 'unaccusative' were suggested by Perlmutter and Postal in various works to distinguish classes of intransitives based on whether the initial structure contained a subject or no subject, respectively. See Perlmutter and Postal (1984) for suggested universal criteria for membership in each class.

¹¹ There are apparent exceptions to this generalisation. Predicates such as *totop* 'be closed', *kante* 'tied up', and others that seem by all accounts and tests to be unaccusatives can take the *a-* prefix, as in:

Labang-nga a-totop.
 door-def AV-close
 'The door is closed.'

I have no explanation for this at this time.

There is also a class of stative transitive predicates (at least semantically transitive) that take no voice morphology. This class includes, but is not limited to, *percaja* ‘believe’, *yaken* ‘be sure’, *kasta* ‘regret’, *tao* ‘know’, *enga* ‘remember’, *baji* ‘hate’. For the most part, verbs of this class that take nominal complements (as opposed to clausal complements) take prepositional complements, as in (14).¹²

- (14) a. *Hasan percaja dha’ Siti.*
 Hasan believe to Siti
 ‘Hasan believes Siti.’
- b. *Amir enga’ dha’ jawab-ba.*
 Amir remember to answer-DEF
 ‘Amir remembered the answer.’

Given the facts of intransitives and stative transitives, it is reasonable to suggest that Madurese is what Klimov (1974) has referred to as an ‘active’ language, as distinct from the classification of nominative versus ergative; that is, the marking system is sensitive (at least to some degree) to the semantics of the predicate, marking ‘active’ predicates differently from ‘non-active’ predicates. (Arka 1998 has argued explicitly for this for Balinese.)

3 ‘Extending’ the voice system

Two suffixes generally omitted from discussion of Madurese voice are *-agi* and *-e*. Like their counterparts in Javanese and Indonesian (*-ake/-i* and *-kan/-i* respectively), *-agi* and *-e* are usually considered derivational suffixes. In the main, they seem to play a role similar to what are referred to as ‘applicative affixes’ in Bantu languages, and in theoretical work on Austronesian languages this term is often applied to these suffixes, for example Arka (1998), Musgrave (2001), Ross (2002). And indeed they seem to have the same function as applicatives. However, I will attempt to argue here that they work in concert with the generally recognised voice morphology in the same way that voice works in the Philippine languages.

In some environments, these suffixes appear to extend the valence of a base predicate or produce a change in the meaning of a lexical item. In (15), each is used to form a causative of a base intransitive predicate.

- (15) a. *Siti nedhung-ngagi ana’-eng.*
 Siti AV.sleep-AGI child-DEF
 ‘Siti put her child to bed.’
- b. *Ali mate-e maleng.*
 Ali AV.die-E thief
 ‘Ali killed the thief.’

In this way, *-agi* and *-e* function on a par with the Madurese causative prefix *pa-*, illustrated in (16).

¹² Frequently the preposition is omitted in speech, perhaps for reasons of redundancy. This occurs with similar predicates in Javanese (Davies 1993).

- (16) a. *Ita ma-senneng Bambang.*
 Ita AV.CS-happy Bambang
 ‘Ita makes Bambang happy.’
- b. *Siti ma-tedhung ana’-eng.*
 Siti AV.CS-sleep child-DEF
 ‘Siti put her child to bed.’

However, *-agi* and *-e* also occur in other apparent valence-altering constructions, constructions in which an argument that normally occurs in a prepositional phrase occurs as a prepositionless noun phrase, what Chung (1976a) referred to as ‘object creating’ constructions for Indonesian. This is the use that is most like the Bantu applicatives. The *-e* suffix is used predominantly when a locative argument occurs as a bare NP, as illustrated in (17b) and (18b).

- (17) a. *Atin entar dha’ Jakarta.*
 Atin go to Jakarta
 ‘Atin went to Jakarta.’
- b. *Atin ng-entar-e Jakarta.*
 Atin AV-go-E Jakarta
 ‘Atin went to Jakarta.’
- (18) a. *Siti nyaba’ buku neng meja.*
 Siti AV.put book at table
 ‘Siti put the book on the table.’
- b. *Siti nyaba’-i meja buku.*
 Siti AV.put-E table book
 ‘Siti put the book on the table.’

In (17b), *Jakarta* no longer occurs with the preposition *dha’*, when the *-e* suffix occurs. Note also that actor voice morphology obligatorily occurs on the verb; this shows that the clause in (17b) is syntactically transitive. Similarly, in (18b), the locative, *meja* ‘table’ occurs without its preposition, and in unmarked word order occurs immediately following the predicate, the position of the ‘primary’ object. Here (as in (19b)) *-e* occurs as *-i* due to the phonological environment. The suffix can also be used with human directional locatives, as with verbs of saying. This is illustrated in (19b), where the addressee *Siti* occurs as a bare NP.

- (19) a. *Ita a-bala dha’ Siti ja’ Hasan badha e kamar.*
 Ita AV-say to Siti COMP Hasan exist at room
 ‘Ita said to Siti that Hasan was in the room.’
- b. *Ita a-bala-i Siti ja’ Hasan badha e kamar.*
 Ita AV-say-E Siti COMP Hasan exist at room
 ‘Ita said to Siti that Hasan was in the room.’

The arguments that *-agi* occurs with are predominantly beneficiaries (20b) and instruments (21b).

- (20) a. *Hasan melle sepatu anyar kaangguy na'-kana*.¹³
 Hasan AV.buy shoe new for RED-child
 'Hasan bought new shoes for the children.'
- b. *Hasan melle-agi na'-kana' sepatu anyar*.
 Hasan AV.buy-AGI RED-child shoe new
 'Hasan bought the children new shoes.'
- (21) a. *Siti noles dha' tembo' bi' po'lot*.
 Siti AV.write to wall with pencil
 'Siti wrote on the wall with a pencil.'
- b. *Siti noles-sagi po'lot dha' tembo'*.¹⁴
 Siti AV.write-AGI pencil to wall
 'Siti wrote with a pencil on the wall.'

In (20b), the beneficiary, *na'kana'* 'children', occurs without its preposition, *kaangguy* 'for', and in (21b), the instrument, *po'lot* 'pencil', occurs without *bi'* 'with'.

One further environment for *-agi* is when the argument denoting the subject matter of discussion¹⁵ for some verbs of report occurs without a preposition, as with *bala* 'say' in (22b).

- (22) a. *Marlena a-bala dha' Ita parkara Bambang*.
 Marlena AV-say to Ita about Bambang
 'Marlena talked to Ita about Bambang.'
- b. *Marlena a-bala-agi Bambang dha' Ita*.
 Marlena AV-say-AGI Bambang to Ita
 'Marlena talked about Bambang to Ita.'

Now, in the (b)-sentences in (17–22), *-e* and *-agi* do not actually increase the semantic valence of the predicate. No new arguments are introduced. Only the syntactic configuration changes. Various theoretical frameworks handle these types of alternations in different ways. In some, such as Lexical-Functional Grammar and Head-Driven Phrase Structure Grammar, the alternations that occur in (17–22) are treated in the lexicon, and it is assumed that the suffixes are a reflex of the application of lexical rules. On the other hand, in frameworks such as Government Binding Theory or Relational Grammar, these types of alternations are handled in the syntax, and the suffixes are a reflex of syntactic rules. Regardless of the framework one adopts, however, in each case the now prepositionless object has been foregrounded to an extent. It has taken on a slightly more significant role in the clause, although surely not the significance of the subject.

¹³ The form *kaangguy* is the involitive form of *angguy* 'use' and is used to mark beneficiaries and some adverbial and complement clauses. Although different from typical Madurese prepositions, it is essentially used as an unanalysed form by current speakers.

¹⁴ *Noles* 'write' is one of a handful of verbs that can take both the *-agi* and *-e* suffixes. When the *-e* suffix is used, the locative argument is selected as the bare NP object, as in .

Siti noles-e tembo' bi' po'lot.
 Siti AV.write-E wall with pencil
 'Siti wrote on the wall with a pencil.'

¹⁵ This is the term used by Schachter and Otanes (1972), which I use here to make the comparison between Tagalog and Madurese more perspicuous.

Additionally, it is important to note the semantic roles of the arguments involved: direction (17), location (18), beneficiary (20), instrument (21), and subject matter of communication (22). These are precisely the types of arguments that Philippine languages have distinct voices for. And in Madurese when the predicates appear in object voice rather than actor voice, these are the arguments that occur as the subject.

- (17) c. *Jakarta e-entar-e Atin.*
 Jakarta OV-go-E Atin
 ‘Atin went to Jakarta.’
- (18) c. *Meja juwa e-saba'-i buku bi' Siti.*
 table that OV-put-E book by Siti
 ‘Siti put the book on the table.’
- (19) c. *Siti e-bala-i Ita ja' Hasan badha e kamar.*
 Siti OV-say-E Ita COMP Hasan exist at room
 ‘Ita told Siti that Hasan was in the room.’
- (20) c. *Na'-kana'-eng e-belli-agi sepatu anyar bi' Hasan.*
 RED-child-DEF OV-buy-AGI shoe new by Hasan
 ‘Hasan bought the children new shoes.’
- (21) c. *Pot'lod-da e-toles-sagi Siti dha' tembo'.*
 pencil-DEF OV-write-AGI Siti to wall
 ‘Siti wrote with a pencil on the wall.’
- (22) c. *Bambang e-bala-agi Marlana dha' Ita.*
 Bambang OV-say-AGI Marlana to Ita
 ‘Marlana talked about Bambang to Ita.’

The effect of the suffixes *-e* and *-agi* together with object voice morphology, then, is to make available as subject essentially the full range of arguments that can be selected as subject in western Austronesian languages possessing much richer voice systems. The key difference between Madurese and Philippine languages is the ‘route’ an argument takes to subjecthood. Viewed derivationally, an oblique argument in Madurese takes two steps to becoming a subject: it first becomes a bare NP object and then becomes a subject. Conversely, all arguments that are candidates for subjecthood in Philippine languages become subjects in a single move in a derivational system. This difference can be characterised as the difference between a language in which subject is the only target of syntactic revaluation, versus a language in which both subject and object are.

However, it is not necessary to take a derivational position, and in fact it obfuscates the similarity of the two languages. In a non-derivational framework of mapping arguments to syntactic positions, oblique arguments in Madurese can be mapped directly to subject position in the same way as in Philippine languages. The difference lies in the morphology that accompanies the mapping. Taking a framework neutral approach to mapping, one can envision a system in which the mapping of argument to syntactic position is specified by linking thematic roles to grammatical functions, so that a representation such as (23a) would be part of the representation of the clause in (23b).¹⁶

¹⁶ The type of analysis pursued here could easily be converted into any of the frameworks that adopt an explicit and articulated theory of mapping, such as LFG’s Lexical Mapping Theory (Bresnan & Kanerva

- (23) a. put <agent theme location>
 | |
 subject object

b. Siti put the book on the table.

The Madurese version of (24) would be identical.

- (24) a. *nyaba'* <agent theme location>
 | |
 subject object

b. *Siti nyaba' buku neng meja.*
 Siti AV.put book at table
 'Siti put the book on the table.'

Distinguishing the English and Madurese examples are the morphological rules that accompany the mapping information. In Madurese, when the agent or actor is mapped to subject, the predicate takes actor morphology. What actor morphology is appropriate with the root *saba'* 'put' (that is, *ng-* or *a-*) is an idiosyncratic property of the lexical item.

Clearly, different mapping is possible with the verb 'put'. In the object voice (or the passive in English), the theme is linked to subject. Thus, the sentence in (25b) would include (25a) as part of its representation.¹⁷

- (25) a. *esaba'* <agent theme location>
 / \
 subject

b. *Buku-na e-saba' Siti neng meja.*
 book-DEF OV-put Siti at table
 'Siti put the book on the table.'

1989) or Gerdt's Mapping Theory (Gerdt's 1993). I have chosen a theory-neutral approach here to underscore the fact that many frameworks are compatible with the proposed analysis. Arka (1998) includes a Lexical-Mapping Theory treatment of Balinese and Davies (1995) a Mapping Theory treatment of Javanese. Each framework is different from the present proposal in including a layer of structure between the thematic and GF layers. In each theory, the mapping proposed here would take place between the posited intermediate layer and the layer that feeds the syntactic representation. All morphological generalisations proposed here could thus be easily translated into either theory.

¹⁷ It is not precisely clear at this point if (25b) should have the mapping representation in (25a) or one in which the agent is linked to an object function as in (i).

(i) *saba'* <agent theme location>
 / \
 subject object

The representation in (i) is that proposed by Arka (1998) for the object voice construction in Balinese. Arka presents data from quantifier float and resumptive pronouns in left-dislocation as evidence that the actor in the object voice has the property of a 'term' (that is, it behaves like other core arguments that take no prepositional marking). Unfortunately, Madurese does not appear to have the same restriction on these properties so that some obliques as well as 'terms' are eligible. There is therefore no compelling reason to represent the actor in object voice as being linked to the object function. Not doing so further underscores the similarity between Madurese and Tagalog.

- (28) a.
- melle*
- <agent theme beneficiary>

| |
subject object

Hasan melle sepato anyar kaangguy na'-kana'.
Hasan AV.buy shoe new for RED-child
'Hasan bought new shoes for the children.'

- b.
- ebelli*
- <agent theme beneficiary>

/ /
subject

Sepato anyar juwa e-belli Hasan kaangguy na'-kana'.
shoe new that OV-buy Hasan for RED-child
'Hasan bought those new shoes for the children.'

- c.
- melleagi*
- <agent theme beneficiary>

| /
subject object

Hasan melle-agi na'-kana' sepato anyar.
Hasan AV.buy-AGI RED-child shoe new
'Hasan bought the children new shoes.'

- d.
- ebelliagi*
- <agent theme beneficiary>

/ /
subject

Na'-kana'-eng e-belli-agi sepato anyar bi' Hasan.
RED-child-DEF OV-buy-AGI shoe new by Hasan
'Hasan bought the children new shoes.'

Inasmuch as the beneficiary is mapped to a grammatical function in (28c, d), *-agi* must be affixed for the predicate to be well-formed.

A brief consideration of argument mapping in Tagalog highlights the similarities between voice marking in Tagalog and the morphology found in Madurese. First off, the analysis for the actor topic clause in (1a) is given in (29).

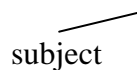
- (29) a.
- magsalis*
- <agent theme location beneficiary>

|
subject

- b.
- Mag-salis ang babae ng bigas sa sako para sa bata.*
-
- AT-will.take.out T-woman G-rice D-sack B-child
-
- 'The woman will take some rice out of a/the sack for a/the child.'

As a language that makes no apparent use of the grammatical function object, only a subject function is included in the argument-function mapping in (29a). Just as in the case of Madurese, in Tagalog mapping the agent to the subject triggers an actor topic rule, and the actor topic morphology appropriate for the verb *salis* 'take out' occurs in (29b). The goal topic clause can be given the same analysis as the object voice in Madurese.

- (30) a.
- aalisin*
- <agent theme location beneficiary>

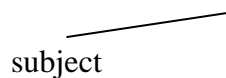


subject

- b. *Aalisin ng babae ang bigas sa sako para sa bata.*
 GT.will.take.out A-woman T-rice D-sack B-child
 ‘A/The woman will take the rice out of a/the sack for a/the child.’

Just as in the object voice structure in Madurese, the theme is linked to the subject. In Tagalog this configuration will trigger goal topic morphology on the verb. Naturally, the direction topic and the benefactive topic structure can be given similar analyses:

- (31) a.
- aalisin*
- <agent theme location beneficiary>



subject

- b. *Aalisin ng babae ng bigas ang sako para sa bata.*
 DT.will.take.out A-woman G-rice T-sack B-child
 ‘A/The woman will take some rice out of the sack for a/the child.’

- (32) a.
- ipagsalis*
- <agent theme location beneficiary>



subject

- b. *Ipag-salis ng babae ng bigas sa sako ang bata.*
 BT-will.take.out A-woman G-rice D-sack T-child
 ‘A/The woman will take some rice out of a/the sack for the child.’

The linking of the location to the subject triggers the direction topic rule in (31), and the linking of beneficiary to subject triggers the benefactive topic rule (32), which basically states that if a beneficiary is mapped to subject the appropriate benefactive topic morphology must occur on the predicate.

Under the foregoing analysis, the similarity between the voice morphology in Tagalog and the morphology in Madurese is unmistakable: when the beneficiary in Tagalog is linked to the subject function the morphological rule that affixes *ipag-* to *salis* is triggered, and when the beneficiary in Madurese is linked to the subject function the morphological rules that affix *e-* and *-agi* to *belli* are triggered. Thus, the function of *e-...-agi* here in Madurese is the same as the function of the benefactive topic voice marking in Tagalog. Viewed in this way, both *e-* and *-agi* are voice markers in the language.

The analysis of these forms in Indonesian (where there has been the most extensive discussion) is somewhat mixed. However, Naylor (1978) explicitly draws the same parallel being made here, identifying *-i* as locative focus and *-kan* as instrument focus and positing a null suffix for goal focus. For his part, Verhaar, in various works but particularly in Verhaar (1984), treats the *-i* and *-kan* suffixes in Indonesian as being part of a single system. He refers to *-i* as the ‘locative role marker’ and describes *-kan* as belonging to five ‘semantic species’, ‘benefactive’, ‘instrumental’, ‘dative-accusative’, ‘accusative-causative’, and ‘noncausative’. Importantly, he takes *men...i* and *men...kan* to be circumfixes ‘deriving *men-* verbs’ (1984:6). He thus considers the actor voice marking and the suffixes *-i* and *-kan* part of a single system, although he does not use the terminology voice to describe this. However, Verhaar does not (to my knowledge) treat

the object voice marker *di-* together with these suffixes as circumfixes. In neither case does positing a circumfix seem appropriate for Madurese. As outlined above, the actor voice, object voice, locative ‘voice’, and benefactive/instrumental ‘voice’ are all morphological manifestations of different argument/grammatical function mappings which work in complementary ways to signal the various voice possibilities in Madurese. Whether ‘voice’ is precisely the appropriate notion for this strikes me as a terminological issue. What is more important here is the recognition that the prefixes and the suffixes (in some of their instantiations) should comprise a system the end of which is the same as the voice system in languages such as Tagalog and Malagasy.

4 Some determinants of voice

While it is relatively uncontroversial that the selection of actor voice and object voice is related to identifying the ‘perpetual centre of the sentence’, what is less clear is what determines which argument a speaker wishes to identify as such. Although the focus of this paper is not to provide a definitive detailed answer to this question, I would like to explore what some of the contributing factors are and suggest that the case for the ‘extended voice’ proposal in §3 gains a bit of support from this. The observations in this section result from the examination of four examples of extended oral narratives, two representing the genre of folktales and two historical narratives.²⁰

4.1 Discourse and syntactic factors

Perhaps unsurprisingly, both discourse and syntactic factors appear to contribute to voice selection. Before examining these factors, it is necessary to delimit the corpus under consideration. The relevant data for analysis are those clauses containing syntactically transitive predicates, the domain where voice marking makes an interesting difference. Clauses with intransitive verbs (e.g. *ajalan* ‘walk’, *ngoca* ‘say’), transitive verbs that take no voice marking (e.g. *andhi* ‘have’, *tao* ‘know’) and non-verbal predicates (i.e. prepositional and nominal predicates) are not relevant to determining voice selection factors, and so they are excluded from consideration. In the texts examined here, 35%–45% of predicates are transitive verbs with voice morphology. In these texts, 45%–53% are marked for actor voice and 47%–55% for object voice (including the bare stem variety, which is rare). So the distribution is fairly equal, though object voice is slightly more common.

4.1.1 Discourse factors

Three types of discourse factors can be identified as playing a role in the determination of voice: foregrounding vs backgrounding, topic continuity, and impersonal structure.

²⁰ The narratives were transcribed from voice recordings collected from speakers in Bangkalan and Iowa City (born and raised in Bangkalan).

Hopper (1979) presents evidence that in Malay discourse²¹ foregrounded information ('the language of the actual story line') occurs in object voice ('passive') and background information ('the language of supportive material which does not itself narrate the main events') occurs in actor voice ('active'). Examination of the specific Madurese texts considered here indicates a similar pattern, although this is by no means set in stone.²² For example, in the introduction to the story *Bang Mera so Bang Pote* 'Onion and Garlic', background information occurs in the actor voice:

- (33) *Reng lake' gelle' ng-rabad-i Bang Pote.*
 person male previous AV-care-E Garlic
 'The man took care of Garlic.'
- (34) *Melana jiya bapa'-eng Bang Pote nyare bine se teppa'.*
 because this father-DEF Garlic AV.seek wife REL right
 'Because of this, Garlic's father was looking for a good wife.'

Action that moves the story forward often occurs in object voice, as in the following examples.

- (35) *Samper gelle' e-tabang bi' Bang Pote.*
 cloth previous OV-search with Garlic
A-longoe-a, aba'eng lo' bisa a-longoe.
 AV-swim-IRR she not can AV-swim
 'Garlic searched for the cloth. She wanted to swim, but she couldn't swim.'
- (36) *Daddi e-toro' bunte' neng penggir-ra.*
 so OV-follow behind at edge-DEF
 'So, she followed along the bank.'

The examples in (35, 36) also illustrate another aspect of discourse that can exert influence on voice selection — topic continuity. The subject of (35), *samper gelle'* 'the cloth', is also the subject in (36). Thus, (36) takes object voice not only because it is foregrounded material but also because it continues the topic of the discourse. This can also be seen in (37, 38), where an element introduced as an object in one sentence is taken up as the topic in the following clause.

- (37) *Oreng lake' gelle' andhi' saba tape lo' pate raja, ne'-kenne'*
 person male previous have field but not really big RED.small
 'The man had a field. But it was not really big. It was quite small.'
- (38) *Jiya se e-garap gabay ng-engon-e ana'-eng bi' aba'eng.*
 this REL OV-work make AV-raise-E child-DEF with him
 'He worked the field to feed his child and himself.'

Example (38) contains background information, so we might expect actor voice on *garap* 'work' instead of the object voice that occurs. However, the sentence takes up the object

²¹ Hopper based his observations on what seems to be relatively formal historical discourse.

²² Hopper (1979) also proposes that the generalisation is applicable to Tagalog, though according to Bell (1988), the foreground/background distinction fails to provide an explanation of voice selection in Cebuano.

of the previous sentence as its topic, and the subject selection carries with it the requirement of object voice.

A third discourse factor is the use of object voice when the narrator wishes to de-emphasise the agent or omit it altogether as in a story-telling mode or in historical narrative, similar to the function of passive in English narration.

- (39) *E-tanam-e padhi saba-na.*
 OV-plant-E rice field-DEF
 ‘The field was planted with rice.’
- (40) *E-careta-agi gitek gelle’ terros noju dha’ Madu Oro.*
 OV-story-AGI raft previous continue north to Madu Oro
 ‘It was told that the raft then continued north to Madu Oro.’

In both of these sentences, the agent is omitted. In (39), the agent is unimportant to the action. The focus is on the field *sabana* and the agent can be inferred from the context and hence is omitted. The sentence in (40) contains a story-telling device. The agent is nonspecific, and authenticity of the report is supported through reference to historical precedent.

4.1.2 Syntactic factors

Syntactic factors also play a role in determining voice selection, at times overriding discourse factors and at times forcing voice selection to coincide with voice selection determined by discourse factors.

As previously established, only subjects may be relativised or clefted. Thus, the voice of the verb of the relativised or clefted clause must be consistent with this constraint. This accounts for object voice selected in background sentences (41) and actor voice selected in foreground sentences (42).

- (41) *Bakto jiya keya barang dagang-an-na se e-tangge’ dhari*
 time this also thing trade-NOM-DEF REL OV-bring from
Palembang e-juwal dha’ Sumenep.
 Palembang OV-sell to Sumenep
 ‘Also at this time, merchandise brought from Palembang was sold to Sumenep.’
- (42) *Aba’eng e-gigir-in bi’ ma’ butha se ngakan-an oreng.*
 she OV-mad-E with father giant REL AV.eat-IT person
 ‘Father giant, who ate people, got mad at her.’

Although the sentence in (41) provides background information in the story, *tangge’* ‘bring’ occurs in the object voice since it is the patient of this verb that is the head of the relative clause. (The matrix verb *juwal* ‘sell’ occurs in the object voice since the agent is unknown, hence non-specific.) In (42), the matrix verb *gigir* ‘mad’ is in the expected object voice since this is a sentence that moves the storyline along. However, since the head of the relative clause *ma’ butha* ‘father giant’ is the agent, *kakan* ‘eat’ occurs in the actor voice.

Obligatory control structures also influence voice selection. There is a marked preference in Madurese for the controllee in a control construction not only to be the subject of the embedded clause, but also the agent.²³ Therefore, the majority of embedded transitive verbs in control structures occur in the actor voice, again regardless of whether the clause is part of backgrounded (43) or foregrounded (44) information.

- (43) *Aba'eng terro lo' a-berri'-a tao dha' Bang Pote.*
 she want not AV-give-IRR know to Garlic
 'She didn't tell Garlic about this.'
- (44) *Bangsa Cara e-soro raja nga-bine-e Raga Padmi.*
 Bangsa Cara OV-order king AV-wife-E Raga Padmi
 'The king ordered Bangsa Cara to marry Raga Padmi.'

In (43), the complement of the predicate *terro* 'want' is *lo' aberri'a tao dha' Bang Pote* 'not tell Garlic', and the embedded predicate *berri* 'give' occurs in the actor voice. Here both the controller *aba'eng* 'she' and the controlled position are actors. In (44), the controller is the patient subject *Bangsa Cara* and the controlled position, the subject of *ngabinee* 'marry (a woman)' is an agent actor, despite the fact that the sentence serves to move the story line along.

Adverbial clauses also appear to exert influence over the voice of the verb. For the most part, when the subject of the adverbial clause is coreferent with a matrix clause argument, the predicate occurs in actor voice.

- (45) *Mare ng-rengkes-se, duli kalowar e-sambi klambi-na se kotor.*
 finish AV-tidy-E soon go.out OV-bring cloth-DEF REL dirty
 'After cleaning the room, she quickly went out and brought his dirty clothes.'

In (45), the controlled subject of the adverbial clause *mare ngrengkesse* 'after cleaning up' is coreferent with the understood agent of the matrix clause. However, in the majority of the cases where there is no coreferent argument in the adverbial clause, the verb is marked for object voice, as is *bukka* 'open' in the adverbial clause *sedang ebukka' labangnga bi' bapa'-eng Bang Pote* 'when Garlic's father opened the door' in (46).

- (46) *Sedang e-bukka' labang-nga bi' bapa'-eng Bang Pote,*
 when OV-open door-DEF with father-DEF Garlic
ebu'-eng bi' Bang Mera temmo-na la mate.
 mother-DEF with Onion meet-DEF already AV.die
 'When Garlic's father opened the door, the mother and Onion were already dead.'

4.2 -agi and -e in discourse

The occurrence of the suffixes *-agi* and *-e* is somewhat rare in the texts; however, where they occur is instructive of the role they play in voice marking. Recall that according to

²³ While it is grammatically acceptable for non-subject agents to be controlled, it is rare to find this construction in daily conversation.

the analysis proposed here, these morphemes have two functions: to derive causatives or to ‘inflect’ verbs for voice, as distinguished in (47) and (48) respectively.

- (47) *Aba’eng mate-e Bangsa Cara.*
 he AV.die-E Bangsa Cara
 ‘He killed Bangsa Cara.’
- (48) *Aba’eng ngerem-e ana’-eng Marlana pesse.*
 He AV.send-E child-DEF Marlana money
 ‘He sent Marlana’s child money.’

Before proceeding, it is important to clarify the sense in which these suffixes have non-causative uses. Wayan Arka (pers. comm.) points out that there are Balinese predicates such as *atur* ‘offer’ which can occur only with the suffix *-ang* or *-in* (the Balinese counterparts of *-agi* and *-e*, respectively). Thus, in actor voice, Balinese allows only (49) or (50). A form *ngatur* does not exist as a well-formed transitive predicate.

- (49) *Cang ngatur-ang pipis-e sig Ida Peranda-ne.*
 I AV.offer-ANG money-DEF to ART priest-DEF
 ‘I offered the money to the priest.’
- (50) *Cang ngatur-in Ida Peranda pipis.*
 I AV.offer-IN ART priest money
 ‘I offered money to the priest.’

What is crucial in these examples is the fact that one of the morphological suffixes is obligatory, in the same way that actor voice morphology is obligatory. This supports the notion that these suffixes are a part of the voice morphology in Balinese. Interestingly, the predicate ‘give’ in Javanese exhibits the same behavior. While there are forms *ngekekake* (or the dialectal variants *ngekekna* and *ngekekne*) and *ngekeke*, there is no form **ngeke(k)*. While I am aware of no examples like this from Madurese, the Balinese and Javanese facts offer some limited support to the notion that *-agi* and *-e* function as part of the voice morphology of the language. It is such a non-causative use that is of importance here.

Note that all verbs taking these suffixes also take either actor or object voice morphology. In the sample, 40% of all verbs taking *-agi* or *-e* (in both causative and noncausative functions) occur with actor voice morphology and 60% occur with object voice morphology. This distribution is comparable to the overall distribution of AV and OV morphology on transitive verbs. However, when the causatives are eliminated, only 20% of these verbs take AV morphology. The overwhelming majority, 80%, occur with OV morphology.

This fact is important because it suggests that *-agi* and *-e* play an important role in voice marking. First, it should be noted that the causative uses of *-agi* and *-e* are largely represented by two verbs that recur frequently due to the subject matter of the stories, *mate-e* ‘kill’ and *ngedhingngagi* ‘listen to’. Second, the fact that the majority of these verbs occur with OV morphology seems to indicate that the primary role of these morphemes in non-causative use is to make it possible to select as subject an oblique argument, whether in support of a discourse function or for purely syntactic reasons. Some examples include making the subject matter of discussion subject, as with *e-tanya-agi* ‘ask about’ and *e-careta-agi* ‘tell about’. In (51), the subject matter *dha’iya* ‘like this’ is the subject of the clause, required by topic continuity from the preceding sentence.

- (51) *E-tanya-agi dha'iya, kancil mekker.*
 OV-ask-AGI like.this deer AV.think
 'When asked this, the deer thought.'

Others make the addressee the subject, as with *e-bala-i* 'say to' and *e-tanya-e* 'ask to'. In (52), the addressee is the subject. Here the addressee is understood.

- (52) *So ma' butha e-tanya-e.*
 by giant AV-ask-E
 'She was asked by the giant.'

Yet other examples illustrate making the location subject, as with *e-entar-e* 'go into' and *e-tanam-e* 'plant in'. In (53), discourse reasons dictate the object voice and the structure with a location subject.

- (53) *Daddi so ebu'-eng Bang Mera e-entar-e Bang Pote.*
 become by mother-DEF Onion OV-go-E Garlic
 'So, Onion's mother went to Garlic.'

A final point of comparison with Philippine languages is suggestive. Bell (1988) provides some statistics regarding voice selection in Cebuano texts. It turns out that the statistics from the four Madurese texts considered here are remarkably similar. Bell's statistics reveal the following distribution of voice (using Bell's terminology):

- (54) Cebuano voice distribution
- | | |
|------------------|-------|
| Agent-topic | 50.7% |
| Object-topic | 31.6% |
| Reference-topic | 12.4% |
| Instrument-topic | 5.0% |
| Total goal-topic | 49.3% |

It should be noted that Goal-topic is a cover term employed by Bell to refer to all voice morphology other than Agent-topic morphology. The following table provides comparable statistics for the Madurese texts, with the OV figure reduced by the number *-agi* and *-e* plus OV combinations and the *-agi* and *-e* figures excluding causative and AV usage.

- (55) Madurese voice distribution
- | | |
|--------------------|-------|
| Actor voice | 48.0% |
| Object voice | 38.3% |
| <i>-e</i> | 9.7% |
| <i>-agi</i> | 4.0% |
| Total object voice | 52.0% |

The 'object voice' number (38.3%) should be compared with Bell's 'object-topic' (31.6%) and the 'total object voice' number (52%) with Bell's 'total goal-topic' (49.3%). While these are casual statistics and should be approached cautiously, all the totals are remarkably similar across the two languages. This similarity certainly lends some credence to the hypothesis that the designated morphology is serving a similar function in the two languages.

5 A final bit of evidence and a conclusion

Alternative object voice morphology available with a subset of predicates provides a further bit of evidence for the proposal that *-agi* and *-e* contribute to the voice system in Madurese. With certain predicates object voice is marked with the prefix *eka-* rather than *e-*. These are all predicates that take prepositional objects rather than simple noun phrase objects. Included in this group are verbs such as *tao* ‘know’, *percaya* ‘believe’, *peggel* ‘angry’, *loppa* ‘forget’, and a host of others. Examples are in (56).

- (56) a. *Hasan eka-peggel-le bi’ Siti.*
 Hasan OV-angry-E with Siti
 ‘Siti is angry with Hasan.’
- b. *Jawap-a eka-loppa-e bi’ sengko’.*
 answer-DEF OV-forget-E with me
 ‘I forgot the answer.’

Interestingly, there is some variation among speakers and within speakers regarding the use of the *eka-* prefix and the voice suffixes. Therefore, one finds the alternations in (57–59).

- (57) a. *Hasan e-bala-agi bi’ Siti dha’ Marlana.*
 Hasan OV-say-AGI with Siti to Marlana
 ‘Siti talked about Hasan to Marlana.’
- b. *Hasan eka-bala bi’ Siti dha’ Marlana.*
 Hasan OV-say with Siti to Marlana
 ‘Siti talked about Hasan to Marlana.’
- (58) a. *Bambang e-kasta-e Ita ja’ aba’eng mangkat.*
 Bambang OV-regret-E Ita COMP he leave
 ‘Ita regretted about Bambang that he left.’
- b. *Bambang eka-kasta Ita ja’ aba’eng mangkat.*
 Bambang OV-regret Ita COMP he leave
 ‘Ita regretted about Bambang that he left.’
- (59) a. *Tarentan-na e-bellis-i Hasan.*
 old.sibling-DEF OV-hate-E Hasan
 ‘Hasan hates his older brother.’
- b. *Tarentan-na eka-bellis Hasan.*
 old.sibling-DEF OV-hate Hasan
 ‘Hasan hates his older brother.’

In the (a) sentences we find the standard object voice morpheme *e-* along with the suffix *-agi* (57) or *-e* (58, 59). In the (b) sentences, however, we find only the object voice morpheme *eka-* and neither *-agi* nor *-e*. In these sentences the morpheme *eka-* appears to be doing all the work of marking the voice; that is, *eka-* indicates that the subject of the sentence is an oblique argument. This is much like the richer voice system of the Philippine languages, and provides some additional suggestive evidence that despite the surface appearance and traditional analyses the voice system of Madurese is every bit as rich as some of its western Austronesian relatives.

References

- Arka, I Wayan, 1998, From morphosyntax to pragmatics in Balinese: a Lexical-Functional Grammar approach. PhD dissertation, University of Sydney.
- Bell, Sarah J., 1988, Voice and foreground/background in Cebuano. In Richard McGinn, ed. *Studies in Austronesian linguistics*, 427–440. Athens, OH: Ohio University.
- Bintoro, 1980, Javanese transitive verbs: a tagmemic analysis. *NUSA Linguistic Studies in Indonesia and Languages in Indonesia* 8.
- Bresnan, Joan W. and Jonni Kanerva, 1989, Locative inversion in Chichewa: a case study of factorisation in grammar. *Linguistic Inquiry* 20:1–50.
- Chung, Sandra A., 1976a, An object-creating rule in Bahasa Indonesia. *Linguistic Inquiry* 7:41–87.
- 1976b, On the subject of two passives in Indonesian. In Li, ed. 1976:57–98.
- Cumming, Susana, 1986, Word order change in Malay. In Geraghty et al., eds 1986:97–111.
- Davies, William D., 1991, Against an ergative analysis of eastern Javanese. *ESCOL '90*, 79–89.
- 1993, Javanese subjects and topics and psych verbs. *Linguistics* 31:239–277.
- 1995, Javanese adversatives, passives, and Mapping Theory. *Journal of Linguistics* 31:15–51.
- Geraghty, Paul, Lois Carrington and S.A. Wurm, eds, 1986, FOCAL I: *Papers from the Fourth International Conference on Austronesian Linguistics*. Canberra: Pacific Linguistics.
- Gerds, Donna B., 1993, Mapping Halkomelem grammatical relations. *Linguistics* 31:591–621.
- Hoekstra, Teun, 1984, *Transitivity*. Dordrecht: Foris.
- Hopper, Paul J., 1979, Aspect and foregrounding in discourse. In Talmy Givon, ed *Discourse and syntax*, 213–241. New York: Academic Press.
- Keenan, Edward L., 1976, Remarkable subjects in Malagasy. In Li, ed. 1976:247–301.
- Klimov, Georgij A., 1974, On the character of languages of active typology. *Linguistics* 131:11–25.
- Kroeger, Paul R., 1993, *Phrase structure and grammatical relations in Tagalog*. Stanford: CSLI Publications.
- Levin, Beth, and Malka Rappaport Hovav, 1995, *Unaccusativity at the syntax-lexical semantics interface*. Cambridge, MA: MIT Press.
- Li, Charles N., ed., 1976, *Subject and topic*. New York: Academic Press.
- Musgrave, Simon, 2001, Non-subject arguments in Indonesian. PhD dissertation, University of Melbourne.
- Naylor, Paz Buenaventura, 1978, Toward focus in Austronesian. In S.A. Wurm and Lois Carrington, eds *Second International Conference on Austronesian Linguistics: proceedings*, 395–442. Canberra: Pacific Linguistics.
- Paul, Ileana, 2000, Malagasy clause structure. Unpublished McGill University PhD dissertation.

- Perlmutter, David M. and Paul M. Postal, 1984, The 1-Advancement Exclusiveness Law. In David M. Perlmutter and Carol Rosen, eds, *Studies in Relational Grammar 2*, 81–125. Chicago: University of Chicago Press.
- Ross, Malcolm, 2002, The history and transitivity of western Austronesian voice and voice-marking. In Fay Wouk and Malcolm Ross, eds *The history and typology of western Austronesian voice systems*, 17–62. Canberra: Pacific Linguistics.
- Schachter, Paul, 1976, The subject in Philippine languages: topic, actor, actor-topic, or none of the above. In Li, ed. 1976:491–518.
- Schachter, Paul and Fe Otones, 1972, *A Tagalog reference grammar*. Berkeley: University of California Press.
- Sneddon, James, 1996, *Indonesian: a comprehensive grammar*. London: Routledge.
- Starosta, Stan, 1986, Focus as recentralisation. In Geraghty et al, eds 1986:73–95.
- Soemarmo, Marmo, 1970, Subject-predicate, focus-presupposition, and topic-comment in Bahasa Indonesia and Javanese. Unpublished UCLA PhD dissertation.
- Stevens, Alan, 1968, *Madurese phonology and morphology*. American Oriental Society.
- Thomas, Michael R., 1980, Verb affixes and focus in Bahasa Indonesia. In Paz B. Naylor, ed. *Austronesian studies*, 63–70. Ann Arbor: University of Michigan.
- Verhaar, John M.W., 1984, Affixation in contemporary Indonesian. In Bambang Kaswanti Purwo, ed. *Towards a description of contemporary Indonesian: preliminary studies*, Part I, 1–26. Jakarta: NUSA.
- Voskuil, Jan, 2000, Indonesian voice and A-bar movement. In Ileana Paul, Vivianne Phillips and Lisa Travis, eds *Formal issues in Austronesian linguistics*, 195–213. Dordrecht: Kluwer.
- Wolff, John U., 1996, The development of the passive verb with pronominal prefix in western Austronesian languages. In Bernd Nothofer, ed. *Reconstruction, classification, description: festschrift in honor of Isidore Dyen*, 15–40. Hamburg: Abera.

10 *Topic continuity, voice and word order in Pendau*

PHIL QUICK

1 Introduction*

This study of topic continuity in Pendau examines several issues. First, does the inverse construction behave more like a typical passive or more like a transitive? Second, can the choice between a *ni*- verb and a *nong*- verb be predicted on the basis of discourse-level information, and, if so, on the basis of what parameters? Third, does Pendau have a true passive construction? Fourth, is the variation of SV/SVO with VS/VOS a result of topic continuity or of other criteria?

The main findings in this paper concern the criteria for selecting active voice and inverse voice transitive constructions in Pendau, and show that both constructions are basic transitives. After a brief introduction to some of the grammar of Pendau, I present a quantitative analysis based on Givón's (1983, 1994) approach to topic continuity. After this, further quantitative results are presented, identifying the functional equivalent of passive in Pendau and addressing the problem of word-order variation (SV/SVO versus VS/VOS).

2 Grammatical background

2.2 Active voice and inverse voice

Transitive verbs can be inflected in either active voice or inverse voice without a change in transitivity.¹ Examples (1) and (2) contrast the *nong*- and the *ni*- transitive verb forms. The verbs in these sentences can be interpreted as *primary transitive verbs* (Andrews

* Pendau is a western Austronesian language group of about 3000–5000 speakers in Central Sulawesi, Indonesia. Interlinear abbreviations used in this paper are: 1SG first singular person; AB absolute case; AV active voice; DY dynamic verb class; GE genitive case; IV inverse voice; PN proper noun; PT primary transitive verb class; RE realis modality; SF augmented stem former; and ST stative verb class.

¹ See Quick (1997, 1999, 2001, 2003) for the background and basis for the pragmatic inverse voice construction and the analysis on which this paper is based.

1985), and they represent active voice and inverse voice clause constructions respectively. Transitive clauses which have an agent (A) and a patient (P) argument such as these can be considered to be prototypical transitive constructions.² In the free translation the capitalised NP indicates the pivot or subject in Pendau.³ The two differences between the a and b sentences in (1)–(2) are in the verbal prefix and in the case markers on the postverbal arguments (see Quick 1997, 2003 for the full discussion of these as active voice and inverse voice respectively). Compare examples (1)–(2) with Table 1, which shows that inverse voice results from the realignment of the macro roles. Examples (1a) and (2a) show the SVO word order for active voice and inverse voice respectively. Examples (1b) and (2b) contrast the same clauses in their VOS word order variants and demonstrate that the same meaning can occur in both active voice and inverse voice in either word order.

Table 1: Macro role realignment

Active voice	Subject <i>actor role</i>	V	Object <i>undergoer role</i>
Inverse voice	Subject <i>undergoer role</i>	V	Object <i>actor role</i>

- (1) a. *Siama'u nonuju siina'u.*
si=ama='u N-pong-tuju si=ina='u
 PN/AB=father=1SG/GE RE-SF/PT-send PN/AB=mother=1SG/GE
Pivot=A non-pivot=P
 'MY FATHER sent my mother.'
- b. *Nonuju siina'u siama'u.*
N-pong-tuju si=ina='u si=ama='u
 RE-SF/PT-send PN/AB=mother=1SG/GE PN/AB=father=1SG/GE
non-pivot=P Pivot=A
 'MY FATHER sent my mother.'
- (2) a. *Siama'u nituju niina'u.*
si=ama='u ni-tuju ni=ina='u
 PN/AB=father=1SG/GE IV/RE-send PN/GE=mother=1SG/GE
Pivot=P non-pivot=A
 'My mother sent MY FATHER.'

² Following Dixon (1979, 1994) and Andrews (1985), agent and patient can be represented as prototypical arguments that have been symbolised as A and O respectively. Others, such as Comrie (1989), have used the same idea with the partially different labels of A and P respectively. In this paper I will follow Comrie's labels A and P to refer to the basic argument positions of transitive clauses.

³ The identification of subject is based on a methodological procedure which requires identifying the pivot first in two clauses of the same sentence (for the mechanics of this procedure see Quick 2003). The use of the term 'pivot' in this paper reflects this preliminary procedure when it is used before identifying the grammatical subject in Pendau. For purposes of understanding this paper the terms 'pivot' and 'grammatical subject' may be understood to mean the same thing.

- b. *Nituju niina'u siama'u.*
ni-tuju ni=ina='u si=ama='u
 IV/RE-send PN/GE=mother=1SG/GE PN/AB=father=1SG/GE
non-pivot=A Pivot=P
 'My mother sent MY FATHER.'

The transitive clauses are contrasted with intransitive clauses in examples (3) and (4). The difference between a and b in each of these examples contrasts SV and VS word order respectively.

- (3) a. *SiYusup neriing.*
si=Yusup N-pe-riing
 PN/AB=Joseph RE-SF/DY-bathe
Pivot= S_A
 'Joseph bathed.'
- b. *Neriing. siYusup.*
N-pe-riing si=Yusup
 RE-SF/DY-bathe PN/AB=Joseph
Pivot= S_A
 'Joseph bathed.'
- (4) a. *SiYusup nanabu.*
si=Yusup no-nabu
 PN/AB=Joseph ST/RE-fall
Pivot= S_P
 'Joseph fell (down).'
- b. *Nanabu. siYusup.*
no-nabu si=Yusup
 ST/RE-fall PN/AB=Joseph
Pivot= S_P
 'Joseph fell (down).'

Table 2 compares the word orders for basic transitive clauses and their associated transitive verb affixes in Pendau. Each verb type has a *rigid* argument position that is postverbal, and each verb type has a *flex* argument position that is in either (a) a preverbal position or (b) in a postverbal position which must follow the *rigid* argument position. The *flex* positions are marked in Table 2 by circles around the arguments which have more than one word order position. However, what is relevant is that this pattern suggests that both the *nong-* verb clause and the *ni-* verb clause have one single underlying word order. The *flex* position is identified as that of the *pivot* since preverbally this is the position in which the *pivot* occurs in relative clauses, and the *rigid* position as that of the *non-pivot*. Word order variation is a pragmatic discourse function that is discussed in §5.

Table 2: A and P argument positions in Pendau transitive clauses

1.	Ⓐ	<i>nong-V</i>	P	
2.		<i>nong-V</i>	P	Ⓐ
3.	⒫	<i>ni-V</i>	A	
4.		<i>ni-V</i>	A	⒫

Turning now to noun phrases, Pendau has two pronoun sets and a noun phrase marking system as seen in Table 3. Noun phrases are either common nouns or proper nouns. There are two sets of pronouns and noun phrase markers, which I will refer to as absolute (AB) and genitive (GE). The distribution of the absolute and genitive NPs in Pendau is different from the expected traditional usage. Genitive NPs are used in two distinct syntactic positions (Table 3): (1) genitive noun phrases, and (2) the A argument of inverse voice *Absolute* NPs are used in all other core argument positions (i.e. ‘elsewhere’), including second objects of ditransitive clauses (except instrumental NPs), the objects of prepositional phrases, and in both argument positions of equative clauses and copula clauses.

Table 3: The core case system in Pendau (pronouns and noun phrase markers)

		Absolute	Genitive ⁴	Instrument
SG.	1	<i>a'u</i>	= <i>'u</i> (<i>'u-</i> , <i>no'u-</i>)	–
	2	<i>oo</i>	= <i>mu</i> (<i>mu-</i>)	–
	3	<i>io</i>	= <i>nyo</i>	–
PL.	1 INC	<i>ito</i>	= <i>to</i>	–
	1 EXC	<i>ami</i>	<i>mami</i>	–
	2	<i>emu</i>	<i>miu</i>	–
	3	<i>jimo</i>	<i>nijimo</i>	–
Proper Nouns		<i>si=</i>	<i>ni=</i>	–
Common Nouns		∅ / (<i>u=</i>)	<i>nu=</i>	<i>nu=</i>

Finally, also shown in Table 3 is the instrument case marker *nu*. Instruments are marked by *nu* when they are not the pivot of the clause. Although *nu* appears on one hand to be preposition-like, it behaves more like a core argument (second object marker) than an oblique argument marker. Andrews (1985:128–130) discusses the ambiguous status of instrument and second objects in general.

⁴ The genitive pronoun set also includes the fronted pronouns *'u-* and *mu-* for first and second person respectively, effectively becoming verbal prefixes. The genitive pronoun set is a mixed set, some are enclitics, and some are free words (distinguishable by phonological criteria).

3 Topic (referential) continuity

3.1 Introduction to topic continuity in Pendau

When people talk, they usually talk about particular things or topics. Some referents are talked about more often than others and for longer stretches at a time. Topic continuity, then, refers not only to whether or not a particular referent remains central to a discourse over a long period of time, but also to how a particular referent which is mentioned more than once is tracked or referred to later on within a discourse. Linguists have often noticed that more continuous referents are realised linguistically in more different ways than less topical elements. Sometimes highly topical elements need not even be realised overtly at all.

Topic continuity manifests itself in a variety of ways in Pendau, as listed in (5).⁵ This ranks the manifestations of continuity from those associated with the most continuous referents to the least continuous (from top to bottom respectively). This ranking fits well with Givón's (1990:917) observations that 'referents that are *already active* require minimal coding.'

- (5) zero anaphora
 agreement in abilitative verbs
 conjugational portmanteaux pronouns (defective paradigm in inverse voice)
 clitic pronouns (genitive pronouns, including mixed clitics and free pronouns)
 independent pronouns (absolute case pronouns)
 full noun phrase
 left-dislocation

3.2 Methodology and background of the four Pendau texts quantified

In order to examine how topic continuity manifests itself in Pendau narrative genres, four narrative texts have been analysed in some detail. The following sections (§3.3–§3.4) give the results of the analysis of the texts with a total of 746 clauses following specific quantitative methodologies proposed by Givón and others (Givón, ed. 1983 and 1994). These four texts are referred to as Mtext 1–4 (or sometimes simply as M1–M4). Table 4 shows the total number of clauses for each text, the title, the author, and whether the text was recorded or composed.⁶ The texts were coded following a tagging convention that has been developed for this kind of analysis, and that can be quantified quickly by a computer program (the Multilinear Discourse Analysis software developed primarily by the author and reported on in Quick 1996). Each participant in every clause was identified and then marked according to its realisation (e.g. zero anaphora or a particular noun-phrase type), its macro-roles/semantic roles and grammatical relations, etc. Other database fields mark categories such as event/non-event, quotation or speech margin, word-order type, etc.

⁵ Compare this with the typological listing in Payne (1997:345).

⁶ 'Composed' means that a story has been authored by a native Pendau speaker who has had some training in writing/authoring skills as well as considerable natural talent to perform what I have judged to be some of the best narrative material in Pendau.

Table 4: Texts analysed and quantified

	Text name in English	Style	Author	# of clauses
Mtext 1 (M1)	The story of the seven men in search of rattan and the two that got lost (fktale01.txt)	oral (recorded)	SiDidi	94
Mtext 2 (M2)	The story of the monkey and the turtle who were friends (turtle.pin)	composed	Josep Piri	221
Mtext 3 (M3)	The story of the flesh-eater who found a monkey to be his friend (troll.int)	composed	Josep Piri	354
Mtext 4 (M4)	The story of the pelican who swallowed the grandfather's grandson (tambao.tst)	composed	Josep Piri	77

3.3 Text profiles according to NP types and participants

This section provides a preliminary look at the raw profile of the four texts before going on to examine specific measures of topic continuity in each text. Figures 1–4 show raw numbers of referent realisations in different linguistic guises for each text. The abbreviations are defined as follows: DEM is a NP coded simply as a demonstrativeN1 a full noun phrase in the absolute case, N2 a full noun phrase in the genitive case (as agent), P1 an absolute case pronoun, P2 a genitive case pronoun (as agent), P3 a pronominal affix (associated with the genitive case pronouns since they function as agent), QN/QF quantifiers or quantifier floating adverb functioning as a simple NP, RC relative clauses, and ZR, marks zero anaphora.⁷ The different graph bars indicate different participants or characters in a story for a particular NP type.

Consistent among these texts are high total scores for the absolute noun phrases (N1), the absolute pronoun case (P1) and zero anaphora. From this raw data one can already note that the high scores for N1 and P1, and those for N1 and ZR, are roughly equivalent. This is significantly different from what the literature generally says about argument realisation. This will be discussed further in the light of topic continuity in later sections.

⁷ In the later quantification methods it was found to be adequate to collapse the lower frequency NPs annotated here as QN/QF, RC, DEM into the appropriate N1 or N2 (usually the former). In my counting of RCs I am quantifying only the head noun as a token that is modified by a RC.

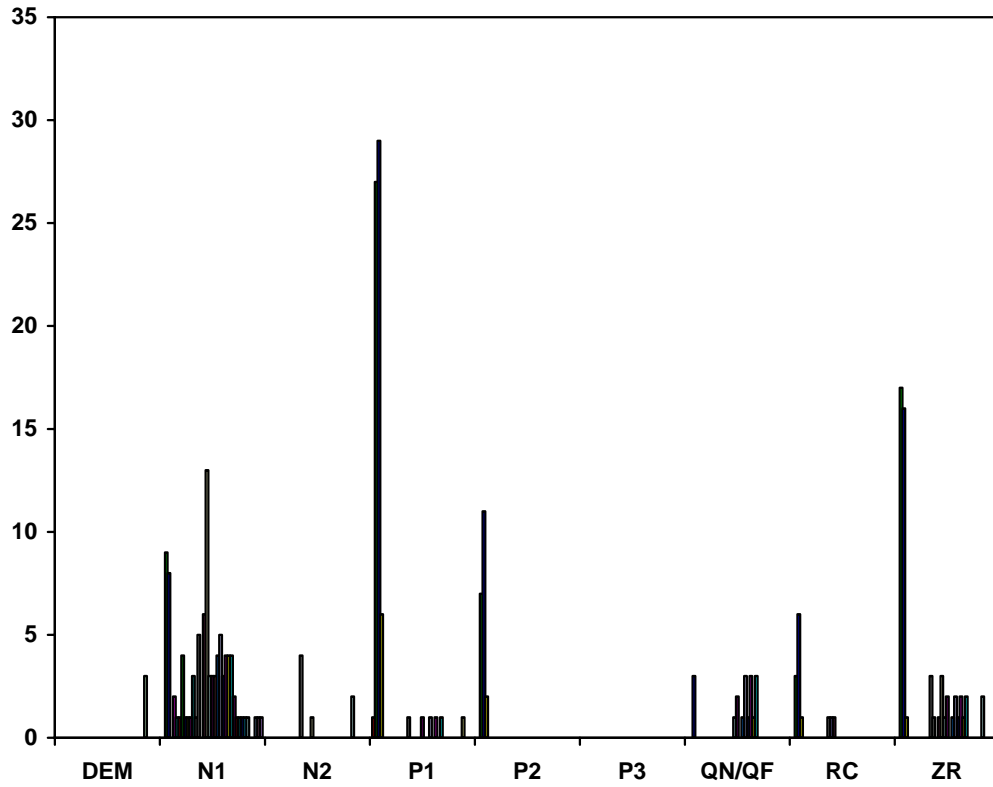


Figure 1: Participants 1-37 NP profile of Mtext 1

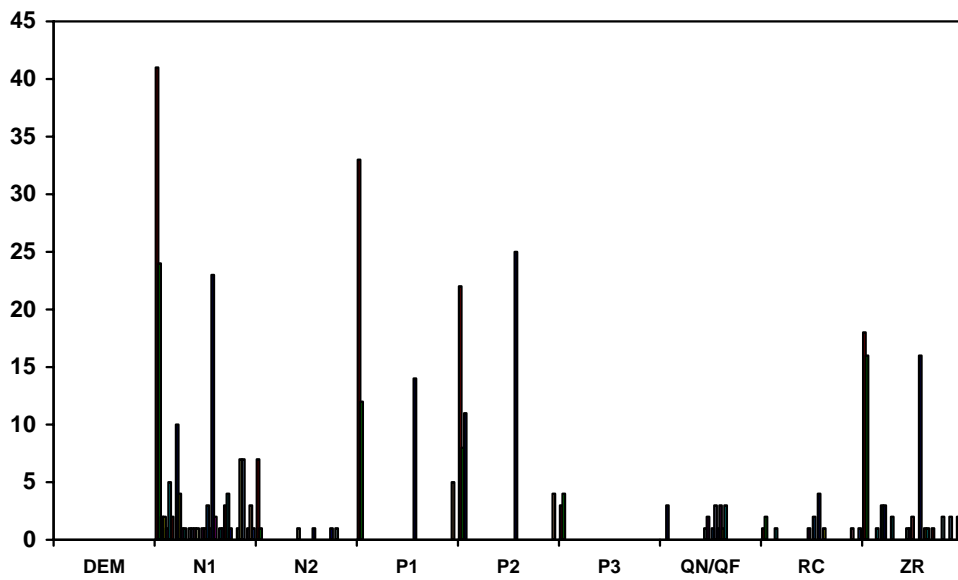


Figure 2: Participants 1-39 NP profile of Mtext 2

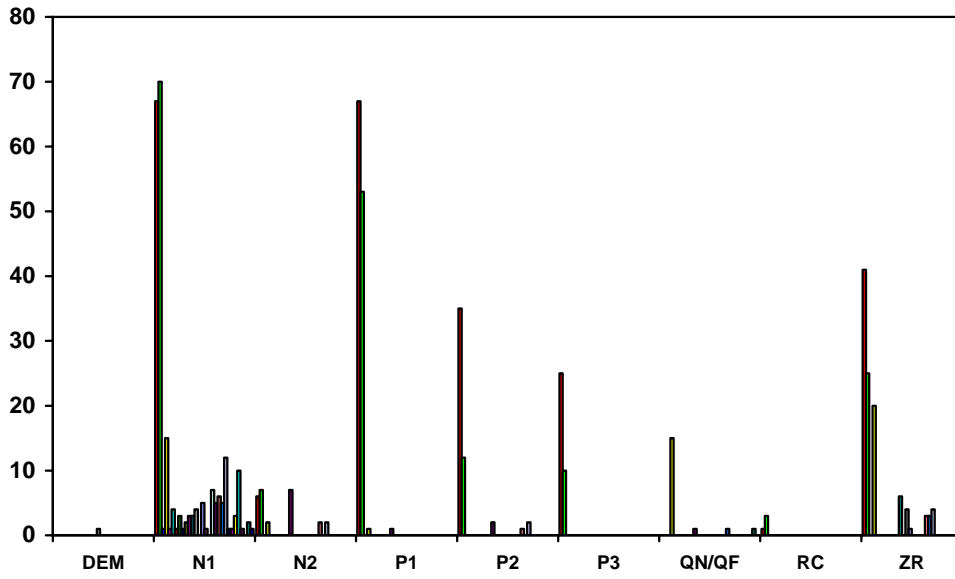


Figure 3: Participants 1–30 NP profile of Mtext 3

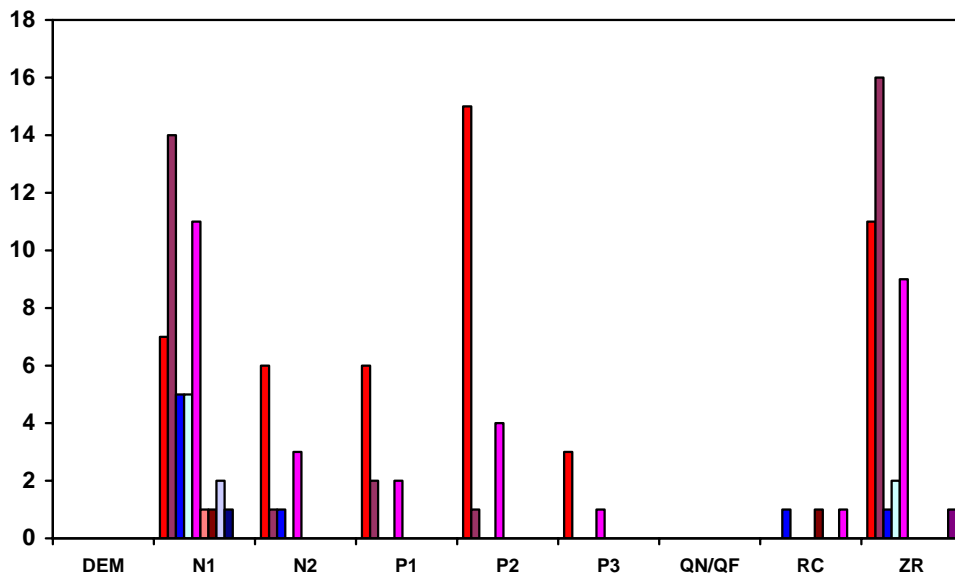


Figure 4: Participants 1–11 NP profile of Mtext 4

3.4 Selection criteria and transitivity evidence from discourse-based quantification

Two major questions are addressed in this section. First, does the inverse construction behave like a typical passive?⁸ Secondly, can the choice between a *ni*-verb and a *nong*-verb be predicted on the basis of discourse-level information, and, if so, on the basis of what parameters? These two problems will initially be explored following Givón's topic continuity methodology in §3.4.1 (Givón 1994; see Quick 1997, 2001, 2003 for previous work on this topic in Pendau). The first question is answered by using Givón's basic methodology. The second question is partially answered by applying Givón's methodology, but it is more fully answered in §3.4.2 by applying Dryer's (1994) supplementary method. Dryer's method is a modification of Givón's.

3.4.1 Applying Givón's method

Topic continuity analysis measures the frequency of occurrence of nominal arguments that are tracked in core argument positions.⁹ Givón states:

These methods are based on the assumption that more topical, (thematically important) referents tend to be both more **anaphorically accessible** ('continuous') and more **cataphorically persistent** ('recurrent'). Neither measure assesses topicality directly. Rather, they measure the referential continuity properties of referents, in two — opposite — textual directions. It is assumed then that the two measures should correlate with the two respective cognitive dimensions of topicality. (Givón 1994:10)

The quantification is carried out by examining each core argument of each transitive clause in a text and counting 'back' to find a match (thus measuring referential distance) and counting 'forward' to find a match (measuring topic persistence)

In this study, measurement of referential distance (RD) is made according to the conventions developed by Givón (1994). Once a core argument has been identified (whether it is expressed overtly or not), the analyst then looks backwards in the text until a previous reference to the same entity is found. One of two different values for RD is then ascribed: (1) distance of 1–3 (the most recent reference was made in any of the three immediately preceding clauses); (2) distance >3 (no reference was made in any of the three preceding clauses).¹⁰

⁸ This has been addressed many times in the literature on the Philippine languages. Shibatani (1988:96) gives a representative opinion (where his use of 'goal-topic' is equivalent to 'inverse' in Pendau):

In conclusion then, it is clear that while the patient nominals in the goal-topic construction and the passive in English and other languages are similar in regards to subject/topic role, these two constructions show far more significant differences. Past analyses that view the Philippine non-actor topic construction passive miss important overall characteristics of this construction that are not shared by the prototypical passives: namely, (i) it is not an agent defocusing mechanism in that it syntactically encodes both agent and patient, just as in active transitive clauses in other languages, and (ii) its functional load of coding a transitive event is as great as that of the actor-topic construction.

⁹ All core arguments required as actor or undergoer (coded here as A and P respectively) in a clause are counted whether they appear overtly or covertly.

¹⁰ Providing a detailed rationale for these particular measurements is beyond the scope of this paper. Givón (1994) provides extensive discussion of why these particular measures (and those used for topic

In this paper the designation *ni-* is used as shorthand to refer to all inverse voice constructions, and the designation *nong-* is used as shorthand to refer to all active voice constructions.

Tables 5a–d show the raw values for the referential distance for the four texts (Mtext 1–4) analysed. A look at these four tables shows that each text patterns in generally the same way.

Starting with the first column for Tables 5a–d, in 66%–74% of instances the P argument of the *ni-* (inverse) construction has a RD of 1–3, indicating high topicality. In only 25%–33% of cases does the P argument of the *ni-* construction display low topicality, i.e. $RD > 3$. The A argument of the *ni-* construction has a RD of 1–3 in 89–96% of cases and is thus even higher in topicality than the P argument in the same construction. In just 3%–17% of cases does the A argument in this construction have a $RD > 3$, i.e. low topicality.

Continuing to the third column, the P argument of the *nong-* (active) construction, the number of high-topicality (RD 1–3) cases ranges from 27% to 57%, whilst it occurs 42%–72% with a $RD > 3$, i.e. low topicality. The statistics for the P argument of the *nong-* construction thus vary widely between these four texts, and it must be concluded that its topicality varies from text to text. However, when we look at the A argument in the *nong-* construction we see that the A is highly topical with a RD of 1–3 80–89% of the time and an $RD > 3$ in only 10%–20% of cases

In sum, in the *ni-* (inverse) construction both P and A display high topicality, A even higher than P. In the *nong-* (active) construction, A is highly topical, whilst P varies in topicality but is always much lower than A.

Table 5a: Referential Distance values–Mtext 1

	<i>ni-</i>	<i>ni-</i>	<i>nong-</i>	<i>nong-</i>
RD	P	A	P	A
1–3	32 (74.42%)	26 (89.66%)	5 (27.78%)	53 (85.48%)
>3	11 (25.58%)	3 (10.34%)	13 (72.22%)	9 (14.52%)
Total	43 (100%)	29 (100%)	18 (100%)	62 (100%)

Table 5b: Referential distance values–Mtext 2

	<i>ni-</i>	<i>ni-</i>	<i>nong-</i>	<i>nong-</i>
RD	P	A	P	A
1–3	36 (66.66%)	54 (93.10%)	9 (33.33%)	33 (80.49%)
>3	18 (33.33%)	4 (6.90%)	18 (66.67%)	8 (19.51%)
Total	54 (100%)	58 (100%)	27 (100%)	41 (100%)

persistence) are appropriate. Givón actually uses three measurements, with the first category ($RD=1-3$) split into two subcategories ($RD=1$, and $RD=2-3$). The less precise measurement is deemed sufficient for current purposes. Due to programming constraints the referential distance measure is tabulated from the fourth clause of the text, and the topic persistence measure stops being tabulated at the tenth clause from the end.

Table 5c: Referential distance values–Mtext 3

	<i>ni-</i>	<i>ni-</i>	<i>nong-</i>	<i>nong-</i>
RD	P	A	P	A
1–3	70 (70.71%)	65 (82.28%)	24 (53.34%)	79 (89.77%)
>3	29 (29.29%)	14 (17.72%)	21 (46.76%)	9 (10.23%)
Total	99 (100%)	79 (100%)	45 (100%)	88 (100%)

Table 5d: Referential distance values–Mtext 4

	<i>ni-</i>	<i>ni-</i>	<i>nong-</i>	<i>nong-</i>
RD	P	A	P	A
1–3	19 (70.37%)	26 (96.30%)	8 (57.14%)	15 (88.24%)
>3	8 (29.63%)	1 (3.70%)	6 (42.86%)	2 (11.76%)
Total	27 (100%)	27 (100%)	14 (100%)	17 (100%)

Figures 5 and 6 compare RDs between the two constructions. Figure 5 shows a scatter-plot display for the RD for *ni-* and *nong-* verb constructions when the RD is 1–3 for either A or P in each of the four texts. The relative topicality of A seems to have little or no bearing on whether on not a *ni-* or a *nong-* form of the verb is used. Topicality of P, though, seems to be a much better predictor of voice. If P has high topicality, the *ni-* construction is chosen, but if it doesn't, the speaker chooses the *nong-* construction.

If the A in the *ni-* clause was actually an oblique of a passive, then the A should be expected to display much lower than in Figure 5. I would also expect it to be lower in topicality than the A in the *nong-* verb constructions (see §4). What appears dramatically here is that the A in the *ni-* verb is actually higher in topicality than the A in the *nong-* verb in three out of four of these texts, although this difference would not appear to be statistically significant.

Figure 5 illustrates that topicality of P as measured by RD in the *ni-* verb construction is much higher than the topicality of P in the *nong-* verb construction measured by the same criterion. This is what I would expect from a transitive construction which makes the P the pivot (or subject).

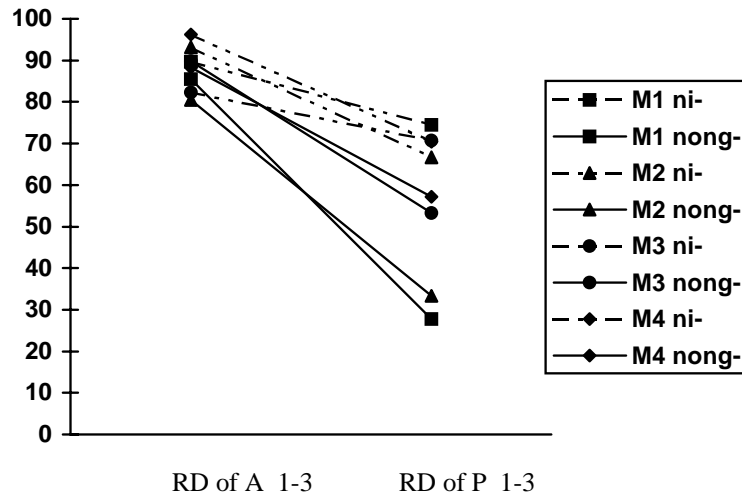


Figure 5: Percentages of *ni-* and *nong-* verb constructions with RD = 1-3 for P and A (Mttexts 1-4)

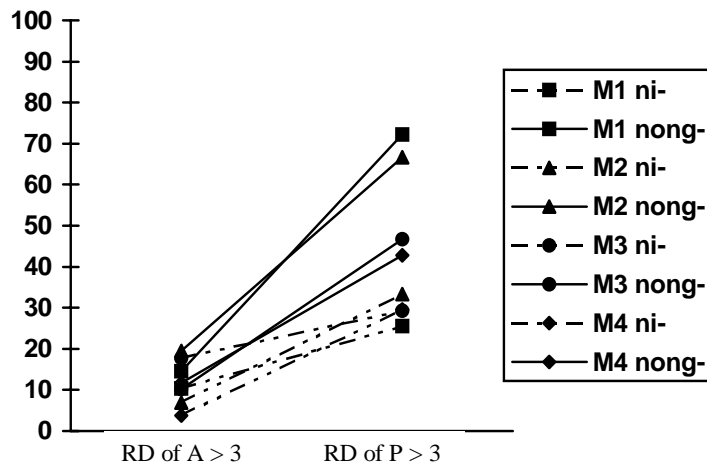


Figure 6: Percentages of *ni-* and *nong-* verb constructions with RD > 3 for P and A (Mttexts 1-4)

Figure 6 shows that for gaps greater than three clauses (i.e. discontinuous topics as against uninterrupted topics in Figure 5) non-topical A arguments are seldom encountered in either voice, but that if the P argument is not topical, then the *nong-* voice tends to be used more frequently than *ni-*. One of the reasons for high topicality of A in both voices is that A arguments generally occur in ‘runs’, i.e. the same participant is frequently a topical A for several continuous clauses. On the other hand, P arguments may or may not be continuous.

3.4.2 Applying the Dryer method

A methodology supplementary to Givón's basic method of quantifying topic continuity was proposed by Dryer (1994). I have also applied Dryer's methodology to the Pendau data since it gives even clearer reasons for voice selection than does Givón's on its own to finding an answer to the question whether the choice between a *ni*- verb and a *nong*- verb can be predicted on the basis of discourse level information, and, if so, on the basis of what parameters.

The Dryer method uses roughly the same counting procedures as outlined above, except that the topicality measures for A and P are compared so that the results obtained are measures of relative topicality between A and P rather than absolute topicality of A or P alone. There are three possible scores for each comparison:¹¹

- (a) A was mentioned in a more recent clause than P.
- (b) A and P were both mentioned most recently in the same clause.
- (c) P was mentioned in a more recent clause than A.

The statistics obtained from the application of the Dryer method can be viewed in two ways.¹² In what Dryer calls the 'vertical' analysis, the relative topicality measures of A versus P are computed as percentages within each clause type. In Dryer's 'horizontal' analysis, the starting point is relative topicality, and selection of clause type is computed as a percentage for each relative topicality measure.

Tables 6a–d show the Dryer vertical analysis of referential distance for each text. These tables also show a similarity in the statistics for all the four texts. These tables underscore and provide background information in understanding Tables 7a–e which give the Dryer horizontal analysis values. These statistics show a clustering effect that allows us to make a provisional statement regarding when a speaker tends to choose a *ni*- verb construction over the *nong*- verb construction. When the RD of A for a given clause is less than P, then the *nong*- verb construction is more frequently chosen (this is expanded further below).

Table 6a: Relative referential distance of A and P's (vertical analysis) Mtext 1

	<i>ni</i> -	<i>nong</i> -
RD of A lower	12 (40.00%)	28 (75.68%)
RD of A and P same	7 (23.33%)	3 (8.11%)
RD of P lower	11 (36.67%)	6 (16.22%)
Total	30 (100%)	37 (100%)

¹¹ Again, as for the application of Givón's methodology, measures of referential distance alone rather than for RD and topic persistence will give fairly clear results, so no measures for topic persistence will be recorded here.

¹² Dryer (1994) presents an additional means of analysing topic continuity data which he calls the 'vertical analysis' and the 'horizontal analysis'.

Table 6b: Relative referential distance of A and P's (vertical analysis) Mtext 2

	<i>ni-</i>	<i>nong-</i>
RD of A lower	15 (22.39%)	17 (40.48%)
RD of A and P same	25 (37.31%)	5 (11.90%)
RD of P lower	27 (40.30%)	20 (47.62)
Total	67 (100%)	42 (100%)

Table 6c: Relative referential distance of A and P's (vertical analysis) Mtext 3

	<i>ni-</i>	<i>nong-</i>
RD of A lower	30 (26.32%)	51 (50.00%)
RD of A and P same	36 (31.58%)	19 (18.63%)
RD of P lower	48 (42.11%)	32 (31.37%)
Total	114 (100%)	102 (100%)

Table 6d: Relative of distance of A and P's (vertical analysis) Mtext 4

	<i>ni-</i>	<i>nong-</i>
RD of A lower	3 (9.38%)	4 (21.05%)
RD of A and P same	15 (46.88%)	6 (31.58%)
RD of P lower	14 (43.75%)	9 (47.37)
Total	32 (100%)	19 (100%)

Tables 7a–d show the Dryer horizontal analysis for each text and Table 7e gives combined totals for all of the texts. Table 7e shows quite clearly that when A is more topical than P, speakers tend to choose a *nong-* construction, but when P is equal in topicality, or greater in topicality than A, speakers tend to choose a *ni-* construction.

Table 7a: Relative referential distance of A and P's (horizontal analysis) Mtext 1

	<i>ni-</i>	<i>nong-</i>	Total
RD of A lower	12 (30.00%)	28 (70.00%)	40 (100%)
RD of A and P same	7 (70.00%)	3 (30.00%)	10 (100%)
RD of P lower	11 (64.71%)	6 (35.29%)	17 (100%)

Table 7b: Relative referential distance of A and P's (horizontal analysis) Mtext 2

	<i>ni-</i>	<i>nong-</i>	Total
RD of A lower	15 (46.88%)	17 (53.13%)	32 (100%)
RD of A and P same	25 (83.33%)	5 (16.67%)	30 (100%)
RD of P lower	27 (57.45%)	20 (42.55%)	47 (100%)

Table 7c: Relative Referential Distance of A and P's (horizontal analysis) Mtext 3

	<i>ni-</i>	<i>nong-</i>	Total
RD of A lower	30 (37.04%)	51 (62.96%)	81 (100%)
RD of A and P same	36 (65.45%)	19 (18.63%)	55 (100%)
RD of P lower	48 (60.00%)	32 (40.00%)	80 (100%)

Table 7d: Relative referential distance of A and P's (horizontal analysis) Mtext 4

	<i>ni-</i>	<i>nong-</i>	Total
RD of A lower	3 (42.86%)	4 (57.14%)	7 (100%)
RD of A and P same	15 (71.43%)	6 (28.57%)	21 (100%)
RD of P lower	14 (60.87%)	9 (39.13%)	23 (100%)

Table 7e: Relative referential distance of A and P's (horizontal analysis)
All texts combined

	<i>ni-</i>	<i>nong-</i>	Total
RD of A lower	60 (38%)	100 (62%)	160 (100%)
RD of A and P same	83 (72%)	33 (28%)	116 (100%)
RD of P lower	100 (60%)	67 (40%)	167 (100%)

Figure 7 summarises the statistics from Tables 7a–d in one graph. These data show that there are clear differences in tendency between the choices of a *nong-* clause and a *ni-* clause. A choice between a *nong-* and a *ni-* clause is based more often than not on the degree of topic continuity. The rule of thumb in choosing between a *ni-* and a *nong-* verb construction can be stated as follows.

- If the P argument is more continuous (RD of P<A) or just as continuous as the A argument (RD of P=A), then the *ni-* verb construction will more often be chosen
- If the A argument is more continuous than the P argument (RD of A<P) than the *nong-* verb construction will more often be chosen

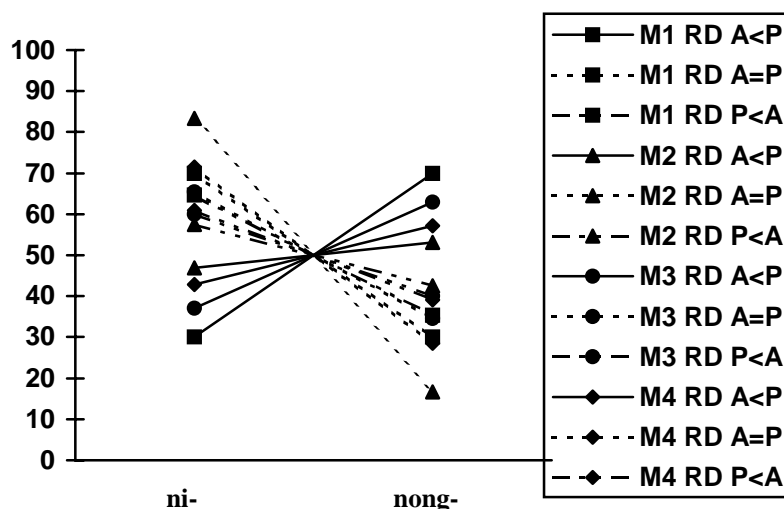


Figure 7: Frequency of *ni-* and *nong-* clauses according to whether the A is equal to P in referential distance (RD A=P), the A is less than P in referential distance (RD A<P), or the P is less than A in referential distance (RD P<A) in Mtexts 1–4

4 Does Pendau have a ‘passive’?

Givón (1994) sets out a number of criteria for determining which structures in a language might be labelled as passive, and how to distinguish passives from inverse constructions. The first criteria are related to topicality and are measured by RD and TP of A. According to Givón, passives are used when A arguments are not topical by either measure. We have already seen in §3 that the *ni-* construction in Pendau fits neither of these criteria, and is thus not a good candidate for a passive.

Givón (1983:23) also states that

... the text frequency of passives is much much lower than that of actives, somewhere between 5–20 percent of all main, affirmative, declarative clauses ... This by itself tags the passive as a *discontinuous* device in discourse, by virtue of its rarity.

In the texts analysed in §3, the *ni-* construction was actually used over twenty per cent more often than the *nong-* construction. Table 7e showed a grand total of 243 *ni-* clauses in all texts versus a grand total of 200 *nong-* clauses. By Givón’s frequency criterion then, the Pendau *ni-* construction makes a better inverse than it does a passive.

Givón (1994:12) gives another quantitative diagnostic for distinguishing passive from inverse voice. This diagnostic is the difference in frequency of omission of A arguments in each kind of structure.

Tables 8a–e display the total occurrences for the four possible occurrences or non-occurrences (labelled as overt and covert respectively) for *nong-* and *ni-* clause constructions. In Tables 8a–e it can be seen that the A argument is rarely omitted (or covert) in the *ni-* clause construction, and in fact the A argument is more often omitted in a *nong-* construction in some of the texts.¹³

¹³ These data also argue against the antipassive interpretation of the *nong-* construction, as an antipassive would be expected to omit the P argument rather than the A argument.

Table 8a: Frequency of overt/covert A/P arguments in Mtext 1

Mtext 1	AV	IV
1 Both A/P Overt	11	13
2 A Overt, (P Covert)	1	9
3 P Overt, (A Covert)	2	5
4 Both A/P Covert	0	1

Table 8b: Frequency of overt/covert A/P arguments in Mtext 2

Mtext 2	AV	IV
1 Both A/P Overt	25	35
2 A Overt, (P Covert)	11	20
3 P Overt, (A Covert)	3	9
4 Both A/P Covert	0	3

Table 8c: Frequency of overt/covert A/P arguments in Mtext 3

Mtext 3	AV	IV
1 Both A/P Overt	28	69
2 A Overt, (P Covert)	16	31
3 P Overt, (A Covert)	15	3
4 Both A/P Covert	5	0

Table 8d: Frequency of overt/covert A/P arguments in Mtext 4

Mtext4	AV	IV
1 Both A/P Overt	6	15
2 A Overt, (P Covert)	1	17
3 P Overt, (A Covert)	8	0
4 Both A/P Covert	2	1

It is clear then that the Pendau *ni-* construction is better treated as an inverse than as a passive by all of the criteria discussed by Givón.

The stative construction is a more likely candidate for passive (since it clearly fits all Givón's criteria for a passive outlined above). Tables 9a–d demonstrate that the 'agent' or

'effector' (E) that brought about the state described in a stative construction¹⁴ is a more likely oblique candidate that fits the profile for frequency of occurrence of the agent of a passive voice construction than the A of the inverse construction. The E argument only occurs 5 times in a main clause while the Su occurs 62 times in all of these texts. Su refers to the stative verb subject which is an undergoer subject.

Table 9a: RD of clause arguments in Mtext 1

Mtext 1	Su	E
1-3	14	0
>3	0	0

Table 9b: RD of stative clause arguments in Mtext 2

Mtext 2	Su	E
1-3	14	2
>3	4	0

Table 9c: RD of stative clause arguments in Mtext 3

Mtext 3	Su	E
1-3	24	3
>3	4	0

Table 9d: RD of stative clause arguments in Mtext 4

Mtext 4	Su	E
1-3	2	0
>3	0	0

Table 9e shows that the relative frequency of stative clauses compared with clauses of other types also fits Givón's criteria for a passive very well. Out of 746 clauses in these four texts only 62 clauses (or 8.3%) were stative clauses.

¹⁴ Stative verbs with an 'effector' adjunct may be a 'middle voice' which differs from stative constructions in which no 'effector' appears. However the statistics here refer to all main clauses which have a stative verb in them.

Table 9e: Total number of stative clauses in Mtexts 1–4

	# of stative clauses	# of total clause types
Mtext1	6 (6.4%)	94
Mtext2	24 (10.9%)	221
Mtext3	30 (8.5%)	354
Mtext4	2 (2.6%)	77
Total Clauses	62 (8.3%)	746

If we were to look for a discourse functional equivalent for a passive in Pendau, a much better candidate for such a structure would be the stative construction.

5 Word order variation in Pendau

This section presents the findings of statistics on word-order variation and attempts to find a motivation for selecting between the SV/SVO and VS/VOS word orders which were discussed in Quick (2003, Ch. 12). The subject can appear either before or after the verb in both intransitive and transitive clauses. The majority of clauses in narrative texts appear with the subject preceding the verb. However, there are a significant number of clauses with verbs from all verb classes in which S follows the verb.

Table 10 displays raw basic word-order scores for Mtext 3 which has a total of 354 clauses. Only clauses that had an overt subject were counted, and the 75 clauses with no overt subject are not included in this tabulation. (These figures do not include clauses that had covert objects but overt subjects.) The figures show that both word orders are quite frequent, but that SV/SVO is more frequent than VS/VSO.

Table 10: Frequency of word order types SV/SVO and VS/VOS

SV/SVO	VS/VOS	Total Clause Population
178 (64%)	101 (36%)	279

Table 11 shows the values for referential distance of the subject (S) in each of the word orders. These figures show that the subject has about the same degree of topic continuity in either word order; that is, it has occurred within the previous three clauses 83% of the time whatever the word order.

Table 11: Comparison of the referential distance of the S between SV/SVO and VS/VOS word orders: Mtext 3

RD	SV/SVO	VS/VOS
1–3	148 (83%)	84 (83%)
>3	30 (17%)	17 (17%)
Total	178 (100%)	101 (100%)

Since it is often mentioned in the literature on topic continuity that pronouns are more topical than full NPs, I then specified a further distinction by quantifying the subjects that occurred as full noun phrases (N1) and those occurring as pronouns (P1). This was done for both word orders and the results are tabulated in Table 12a for VS/VOS and Table 12b for SV/SVO.

Table 12a: Comparison of the S as N1 and P1 in VS/VOS word order

RD VS/VOS	N1	P1
1–3	57 (77%)	27 (100%)
>3	17 (23%)	0 (0%)
Total	74 (100%)	27 (100%)

Table 12b: Comparison of the S as N1 and P1 in SV/SVO word order

RD SV/SVO	N1	P1
1–3	72 (72%)	76 (97.5%)
>3	28 (28%)	2 (2.5%)
Total	100 (100%)	78 (100%)

Rau (1997:382) in her discussion of word-order variation in Atayal (Austronesian, Taiwan) states that:

The results of VARBRUL runs indicate that topicworthiness is the only factor that has any significant effect on word order variation. Proper nouns strongly favor the SV order, followed by common nouns, while personal pronouns strongly disfavor the SV order. Topicworthiness corresponds well with Givón's topic continuity. In other words, VS order is associated with topic continuity while SV order is associated with topic discontinuity.

However, in Pendau, both word orders reflect similar statistics. This is contrary to the assumed expectation that the postverbal subject would reflect greater topic continuity. Not surprisingly nearly all of the absolute case pronouns (P1) were referential within the last three clauses. However, what is surprising is that 77% of 74 tokens for the full noun subjects in VS/VOS orders (N1) had a RD of less than four.

Table 12c shows that full noun phrase subjects occur more commonly than pronominal subjects in the VS/VOS word order, while in the SV/SVO word order there are roughly equal numbers of full noun phrase and pronominal subjects.

Table 12c: Comparison of total occurrences of N1 and P1 in SV/SVO and VS/VOS word orders

	SV/SVO	VS/VOS
N1 (etc.)	89 (52%)	74 (73%)
P1	81 (48%)	27 (27%)
Total	170 (100%)	101 (100%)

When the grammatical subject is in the postverbal position it is more likely to occur as a full noun phrase than as pronoun. This distinction demonstrates that the full noun phrase is favored in about a 3:1 ratio when the subject is in the postverbal position. Since the referential distance of a subject is approximately the same when comparing preverbal and postverbal subject clause constructions, it must be (provisionally) concluded that the difference between the use of VS/VOS and SV/SVO word orders has nothing to do with topic continuity. This indicates that there must be an independent factor apart from topic continuity which results in the difference between word orders. Further research is needed to determine what might cause word-order variations. Obvious possibilities include the placement of NPs in different positions for the purposes of emphasis, or the positioning of heavy NPs in places where their use is less awkward.

References

- Andrews, Avery, 1985, The major functions of the noun phrase. In Timothy Shopen, ed. *Language typology and syntactic description, 1: Clause structure*, 62–154. Cambridge: Cambridge University Press.
- Comrie, Bernard, 1989, *Language universals and linguistic typology* (second edition). Chicago: University of Chicago Press.
- Dixon, R.M.W., 1979, Ergativity. In *Language* 55:59–138.
- 1994, *Ergativity*. Cambridge: Cambridge University Press.
- Dryer, Matthew S., 1994, The discourse function of the Kutenai inverse. In Givón, ed. 1994:65–99. Amsterdam: John Benjamins.
- Givón, Talmy, 1983, Introduction. In Givón, T., ed. *Topic continuity in discourse: a quantitative cross-language study*, 1–41. Amsterdam: John Benjamins.
- 1990, *Syntax: a functional-typological introduction*, vol 2. Amsterdam: John Benjamins.
- 1994, The pragmatics of de-transitive voice. In Givón, ed. 1994:3–44.
- Givón, Talmy, ed., 1994, *Voice and inversion*. Amsterdam: John Benjamins.
- Payne, Thomas, 1997, *Describing morphosyntax: a guide for field linguists*. Cambridge: Cambridge University Press.
- Quick, Philip, 1996, Multilinear discourse analysis software demonstration. In H. Andrew Black, Alan Buseman, David Payne and Gary F. Simons, eds *Proceedings of the 1996 general CARLA conference, November 14–15, 1996*, 291–309. Waxhaw, NC/Dallas: JAARS and Summer Institute of Linguistics.
- 1997, Active and inverse voice selection criteria in Pendau, a Western Austronesian language. In Cecilia Odé and Wim Stokhof, ed. *Proceedings of the Seventh International Conference on Austronesian Linguistics*, 461–481. Amsterdam: Editions Rodopi B.V.
- 1999, Sekilas tata bahasa Pendau. In Bambang Kaswanti Purwo, ed. *Panorama Bahasa Nusantara*, 1–44. Jakarta: Referensia.
- 2001, A sketch of the primary transitive verbs in Pendau. In Fay Wouk and Malcolm Ross, eds *The historical and typological development of western Austronesian voice systems*, 101–122. Canberra: Pacific Linguistics.

- 2003, A grammar of the Pendau language. PhD dissertation. The Australian National University.
- Rau, Der-hwa V, 1997, A discourse analysis of Atayal word order variation. In *Papers for the Eighth International Conference on Austronesian Linguistics*, 381–399. Taipei: Unpublished Proceedings.
- Shibatani, Masayoshi, 1988, Voice in Philippine languages. In Masayoshi Shibatani, ed. *Passive and voice*, 85–142. Amsterdam: John Benjamins.

11 *Semantic analysis of the Moronene verbal prefix moN-*

SUREE ANDERSEN AND T. DAVID ANDERSEN

1 Introduction

1.1 Moronene clause types

Moronene, a Bungku-Tolaki language of Southeast Sulawesi, has a number of different clause types which fall into the category of dyadic semantically transitive clauses.¹ These clauses have an agent which is manifested as transitive actor and an undergoer which is manifested as object.² This paper will focus on the semantic contrast between clauses in which the undergoer/object is indexed on the verb and those in which it is not.

Before we turn to the matter at hand, it will be helpful to give a brief summary of some of the basic intransitive and transitive clause types in Moronene. The main factor distinguishing the types is the case of the person markers which are cliticised to the verb or auxiliary. These are either nominative (NOM), indexing either the actor of a transitive

¹ We want to express our gratitude to David Mead, who encouraged us to present this paper at the Ninth International Conference on Austronesian Linguistics, and gave us valuable feedback to improve it. Our thanks also go to many Moronene speakers who provided the data on which our analysis is based, in particular to the storytellers Wede, Ndasi, and Narson, who provided many of our texts, to Tipu and Ndoke for their recorded discussion of rice rituals, for Agus Poli and Sultan Wudhawie for hundreds and hundreds of sentences for the Moronene lexicon, and to Estelita for natural Moronene conversations.

The data upon which this paper is based was gathered from 1991 to 2001 mainly from the villages of Taubonto (Rarowatu subdistrict) and Pusu'ea (Poleang Timur subdistrict) in the Buton district of Southeast Sulawesi. This research was conducted under cooperative agreements between the Summer Institute of Linguistics and the Indonesian Government, specifically the Department of Education, the Indonesian Institute of Sciences (LIPI) and the Village Community Development Office (PMD). We are grateful to all these organisations for their facilitation of our work among the Moronene people.

² In this paper the following terminology will be used: actor, the surface constituent which refers to the initiator of the predication in a transitive clause; object, the surface constituent which refers to the undergoer of a transitive clause; subject, the surface constituent constituting the sole core argument in an intransitive clause (when referring to the terminology of others, subject may also refer to the transitive actor); agent, the underlying initiator of an event; undergoer, the underlying entity affected by an event (cf. Matti 1994:84, n.10).

clause or the subject of an intransitive clause, absolutive (ABS), indexing the object of a transitive clause or (like NOM) the subject of an intransitive clauses, or possessive (POS), indexing the actor of a transitive clause with undergoer focus. It is beyond the scope of this paper to discuss the features of all these clause types. The important question of the presence and order of noun phrases corresponding to actor and object is ignored. The list of clause types is not exhaustive. Each clause type will be given an acronym in parentheses, and all the subsequent examples will be marked with the acronyms of all the clause types found in them. When there is more than one clause the acronyms are separated by semicolons. This may aid those who wish to compare similar clause types found in the data.

1.1.1 Intransitive clause types

The subject of intransitive clauses may be marked either with an absolutive clitic or a nominative clitic. Among the factors influencing the choice of case are the presence or absence of conjunctions and other grammatical functors in the clause. This will not be discussed here.³

1. Unmarked verb (V)⁴

- (1) *Is te-tii ma'e.*
 start RES-descend uncle
 'First uncle arrived.' (laku113)

³ Abbreviations of glosses: ABS absolutive; AF action focus; AF/NF action focus/non-finite; ANTIPASS antipassive; APPL applicative; ARF actor focus; ART article; BEN benefactive; CAUS causative; CERT certainty; COMPL complementiser; CMS commiserative; CTR contra-expectation; CTRS contrastive emphasis; e exclusive; FUT future; h honorific; i inclusive; IMPF imperfective; IMPV imperative; INT intransitive; LG ligature; LOC locative applicative; NEG negative; NOM nominative; NR nominaliser; p plural; PASS passive; PI person indicator; PL plural; POS possessive; PRF perfective; PROG progressive; Q question marker; RED reduplication; REL relativiser; s singular; STV stative; SUP superlative; TR transitive; UF undergoer focus; V verb; VRt transitive verbaliser; 1, 2, 3 first, second, third person.

Abbreviations of texts: **coco1** (Moweweu epu. [Making coconut oil] Pusuea 1991), **col** (Colisi. [Pinkie] By Ndasi. Taubonto 1992), **dahu** (Tulurano Dahu ronga Koe. [Story of the Dog and the Egret] By Ndasi. Taubonto 1992), **dic** (Moronene lexicon. Compiled by T. David Andersen, Agus Poli and Sultan Wudhawie), **diu** (Petampu'uno Ica Diu. [The Origin of the Dugong] By Wede. Taubonto 1994), **kol** (Kolopua ronga Ndoke. [The Turtle and the Macaque] By Rambe. In Muthalib et al. 1991), **laku** (Laku and Manu. [Conversation about civet and chickens] By Sudin, Sese, Maa, and Rondo. Pusuea 1991), **lang** (Langko'oheo. By Wede. Taubonto 1994), **Maegani** (Tinano I Maegani, Miano Cantete. [Maegani's Mother, the Person in the Latrine] By Wede. Taubonto), **ndoke** (Ndoke hela Buea. [The Macaque and the Crocodile] By Narson. Kendari 1994), **perc** (Percakapan. [Conversations] By Estelita. Taubonto), **roo** (Roo Mpae. [Rice Medicine] By Tipu and Ndoke. Taubonto), **sio** (Wuu Sio Ropa. [Nine Fathom Hair] By Wede. Taubonto 1992?), **wola** (Tulurano Wola ronga Kore. [Story of the Rat and the Frog] By Ndasi. Taubonto 1992.)

⁴ Unmarked here means unmarked by a person marker. There is a wide variety of derivational intransitive verb prefixes, but those are ignored here.

2. Verb + ABS (VAb)

- (2) *Nta lako-kami hai peta.*
 FUT go-1peABS at ricefield
 ‘We’re going to go to the ricefield’ (perc)

2a. Auxiliary + ABS + verb (AuAbV)

- (3) *Nde’e i Tadi ari-o tii di-ceena?*
 indeed PI Tadi already-3sABS descend this-there
 ‘Has Tadi already come down here?’ (laku26)

3. NOM + verb (NV)⁵

- (4) *Ka-i leu korua i Arbaa.*
 then-3sNOM come below PI Arba
 ‘Then Arba came down.’ (laku120)

3a. NOM + auxiliary + verb (NAuV)

- (5) *Ka-u daa um-owu.*
 then-2sNOM be INT-slash
 ‘Then you slash.’ (roo3)

1.1.2 Passive clause types

Passive clauses involve the use of the undergoer focus prefix *ni-* (or infix *-in-*) on the verb. In other respects they are similar to intransitive clause types 2 and 2A. In rare cases, the agent may be present as an oblique, as in example (7).

4. *ni*-verb + ABS (ABS = undergoer) (nVAb)

- (6) *T[in]arima-a moico-si langa-ku.*
 [UF]accept-3sABS good-CTR brideprice-1sPOS
 ‘My brideprice was well accepted.’ (w:46)

4a. Auxiliary + ABS + *ni*-verb (ABS = undergoer) (AuAbnV)

- (7) *Daa-ko nta ni-rako hai polisi.*
 be-2sABS FUT UF-catch at police
 ‘You may be caught by the police.’ (7a)

1.1.3 Transitive clause types

One subtype of transitive clause is characterised by indexing the undergoer/object with absolutive clitics on the verb stem. There are several types:

⁵ This includes cases in which the third singular nominative clitic *-i* is assimilated by the conjunctions *ki* ‘if’ and *hi* ‘COMPLEMETISER’ (*ki+i = ki*; *hi + i = hi*)

5. Verb + ABS (ABS = undergoer) (tVAb)

(8) *Lako-mo onto-o koie lere.*
 go-IMPV see-3sABS that sedge.grass
 ‘Go look at that sedge-grass.’ (diu 127)

(9) *Sie-mo ka’asi po-pate-kami!*
 don’t-IMPV CMS CAUS-kill-1peABS
 ‘Please don’t kill us!’

5a. Auxiliary + Verb + ABS (ABS = undergoer) (AuVAb)

(10) *Yo laku ari kea’-o manu.*
 ART civet already bite-3sABS chicken
 ‘The civet bit the chickens.’ (laku11)

6. Auxiliary + ABS + verb + ABS (1st ABS = agent; 2nd ABS = undergoer) (AuAbVAb)

(11) *Ari-aku-mo wowa-hira.*
 finish-1sABS-PRF bring-3pABS
 ‘I already brought them.’ (laku20)

7. NOM + verb + ABS (NOM = agent; ABS = undergoer) (NVAb)

(12) *Mi-m-pasipole-e ile.*
 2pNOM-PL-undertake-3sABS tomorrow
 ‘You take care of it tomorrow.’ (laku75)

7a. NOM + auxiliary + verb + ABS (NOM = agent; ABS = undergoer) (NAuVAb)

(13) *Taba tina manu ku-da’a nta balu-’o.*
 only female chicken 1sNOM-NEG FUT sell-3sABS
 ‘But the hens I won’t sell.’ (laku49)

Another subtype of transitive clause is that with undergoer focus, where the agent is marked on the verb with a possessive suffix.

8. *ni*-verb + POS (POS = agent) (nVG)

(14) *Da-hoo pe’ico raro-m-punti da in-onto-ngku.*
 be-3sABS yon group-LG-banana REL UF-look-1sPOS
 ‘Over there there is a group of banana trees which I can see.’ (sio44)

A third subtype of transitive clause does not have the undergoer indexed on the verb. A transitive verb without such an absolutive clitic to mark the object occurs with the *moN*-verb prefix. In almost all cases, a *moN*- prefix and an absolutive clitic referring to an object cannot co-occur. There are two forms of the prefix, *moN*- and *poN*-. The former is a non-finite form (Mead (1998:173) calls it a participle), whereas the latter is finite. The *moN*- prefix will be glossed AF/NF ‘action focus/non-finite’, whereas the alternative *poN*- prefix will be glossed AF ‘action focus’ (implicitly finite). The reasons for this gloss will become clear by the end of the paper. Etymologically, the *poN*- form is more basic, and the *moN*- form arose from coalescence of the participle marker *-um-* with the *poN*- (Mead

1998:173). But in Moronene the *moN-* form is more frequent and the *-um-* infix no longer exists as a participle marker, so for convenience we will refer to verb forms with either prefix collectively as *moN-V* forms. The various clause types with *moN-V* forms are set out below.

9. Unmarked *moN-Verb* (mV)

- (15) *Mo-'ita puhu, mo-'ita in-isa.*
 AF/NF-ask corn AF/NF-ask UF-pound
 'He asked for corn, he asked for husked rice.' (laku153)

9a. Unmarked *poN-Verb* (pV)

- (16) *Po-wawa yo boru.*
 AF-carry ART umbrella
 'Take an umbrella.' (perc)

9b. Auxiliary + *moN-Verb* (AumV)

- (17) *Osie-mo mo-wada koie karambau-ku.*
 don't-IMPV AF/NF-pay that buffalo-1sPOS
 'Don't pay for that buffalo of mine.' (Maegani 224)

10. *moN-Verb* + ABS (ABS = agent) (mVAb)

- (18) *Nta mo-weweu-kita yo wala.*
 FUT AF/NF-make-1piABS ART fence
 'We're going to make a fence.' (perc)

10a. Auxiliary + ABS + *moN-Verb* (ABS = agent) (AuAbmV)

- (19) *Da-hira-po mo-hoho kawasa-no mokole-do.*
 be-3pABS-IMPF AF/NF-worship rich-3sPOS king-3pPOS
 'They still worship the glory of their king.' (dic:hoho)

10b. *poN-Verb* + ABS (ABS = agent) (pVAb)

- (20) *Po-lesa-ko-mo ta'i sapi.*
 AF-step.on-2sABS-PRF faeces cow
 'You have stepped on some cow manure.' (perc10)

11. NOM + *poN-Verb* (NOM = agent) (NpV)

- (21) *Ka-i po-ndo'u penda.*
 then-3sNOM AF-drink again
 'Then he drank again.' (Maegani 118)

11a. NOM + auxiliary + *moN-Verb* (NAumV)

- (22) *To-daa nta mo-weweu epu.*
 1pi-be FUT AF/NF-make coconut.oil
 'We're going to make coconut oil.' (coco1)

The *N* in the prefixes symbolises a nasal ligature which appears before roots starting with voiceless stops. For example:

<i>mom-pa'u</i>	'to pour'
<i>mom-pepe</i>	'to pound'
<i>mon-tabe</i>	'to forbid'
<i>mon-totai</i>	'to slice up'
<i>mong-kela</i>	'to tilt'
<i>mong-kerusi</i>	'to sweep'

Since the nasal ligature only appears before three consonants, in most cases the prefix appears as *mo-* or *po-*. It is important to distinguish the *moN-* and *poN-* prefix from other homophonous *mo-* and *po-* prefixes which can co-occur with the object clitic. These latter prefixes occur with derived causative verbs and verbs of ability and in a small number of other verbs where its meaning is unclear, being a frozen form. Examples:

<i>mo-ko-mosele-'o</i>	'to cause it to be wet'
<i>mo-ko-sangke-'o</i>	'to be able to harvest it'
<i>mo-turusi-o</i>	'to let him'
<i>mo-dea-ho</i>	'to hear it'
<i>mo-wee-ho</i>	'to give him'

The transitive *moN-* and *poN-* prefixes contrast with other prefixes marking intransitive verbs, such as *me-*, *ma-*, *um-* and *ko-*. For example:

<i>me-baho</i>	'bathe oneself'	<i>mo-baho</i>	'bathe someone'
<i>ma-daga</i>	'be on guard'	<i>mo-daga</i>	'guard something'
<i>um-ahi</i>	'draw water'	<i>mo-'ahii</i>	'fill'
<i>ko-'onto</i>	'able to see'	<i>mo-'onto</i>	'see'

Objects marked by absolutive clitics tend to be important participants in the story or important props, definite or specific noun phrase objects. For example:

- (23) *Ka-i leu susu-o ana-no.*
 then-3sNOM come milk-3sABS child-3sPOS
 'Then she came and breast-fed her child.' (diu 37) (NV;tVAb)
- (24) *Ka-i wiso-'o koie yo kuli ng-kuliri-no, yo*
 then-3sNOM enter-3sABS that ART skin LG-lory-3sPOS ART
babu-no koie yo tina sangia.
 clothes-3sPOS that ART woman fairy
 'Then he put in that lory skin, the fairy's clothes.' (lang26) (NVAb)

In contrast, objects occurring with *moN-* verb prefix tend to be ordinary props, general or non-specific noun-phrase objects. For example:

- (25) *Da-ho-mo sisimbu mo-'u'ungke da nta t[in]ora-ako-mo.*
 be-3sABS-PRF wander AF/NF-search REL FUT [PAS]live-APPL-PRF
 'He was wandering around looking for something to live off.' (wola4)
 (AuAbV;mV)

- (26) *Nilako-no-mo mong-kuli-si me'asa punti*
 immediately-3sPOS-PRF AF/NF-skin-LOC one banana
 'He immediately peeled one banana.' (kol33) (mV)

Yet there are quite a number of significant examples of other possibilities such as:

A V-ABS form with an object noun phrase which is semantically indefinite:

- (27) *Ka-i te-leu awa-a eo wai.*
 then-3sNOM INT-come find-3sABS vine rattan
 'Then he arrived and found a rattan vine.' (lang49) (NV;tVAb)
- (28) *Ka-i awa-a olumpu-'ute.*
 then-3sNOM find-3sABS hut-little
 'Then he found a little hut.' (sio52)⁶ (NVAb)

A *moN-V* form with an object noun phrase which is definite:

- (29) *Da-hoo kolopua tangasa mo-aha pali-no.*
 be-3sABS turtle PROG AF/NF-sharpen axe-3sPOS
 'The turtle was sharpening his axe.' (kol20) (AuAbmV)

Even later in this text, this 'axe' has no possessive clitic, yet it is definite:

- (30) *Mo-hapa-a ka-u daa mo-'aha pali?*
 AF/NF-what-3sABS then-2sNOM be AF/NF-sharpen axe
 'Why are you sharpening the axe?' (kol21) (NAumV)

In other examples, the object of a *moN-V* form is marked as definite by the use of demonstratives:

- (31) *ka-i po-'oli koie ica mokokondo'u e'e*
 then-3sNOM AF-buy that fish thirsty water
 'Then he bought that "thirst fish".' (diu9) (NpV)
- (32) *Hai hapa ari-a-u mo-'ala co'o ana n-tina-'ate*
 at what finish-LOC-2sPOS AF/NF-take you child LG-woman-little
koie yo arataa?
 that ART treasure
 'Little girl, where did you get that treasure?' (diu112) (mV)

From this evidence one can conclude that Moronene verbs can have object agreement marked not only for definite objects, but also for indefinite specific objects. But how can one explain *moN-V* forms with a definite object?

1.2 Similar verb forms in Sulawesi languages

Many Sulawesi languages have forms which seem to be functionally similar to and often cognate with the *moN-V* and V-ABS forms in Moronene. The verb form equivalent to the *moN-V* form generally has a transitive prefix added to the transitive verb stem. In

⁶ This example is discussed in Mead (1998:179).

reference to western Austronesian languages Kroeger (1996:33) says ‘*PoN-* appears in a variety of languages as a marker of ‘transitivity’ and has sometimes been analyzed as an antipassive marker’.

The *moN-V* form in Bungku-Tolaki languages, including Moronene, is regarded as an antipassive form by Mead (1998:172–183). In Mori Bawah, according to the analysis of Esser (1933), active forms (e.g. V-ABS) are used so long as the patient is a specific (whole) entity, whereas non-specific patients require the antipassive (Mead 1998:174). In Tolaki, however, the antipassive can be used with specific patients when they are first mentioned in the discourse. Active forms are used with definite patients (Mead 1998:176, citing data from Scott Youngman). In Kulisusu, also, the antipassive is used with indefinite patients (Mead 1998:180).

The Balantak language of Central Sulawesi has an irrealis actor focus prefix *mVng-*. Busenitz (1994:1) says ‘the primary function of the AF prefix is to mark the actor as the topic of the clause’. This prefix, as with Moronene *moN-*, cannot co-occur with object agreement.

In the Da’a language of Central Sulawesi, the realis actor focus prefixes are *naN-/neN-/noN-* and the irrealis actor focus prefixes are *maN-/meN-/moN-* (Barr 1988a:19, 1988b:122). Actor focus is used when there is a change of actor or when props or minor participants are introduced as objects (Barr 1988b:84, 91). But this pattern is overruled when the object is a thematic prop, in which case goal focus is used (Barr 1988b:97).

In the Uma language of Central Sulawesi, the actor focus is marked by the prefix *mpo-*. It is used primarily ‘when the actor or the activity of the actor is the topic of the sentence or when the actor is being contrasted or highlighted’ (Martens 1988:173). In contrast to Moronene, it is possible for the object to be marked with a clitic pronoun on such actor focus verbs. For example:

- (33) *Uma-a dota m-po-'ubu'-ko*
 not-I want ARF-TR-carry-you
 ‘I don’t want to carry you.’ (Martens 1988:173)

In the Konjo language of South Sulawesi, there are two verbal prefixes with the form *ang-*. The first is attached to the infinitive and affects the initial consonant of the verb stem. It ‘indicates a semitransitive verb with actor focus; there may or may not be an object, but it cannot be a definite object, i.e. one that is clearly specified’ (Friberg 1991:108). If an absolutive clitic occurs on the verb, it refers to the actor. The second *ang-* prefix attaches to the basic root or derived stem and does not affect the initial consonant of the stem. It ‘indicates a transitive verb with goal focus; i.e., the goal (or object) is fully specified’ (Friberg 1991:108). The absolutive clitic on the verb refers to the object.

Among Sulawesi languages some terms used for verb forms equivalent to V-ABS in Moronene and contrasting with the subject/actor focus forms include ‘object focus’ (Friberg 1996:144) or ‘goal focus’ (Barr 1988b:78) or ‘undergoer focus’ (Himmelmann 1996:118).

A number of descriptions of Sulawesi languages give examples of constructions described as actor focus or subject focus or antipassive occurring with definite objects, similar to the Moronene examples cited above. The usual case is for undergoers of such verbs to be indefinite/low in topicality, but authors who describe these languages acknowledge that exceptionally sometimes the patient can be definite. What explanations are put forward for these non-prototypical constructions?

In Konjo, Friberg uses the explanation that the formally definite noun phrase with the subject focus verb form is indefinite in its meaning, i.e. one of many. Note the following Konjo example:

- (34) *L-angng-inrang-a berangta.*
 FUT-VRt-borrow-1ABS knife-2hPOS
 ‘I want to borrow (one of) your knives.’ (Friberg 1996:144)

In the Mamasa language of South Sulawesi, there is an antipassive construction which uses an actor focus prefix. Here is an extract of Matti’s (1994:74–75) discussion of this construction:

Although the antipassive is most often used with indefinite objects, occasionally it occurs with definite objects.

- (35) *Ung-kolo(ng)-mi ade’ adi-[n]na.*
 ARF-carry.on.ones.back-PRF/3 it.is.said younger.sibling-3
 ‘She carried her younger sibling on her back.’

Although the object *adinna* ‘her younger sibling’ is definite due to the possessive suffix *-na*, its definiteness is not being stressed; rather this is the only way that the younger sibling could be referred to due to the nature of the noun *adi*. Here the object is being reintroduced into the story, but once again it is not in ‘focus’.

Here are some similar examples of antipassive constructions with definite objects found in the Mamuju language of South Sulawesi:

- (36) *Ma(l)-lalle-a’ arloji-ku’.*
 TR-look.for-1ABS watch-1POS
 ‘I’m looking for my watch.’ (Strömme 1994:107)
- (37) *Na ma-[n]appasa-a’ poo-poo-na Cia.*
 FUT TR-wash-1ABS RED-nappy-3POS Cia
 ‘I’m going to wash Cia’s nappies.’ (Strömme 1994:108)

In relation to the last example, Strömme (1994:108) says, the object is also definite (Cia’s nappies). So why is the antipassive construction with absolutive actor still used? The answer is probably that ‘Cia’s nappies’ here is just used as a cover for or generalisation for all the other things the person is going to wash.

In Tolaki, Mead (1998:177) cites the following example of an antipassive being used with an object which is specific and known to both the speaker and his audience (e.g. definite):

- (38) *Tewali-’i-ki ku-onggo [m]o-saru la’usa-miu?*
 be.possible-3s-CERT 1s-want PART:ANTIPASS-borrow ladder-2p
 ‘May I certainly borrow a ladder of yours?’

Mead suggests that this construction is motivated by considerations of politeness: referring to the item to be borrowed as if it were indefinite is more polite than using a V-ABS form.

The aim of this paper is to examine Moronene clauses with *moN-V* forms which are similar to those cited above for other Sulawesi languages, and to posit a semantic analysis which accounts for the usage. In contrast to descriptions which mention only such ‘definite

object' constructions as exceptions, we aim to undertake a more systematic investigation, and to use a cognitive approach to present a more integrated account.

2 Prototypical meaning

Our aim in this paper is to clarify the meanings and functions of *moN-V* verb forms in Moronene, especially vis-a-vis the transitive V-ABS construction. The *moN-V* forms will be analysed in terms of a prototypical meaning and various extended meanings. We will examine examples showing various kinds of meaning differences correlating with the above-mentioned grammatical marking distinction. We will survey these meaning differences to trace other grammaticalised semantic patterns. And to obtain a plausible interpretation, we shall examine a cognitive grammar approach to verb semantics.

The first step is to examine various aspects of the prototypical meaning of *moN-* verbs in Moronene.

2.1 Act on nonspecific object versus act on specific object

The basic pattern of contrast between the two verb forms is that the *moN-V* form designates an act on a nonspecific object whereas the ABS-V form designates an act on a specific object. The large majority of cases can be explained using this pattern.

Some examples which contrast the *moN-V* and V-ABS forms:

mo-rako 'catch' versus *rako-'o* 'catch it/him/her'

- (39) *saba nta mo-rako pera-no kadadi da in-onto-no.*
 appear FUT AF/NF-catch all-3sPOS animal REL PAS-see-3sPOS
 'It appeared and was going to catch any animal it saw.' (col 78) (V;mV;nVG)

- (40) *Ka-ndo men-teka-'u'ungke nde'e miano me'alu rako-'ira*
 then-3pNOM PL-?-search indeed person many catch-3pABS
ko'ira miano m-po-nonako.
 those person LG-NR-steal
 'Then the crowd of people searched and searched and caught those thieves.'
 (col 48) (NV;tVAb)

mong-kea 'bite' versus *kea-'o* 'bite it/him/her'

- (41) *Da-hoo nta mong-kea miano.*
 be-3sABS FUT AF/NF-bite person
 'It will bite someone.' (col85) (AuAbmV)
- (42) *Iso tealo kea-'o yo wontu.*
 start pass bite-3sABS ART mosquito
 'Just then a mosquito passed by and bit him.' (lang65) (V;tVAb)

Some other examples illustrating the use of *moN-V* forms with non-specific objects:

- (43) *Kana'umpe cami hi-co me-lako pong-kaa wua ng-keu.*
 how we.ex COMPL-1peNOM PL-go AF-eat fruit LG-wood
 'How about if we go and eat fruit.' (ndoke11) (NV;pV)
- (44) *Da-haku mo-'awa doi hai ai-ai boku-'u.*
 be-1sABS AF/NF-find money at RED-gap book-2sPOS
 'I found money in between the leaves of your book.' (dic:ai) (AuAbmV)

2.2 Antipassive voice

One way of explaining this prototypical meaning is to call it antipassive voice. Mead (1998:172–183) claims that the *moN-V* forms in Bungku-Tolaki languages, including Moronene, are antipassive forms.

Some scholars have a narrow definition of antipassive based on grammatical features. It can also be defined more broadly on the basis of discourse function. According to a narrower grammatical definition, antipassive clauses display characteristics of intransitive clauses. They may be objectless, or the object may be verb-incorporated, or it may be marked as oblique. In ergative languages, the agent of this construction carries absolutive marking, and the verb is coded as intransitive. Foley and van Valin (1984:110–111) describe an antipassive construction as one that allows the actor to function as pivot. ('Pivot' is defined as a noun phrase crucially involved in the grammatical construction.)

An example of a broader definition is that of Mead (1998:155), who defines antipassives as 'constructions in which the patient has lowered referentiality or topicality.' According to Givón (1990:566–567), in the prototypical antipassive clause, the agent has high topicality (continuous from the preceding text and persistent in the subsequent text) and the topicality of the patient is much lower than in the active voice, and the patient tends to be thematically unimportant.

In the following sections we will evaluate the extent to which the *moN-V* form matches the narrower or broader definition of antipassive.

2.3 Absolutive actor construction

There is a particular construction in Moronene which has some similarities to an antipassive construction as defined grammatically. It involves the use of a *moN-V* form. We shall call it the absolutive actor construction. The actor is marked with absolutive rather than the more common nominative marker. Elsewhere in the language, the absolutive subject marker is normally used in an intransitive clause. For example:

- (45) *Nta lako-'aku-mo.*
 FUT go-1sABS-PRF
 'I'm going to go.' (VAb)

The types of tense-aspect found with the absolutive actor construction are only the definite future tense (*nta + mo-V*) and the contra-anteriority aspect *-mo* on the finite verb (*po-V*), which do not use auxiliaries. Some examples:

- (46) *Nta mo-'ala-aku-mo e'e.*
 FUT AF/NF-take-1sABS-PRF water
 'I'm going to fetch water.' (mVAb)
- (47) *Nta mom-pule-kami bede.*
 FUT AF/NF-restore-1peABS dike
 'We're going to repair dikes.' (perc) (mVAb)
- (48) *O kia, po-'awa-aku-si ku-'ari lako d[um]ahu.*
 O friend AF-get-1sABS-PRF 1sNOM-finish go [INT]dog
 'Hey, friend, I got something when I went hunting.' (perc7) (pVAb;NAuV;V)
- (49) *Po-lesa-ko-mo ta'i sapi.*
 AF-step.on-2sABS-PRF faeces cow
 'You have stepped on some cow manure.' (perc10)

This construction resembles an antipassive construction in that it marks the actor of a transitive verb with an absolutive marker. However it does not fulfil the criteria of a prototypical antipassive construction with regard to its undergoer. As mentioned above, typical antipassive clauses may be objectless, or the undergoer may be verb-incorporated, or it may be marked as oblique. These strategies of dealing with the undergoer are related to the categorisation of the clause as intransitive. None of these characteristics match the Moronene absolutive actor construction. The undergoer appears in the normal position for an object after the verb without any oblique marking (in Moronene, oblique constituents can be marked with the general preposition *hai* as in examples (7) and (44)). Nor is this construction restricted to thematically unimportant objects. Note the following examples with definite objects. The second example below uses zero anaphora (shared argument) with a topical object.

- (50) *Nta mom-po-'engka-'o-mo tama-no Eka, yo laica-no.*
 FUT AF/NF-CAUS-stand-3sABS-PRF father-3sPOS Eka ART house-3sPOS
 'Eka's father is going to build his house.' (c28) (mVAb)
- (51) A: *Daa-ko mo-wawa sica?*
 be-2sABS AF/NF-carry brush
 B: *Da-haku. Nde'e nta mo-'ala-ko?*
 be-1sABS Q FUT AF/NF-take-2sABS
 A: 'Do you have a brush with you?' (AuAbmV)
 B: 'Yes I do. Do you want to borrow it?' (perc2)⁷ (mVAb)

These considerations indicate that the Moronene absolutive actor construction differs in some ways from a prototypical antipassive construction. With regard to the somewhat similar subject focus construction in Konjo (including example (34) above), Friberg (1996:144) says 'These types of constructions could be viewed as antipassive, but the antipassive analysis is not justified in Konjo'. With regard to similar constructions in Uma which Martens (1988:176–177) called antipassive, Himmelmann (1996:119) cautions as follows: 'With an "ergative" analysis, the antipassive solution is problematic, since the

⁷ This example is discussed in Mead (1998:180).

UNDERGOER (P) is not marked as an oblique participant (which is especially awkward in examples ... where it is definite and specific)'. The Moronene constructions are best handled in the same way. They are best regarded as transitive clauses, not intransitive. Using a syntactic definition of antipassive, the Moronene *moN-V* forms, even when occurring in the absolutive actor construction, are best regarded as something other than antipassive.

2.4 Topicality pattern of *moN-V* forms⁸

To complement the evaluation of the grammatical features of *moN-V* constructions, it is helpful to investigate the topicality pattern of *moN-* verbs in discourse. Various studies have found correlations between measures of referential distance and topic persistence and different voice constructions, such as active, passive, inverse, and antipassive. A major function of voice is to indicate the relative topicality of the agent and patient. The four voices can thus be defined pragmatically in terms of the relative topicality of the agent and patient as shown in Table 1:

Table 1: Relative topicality of the agent and patient in the four main voices (Givón 1994:8)

Voice	Relative topicality
active-direct	AGT > PAT
Inverse	AGT < PAT
Passive	AGT << PAT
Antipassive	AGT >> PAT

Givón gives the following pragmatic definitions for each of the voices:

Active-direct: The agent is more topical than the patient, but the patient retains considerable topicality.

Inverse: The patient is more topical than the agent, but the agent retains considerable topicality.

Passive: The patient is more topical than the agent, and the agent is extremely non-topical ('suppressed', 'demoted').

Antipassive: The agent is more topical than the patient, and the patient is extremely non-topical ('suppressed', 'demoted'). (Givón 1994:8–9)

2.4.1 Topicality patterns of voice

Givón and other scholars have developed statistical measures which can have a diagnostic function, helping the researcher to decide which constructions in a particular language are fulfilling topicality functions similar to active, passive, inverse, or antipassive constructions in other languages. These are the measures of referential distance and topic persistence. In Givón, ed. (1994) are found detailed studies of these measures for eleven

⁸ The following sections are adapted from Chapter 6 of Andersen (2001)

languages. The way these studies have linked the measures of referential distance and topic persistence to different voices is as follows:

The assumption is that the relative topicality of the agent and patient helps determine the choice between different voices. As shown in Table 1, one would expect that active-direct voice would usually have a highly topical agent and a patient of medium topicality. Inverse voice would usually have a highly topical patient and an agent of medium topicality. Antipassive voice would usually have a highly topical agent and a patient of low topicality. Passive voice would usually have a highly topical patient and an agent of low topicality.

The use of the word ‘usually’ in the previous sentences indicates that not every use of a particular voice will follow the prototypical pattern. Hence one would expect that a large percentage of clauses in a particular voice would follow the expected topicality pattern, but there would be a small percentage which departed from the expected pattern. These percentages have been calculated in the studies found in Givón, ed. (1994).

For each voice, the percentages of clauses with low, medium or high referential distance is calculated. A referential distance of one, which indicates that the referent was mentioned in the previous clause, is defined as low. Low referential distance is associated with high topicality. A referential distance of two or three is defined as medium. A referential distance of more than three is defined as high. High referential distance is associated with low topicality.

For topic persistence, two categories are used, high and low. A topic persistence of zero to two is defined as low. This means that the referent is mentioned only zero to two times in the following ten clauses. Low topic persistence indicates low topicality. A topic persistence of three to ten is defined as high. This means that the referent is mentioned three or more times in the following ten clauses. High topic persistence indicates high topicality.

2.4.2 Antipassive voice in Moronene

Mead (1998:172–180) suggests that in Moronene, verbs with *moN-* or *poN-* prefixes function as an antipassive voice. In order to test this hypothesis, the referential distance and topic persistence was calculated for semantically transitive clauses in three Moronene narrative texts. They consisted of two prose texts, *Petampu'uno Ica Diu* (392 clauses long) and *Colisi* (229 clauses long) and a 566 clause extract from one narrative epic poetry text, *Kada*. In the set of semantically transitive clauses from the three narrative texts, there are 47 clauses with *moN-* or *poN-* prefixes. The referential distance and topic persistence percentages for these 47 clauses are set out in Table 2.

Table 2: Topicality values for *MoN-v/pon-v* clauses

Referential distance			Topic persistence		
Topicality	Agent	Patient	Topicality	Agent	Patient
High (1)	74%	30%	High (>2)	85%	17%
Medium (2–3)	23%	15%			
Low (>3)	2%	55%	Low (0–2)	15%	83%

What we expect in an antipassive voice is a highly topical agent and a patient with low topicality. Comparing the figures in Table 2 (Morenene) with the typical values set out in Table 3 below (Bella Coola [Salishan, Canada]), it is clear enough that this construction does act like an antipassive voice. The referential distance figures show that the agent is more topical than the patient. There are 74% of agents which have referential distance of 1 (high topicality) versus 30% of patients. Almost no agents (2%) have referential distance over 3 (indicating low topicality), whereas 55% of patients do.

The topic persistence figures are even more decisive: 85% of the agents are topical (high topic persistence) whereas only 17% of the patients are topical (most have low topic persistence). These figures are fairly close to those for the antipassive in Bella Coola. In Table 3 we give the figures for the *-m* antipassive in Bella Coola (Forrest 1994:160) as a basis for comparison. One difference is that the Moronene forms have slightly more topical patients than is the case in the Bella Coola antipassive.

Table 3: Topicality values for antipassive voice in Bella Coola

Referential distance			Topic persistence		
Topicality	Agent	Patient	Topicality	Agent	Patient
High (1)	80%	25%	High (>2)	85%	10%
Medium (2–3)	5%	10%			
Low (>3)	15%	65%	Low (0–2)	15%	90%

In terms of topicality, *moN-V* forms in Moronene are somewhat similar to antipassive verbs found in some other languages. But there are also some less typical characteristics. While the patient is usually non-topical, there is a significant minority of topical patients.

3 Cognitive grammar analysis

Before we examine the Moronene data in more detail, it will be helpful to review some aspects of verbal semantics based on cognitive grammar. According to the approach of cognitive grammar expounded by Ronald Langacker, a verb is a symbolic expression whose semantic pole designates a process which involves a continuous series of changes of state through a span of conceived time. To constitute a processual predication ‘a series of component states (not just a single state) must be profiled’ (Langacker 1987:246). Yet at the same time ‘the profiled process is prototypically ... construed as constituting a single event’ (Langacker 1991b:212). This helps us bear in mind that there are a number of possible components in internal semantic structure, which include elements relating to action, participants, change of state and conceived time. One predication may be a single event of ‘kill’ in which phases of one person causing someone to die and the other person dying are combined within a unity of time, whereas in another case it may be just the phase of causing someone to die which is profiled. Choices of verb forms are available on the basis of what should be included within the scope of predication, which can include participants, setting, clusters of acts or actions or the whole process. Or the choice of a particular verb form may be due to specific variants in the internal semantic structure of a verb itself. There are various possible variants of each component. A verb form may be chosen to designate a particular variant of one semantic component or to designate a

combination of variants from more than one semantic component, such as the temporal component and the quality of action.

Before doing any further interpretation, we need to examine verbs using the approach of cognitive grammar. A verb denotes either an act, or series of actions or a state. Lyons suggests that the proposition of a clause should be called a situation, which may be an event, a process or a state. A situation may be static or dynamic (Lyons 1977:483). Let us take a look at a dynamic situation. A situation is dynamic because, if we describe it cognitively according to Langacker, there is involvement of energy, either physical or mental. When there are two participants, a so-called two-place predicate, there is an energy flow from the energy source to the energy sink (Langacker 1991a:292), that is from the actor to the undergoer, and it implies a change of state component, or the affectedness of the object. Affectedness of the object may be considered as the final or end point of an action, giving a clear boundary or boundedness of an action.

The following diagrams may help clarify this matter:

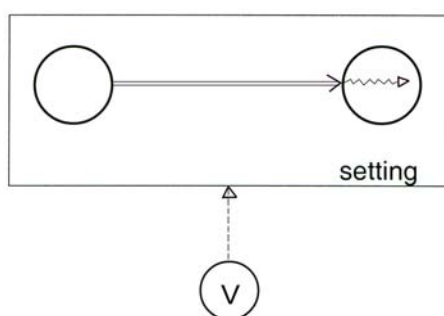


Figure 1: (Langacker 1991b:211)

In reference to the above diagram Langacker (1991b:210–211) says:

we arrive at a more complex conceptualization representing the ‘normal observation of a prototypical action’, whose essential content is diagramed in Figure 1. The stage model contributes the notion of a viewer (V) observing an event from a vantage point external to its setting. ...The event constitutes a prototypical action when it focuses on two participants construed as instantiating the maximally opposed role archetypes, namely agent and patient. The transmission of energy is depicted by a double arrow, whose direction serves to distinguish the agent and patient participants. The squiggly arrow indicates the patient’s resulting change of state.

Langacker (1991b:216) gives a number of English example sentences which he diagrams (1991b:217) in Figure 2.

- (5) a. Floyd broke the glass (with the hammer).
b. The hammer (easily) broke the glass.
c. The glass (easily) broke.
- (6) a. Floyd hit the glass (with the hammer).
b. The hammer hit the glass.
c. Floyd hit the hammer against the glass.

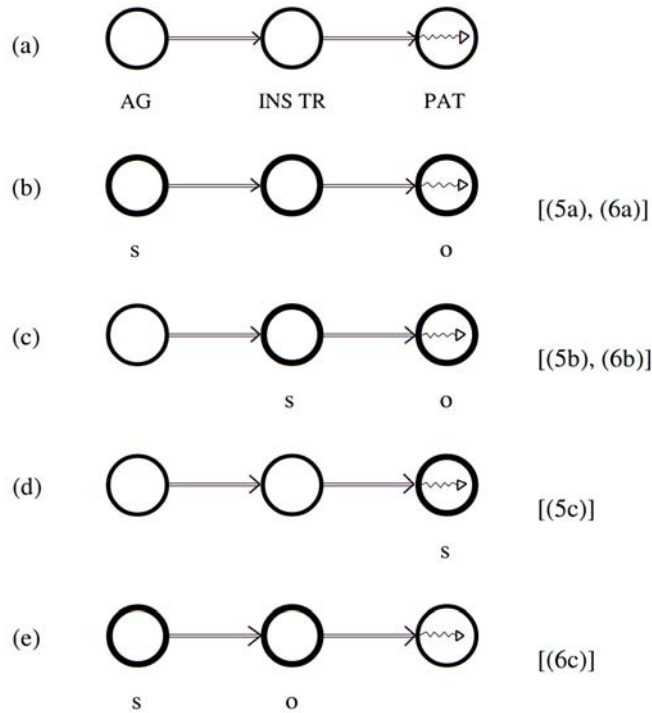


Figure 2: (Langacker 1991b:217)

Langacker (1991b:216–217) explains the diagram as follows:

In (5a), *break* profiles the entire action chain connecting the agent and patient, and the agent is selected as subject. In (5b), which designates the interaction between the instrument and the patient, the subject is the instrument. Only the patient's change of state is profiled by *break* in (5c), and the patient is chosen as subject. ... Though an oversimplification, it is sufficient here to analyze (6c) as profiling only the first segment of the action chain, as shown in Figure [2(e)]; *hit* is thus attributed the approximate value 'wield in a hitting-type motion', and the prepositional phrase specifies the resulting contact with the glass. Except for the change from *break* to *hit*, (6a) and (6b) are parallel to their counterparts in (5).

On the basis of the above cognitive explanation, we hypothesise that when one uses the form of verb with ABS object agreement, it designates the assertion of a vivid or a dynamic action, a kind of action which involves energy flow in a way which is instantaneous, punctual, direct, and forceful. In contrast we hypothesise that *moN-V* denotes a less dynamic action, which is less punctual, less direct, less volitional. This can be reflected in the lexical meaning itself, either as a less concrete action or as a more static process.

4 Patterns of semantic variation

In this section we will present various hypotheses as to factors influencing the choice of a *moN-V* form as against a V-ABS form which can explain seemingly anomalous uses — that is, when the object is definite and more topical.

The discourse function of an antipassive is to show low topicality of the object. It was demonstrated above that this is indeed the case with the large majority of objects used with *moN-V*. But what about the others? In the texts studied, 17% of the *moN-V* clauses had objects with high topicality. Why were *moN-V* forms used in these clauses? Let us examine some specific examples.

The clause in the following example occurs three clauses after the ‘thirst fish’ has been introduced as an important topical prop. In fact the events surrounding the ‘thirst fish’ represent the instigating incident of the whole story. Note the use of a definite noun phrase object with the *moN-V*.

- (52) *Ka-i po-'oli koie ica mokokondo'u e'e.*
 then-3sNOM AF-buy that fish thirsty water
 ‘Then he bought that “thirst fish”.’ (diu9) (NpV)

Three clauses later the fish is the topical object by zero anaphora of two *moN-V* forms. In this part of the text, the fish is a participant in eight successive clauses.

- (53) *Ka-i po-nahu arumai ka-i pong-kaa.*
 then-3sNOM AF-cook heard then-3sNOM AF-eat
 ‘Then she cooked it and ate it.’ (diu12) (NpV;NpV)

The following example shows another important prop, the treasure, as the object of a *moN-V* form. At this point in the text the treasure is a participant in three of the previous four clauses as well as the following two clauses.

- (54) *Hai hapa ari-a-u mo-'ala co'o ana*
 at what finish-LOC-2sPOS AF/NF-take you(sg) child
n-tina-'ate koie yo arataa?
 LG-woman-little that ART treasure
 ‘Little girl, where did you get that treasure?’ (diu112) (mV)

The following example shows another important prop, the cooking spices, as the object of a *moN-V* form. At this point in the text these cooking spices have been mentioned four times previously.

- (55) *Dadi-si koma i-tonia ki-ku daa kaa-hira ka'iaa diie*
 become-CTR right at-new if-1sNOM be eat-3pABS than this
hi-ku ari-mo lako mo-'ala nta po-gule-ku.
 COMPL-1sNOM finish-PRF go AF/NF-take FUT NR-curry-1sPOS
 ‘It would have been better just now if I had eaten them rather than what I’ve done going and getting my cooking spices.’ (Maegani 68) (NAuVAB; NAuV;mV)

In these examples, the prototypical meaning of *moN-* verbs does not fit. The object is specific and definite rather than nonspecific. It is topical rather than nontopical. Hence we need to discover other extended meanings of the *moN-* verb form which will explain these usages.

In the following sections, various semantic factors will be identified which are associated with the use of *moN-V* forms with definite objects. Examples of such clauses will be presented. It must be understood that in each case a large number of examples of

V-ABS clauses having the same semantic features could have been given. It is usually the case that when there is a specific definite object, a V-ABS form is the default choice. The reason few examples of V-ABS clauses are given below is because such clauses are not the focus of this paper.

4.1 Whole process versus act/action

Note the two constructions below, when the objects for both verb stems are definite, but different verb forms are used:

- (56) *Ka-i po-nahu arumai ka-i pong-kaa.*
 then-3sNOM AF-cook heard then-3s NOM AF-eat
 ‘Then she cooked it and ate it.’ (diu12) (NpV;NpV)
- (57) *Leu-ho-mo kolopua ala-a bangke-no koie ndok*
 come-3sABS-PRF turtle take-3sABS corpse-3sPOS that monkey
ka-i sampali-o naamo nahu-o.
 then-3sNOM butcher-3sABS and cook-3sABS
 ‘Then the turtle came and took the corpse of that monkey and cut it up
 and cooked/boiled it.’ (kol 45) (VAb;tVAb;NVAb;tVAb)

In both cases the objects of *nahu* ‘cook’ are definite. In example (56) with the *moN-V* form (cited earlier as example (53)), the definite object is represented by zero anaphora. In the V-ABS construction in the example (67), the definite object is represented by an absolutive clitic.

A further way of clarifying the difference is by looking at the inherent semantic structure: *monahu* is a whole process while *nahu-o* is one possible stage in cooking or a more distinctive way of cooking.

Hence the choice between the verb *monahu* and *nahu-o* is a matter of chunking clusters of actions as against one action, that is the process as a whole versus one act/action. In a similar way, the process of loading can be divided up into acts such as an act of grabbing, or lifting up, or putting down. Other verbs which use the prefixed *moN-V* form or clitic V-ABS form to represent different semantic components of verbs include:

- mongkaa* ‘eat’ vs. *kaa-ho* ‘eat it’
mo-ndo’u ‘drink’ vs. *ndo’u-o* ‘drink it up’
mo-’ulea ‘serve(food)’ vs. *ulea-ho* ‘load it on’

The *moN-V* form profiles the entire action series, whereas the V-ABS form specifies one act or action, especially the one seemingly significant key action, such as loading food on the plate, not counting lifting food towards the eater. An act of the mouth opening for water or food is more vivid than the word representing the entire action of eating or drinking.

The following examples illustrate the difference between *mong-kaa* ‘eat’ and *kaa-ho* ‘eat it’:

- (58) *Ku-’ari mong-kaa induwa ni-wuhai ondo-no i Suriani*
 1sNOM-finish AF/NF-eat yesterday UF-steam yam-3sPOS PI Suriani

- kana to'u hi-cu daa mo-dea-dea.*
 like very COMPL-1sNOM be STV-RED-drugged
 'Yesterday I ate Suriani's steamed wild yams, it was as if I was drugged.'
 (dic:dea) (NAumV;NAuV)
- (59) *Ku-'ari mong-kaa yo roo mo-lombi da ni-owa-no*
 1sNOM-finish AF/NF-eat ART medicine STV-fat REL UF-bring-3sPOS
i Sultan hai Kandari
 PI Sultan at Kendari
 'I ate the health tonic that Sultan brought from Kendari.' (dic:dengke)
 (NAumV;nVG)
- (60) *Nilako-no-mo koie tuuna-ni-akono-'o kolopua*
 immediately-3sPOS-PRF that fall-LOC-3sBEN-3sABS turtle
ko-kona-a bolo nganga-no, laulau-no-mo na'a-na
 RED-hit-3sABS in mouth-3sPOS direct-3sPOS-PRF also-3s
kolopua kaa-ho.
 turtle eat-3sABS
 'Immediately he dropped it down to the turtle and it went straight in his
 mouth, and the turtle ate it up right away.' (kol 34) (tVAb;tVAb;tVAb)
- (61) *Ya-ho-po nde'e koie harimau ka-i saisai-no*
 then-3sABS-IMPF indeed that tiger then-3sNOM direct-3sPOS
saba rako-'o koie dara ka-i kaa-ho.
 appear catch-3sABS that horse then-3sNOM eat-3sABS
 'Then that tiger suddenly appeared, caught that horse and ate it up.' (col 71)
 (NV;tVAb;NVAb)
- (62) *Nilako-no kaa-hira orua-'ira.*
 immediately-3sPOS eat-3pABS two-3pABS
 'it immediately ate up both of them.' (wola 28) (tVAb)

When you are talking about a routine meal, one uses *mong-kaa* to describe the whole process. The object tends to be normal routine food. The use of *kaa-ho* focuses on the action of actually putting the food in one's mouth and devouring it. The type of food tends to be less predictable, less routine. The temporal component also tends to differ, with *mong-kaa* more likely to be durative or habitual and *kaa-ho* more likely to be punctual.

The following examples illustrate the difference between *mo-ndo'u* 'drink' and *ndo'u-o* 'drink it'. Note that in both examples the object is definite:

- (63) *O kia mi-'osie mo-ndo'u akoie tua mo-silu, ntada'a*
 O friend 2hNOM-don't AF/NF-drink that toddy STV-sour perhaps
tua lembahi.
 toddy long.time
 'Friend, don't drink that sour palm wine, perhaps it is old.' (dic:silu) (NAumV)
- (64) *Da-hoo lawa ndo'u-o.*
 be-3sABS receive drink-3sABS
 'She would receive it and drink it up.' (diu15) (AuAbV;tVAb)

- (65) *Sa-pura-no-mo susu da hai tonde-no hi ari*
 when-used.up-3sPOS-PRF milk REL at glass-3sPOS COMPL finish
ndo'u-o ka-i pe-tiani-ako penda e'e n-tee.
 drink-3sABS then-3sNOM INT-add-APPL again water LG-tea
 'When he was finished drinking the milk in his glass he filled it again with
 tea.' (dic:tianio) (NAuVAb;NV)

Note that in the above examples with *ndo'u-o*, the action series is broken into two acts.

The following examples illustrate the difference between *mo-'ulea* 'serve(food)/load' and *ulea-ho* 'load it':

- (66) *Sa-mo-taha-no ni-nahu-no ka-i*
 when-STV-cooked-3sPOS PAS-cook-3sPOS then-3sNOM
po-'ulea-hakondo hai piri.
 AF-serve-3pBEN at plate
 'When the food was cooked, he served it up for them on a plate.'
 (dahu19) (NpV)
- (67) *Ka-i ala-a ulea-ho hai bungku-no.*
 then-3sNOM take-3sABS load-3sABS at back-3sPOS
 'Then he took him and loaded him on his back.' (ndoke4) (NVAb;tVAb)
- (68) *Hai apa daha-no mpe mo-'ulea ue akoie oto*
 at what place-3sPOS often AF/NF-load rattan that car
da sadia tealo?
 REL always pass
 'Where is the place he keeps loading rattan onto that car that keeps passing
 by?' (dic:ulea) (mV;V)

In the first example the object is definite and is represented by zero anaphora. In the second example, the object is also definite, and the action series is broken up into two acts, each of which uses the V-ABS form. There seems to be a partial lexicalisation of the distinction between these two forms. Whereas the meaning 'load' can be found with either the *moN-V* form or the V-ABS form, the meaning 'serve' seems to be virtually restricted to the *moN-V* form. See further discussion of lexicalisation in §4.6 below.

The next examples relate to the verb *mo'ihii* 'to fill up'. Both examples come from the same text in which the crew of a boat want to fill up their boat with goods. The first example is the first occurrence of this verb in the text and represents the process as a whole, using a *moN-V*:

- (69) *Sampe arumai ka-i ari koie bangka, ka-ndo*
 until heard then-3sNOM finish that boat then-3pNOM
m-po-'ihii bangka-do.
 PL-AF-fill boat-3pPOS
 'When the boat was finished, they filled up their boat.' (Maegani 139)
 (NV;NpV)

The following clauses in the text describe a discussion about the goods they should buy to fill the boat up with. Then the process is described again, broken down into separate

actions of buying and filling. In this case the V-ABS form *me'ihii-ho* is used, since it represents the last act of the whole process.

- (70) *Ndo-sabe penda m-pe-balanda kokeena m-po-'oli yo*
 3pNOM-go.up again PL-INT-shop there PL-AF-buy ART
mina-mina pera-no da mo-wondu, yo sabu, pera arataa
 perfume all-3sPOS REL STV-fragrant ART soap all wealth
kokeena hai wonua, ka-ndo me-'ihii-ho koie bangka-do.
 there at place then-3pNOM PL-fill-3sABS that boat-3pPOS
 'They went up again to go shopping there and bought all sorts of fragrant
 perfumes, soap, all sorts of costly goods at that place, then they filled up
 their boat.' (Maegani 143) (NV;V;mV;NVAb)

4.2 Durative action versus punctual action

A second factor influencing the choice of verb form involves a distinction in the temporal element of the internal structure of the verb. The *moN-V* form designates a more durative action whereas the V-ABS form designates a more punctual action. Examples:

mo-wawa 'bring (with)' versus *wawa-a* 'take it away, bring it to'

- (71) *Daa-ko mo-wawa sica-'u?*
 be-2sABS AF/NF-bring brush-2sPOS
 'Do you bring your brush?' (perc2) (AuAbmV)
- (72) *Da-haku mo-wawa diie watu m-pe-'eso ka-u*
 be-1sABS AF/NF-bring this stone LG-INT-rub then-2sNOM
eso-aku itea ki-to pe-baho.
 rub-1sABS soon if-1piNOM INT-bathe
 'I am bringing this rubbing stone so you can rub me shortly if we bathe.'
 (dic:eso) (AuAbmV;NVAb;NV)

The above examples involve non-telic situations; the bringing of the brush or stone has no explicit end point. This contrasts to the examples below with V-ABS forms:

- (73) *Bawa-hira orua-'ira hai otu-ng-keu.*
 bring-3pABS two-3pABS at end-LG-wood
 'It carried the two of them to the top of the tree.' (wola27) (tVAb)
- (74) *Koie miano da mo-pusu me-'e'eta hai aku ka-ku*
 that person REL STV-blind INT-ask at me then-1sNOM
wawa-a lako hai garega.
 bring-3sABS go at church
 'That blind man asked me to take him to church.' (dic:e'eta) (V;NVAb;V)

The durative nature of the next example is indicated by the use of the auxiliary *daa* 'be', which gives the sense of "he was picking" as well as the following clauses of the narrative which describe other events that occurred as he was picking the fruit. Because *moN-V*

portrays an action in an overall holistic way, it has a greater chance to be selected to represent setting.

- (75) *Mo-mone ka-i daa mo-'upui wua-no koie*
 INT-climb then-3sNOM be AF-pick fruit-3sPOS that
to-tandai-ho ki daa k[in]aa.
 RED-try-3sABS if be [UF]eat
 'He climbed up and picked that fruit and tried it whether it was edible.'
 (Maegani 84) (V;NAumV;tVAb;NAunV)

Some further examples of *moN-V* clauses with definite objects with a durative meaning:

- (76) *Pempi-pempia-haku-mo i'aku ari mo-daga-daga koie*
 RED-how.many.times-1sABS-PRF I finish AF/NF-RED-guard that
ana-'ate.
 child-small
 'Time and time again I have taken care of that little child.' (dic:daga) (AumV)
- (77) *Koie Maegani hi daa men-tade mo-rede*
 that Maegani COMPL be INT-stand AF/NF-make.staccato.sound
pana'api koie hai rope-no banka.
 gun that at bow-3sPOS boat
 'That Maegani who was standing shooting off the gun at the bow of the boat'
 (Maegani 156) (NAuV;mV)

In case of temporal variation, between durative and punctual, cognitively they differ in space and time. In other words, a punctual action is more specific in the space and time domain of a verb than a durative action is. This is reflected in the contrast between *wawa-a* 'bring to' at a particular place and time and *mo-wawa* 'bring (with)' at a nonspecific place and during a duration of time.

It is equally possible to find V-ABS clauses in durative contexts when the object is definite. This simply means that the primary transitive usage of V-ABS takes precedence over a possible secondary usage of *moN-V*. An example:

- (78) *mau te'asa mincu ki-u dio-ho koie labu ...*
 even one week if-2sNOM keep-3sABS that pumpkin
 'even if you keep that pumpkin for a week ...' (dic:baku) (NVAb)

4.3 Non-volitional action versus volitional action

Another factor influencing the choice of verb form is volition. In this pattern, the *moN-V* form designates a less volitional action whereas the ABS-V form designates a more volitional action. This pattern is applicable only to verbs designating actions which commonly happen accidentally.

mo-lesa 'step on (accidentally)'
 vs. *lesa-'o* 'tread on it, step on it (on purpose)'

Some examples:

- (79) *Hawiako-hira ko'ira beke-beke-no tari da-hoo nta*
 throw-3pABS those RED-joint-3sPOS bamboo be-3sABS FUT
mo-lesa tuai-u ka-i bela karu-no.
 AF/NF-step.on younger.sibling-2sPOS then-3sNOM injured foot-3sPOS
 'Throw away those bamboo joints; your little sister will step on them and injure her foot.' (dic:beke) (tVAb;AuAbmV;NV)
- (80) *Osie-mo mo-'ala keu n-tangkalasi ki sa-u da'a*
 don't-IMPV AF/NF-take wood LG-thorn.tree if NEG-2sNOM NEG
to'ori te'iaa-mo ka-u po-lesa riu-no.
 know only-PR then-2sNOM AF-step.on thorn-3sPOS
 'Don't take that thorn tree wood; if you don't know it, you'll just step on its thorns.' (dic:tangkalasi) (AumV;NAuV;NpV)
- (81) *Ka-u osie to'u-o pendu-pendua lesa-'o bolo*
 then-2sNOM don't really-3sABS RED-twice step.on-3sABS in
laica-no i Dodi.
 house-3sPOS PI Dodi
 'Don't you ever set foot in Dodi's house again!' (dic:haramu) (NAuVAb)

The prototypical meaning of *moN-* can override this distinction between volitional and non-volitional. In the following example, the topicality of the object influences the choice of the V-ABS form, although the action is non-volitional.

- (82) *Tende me-bosi-no pe'ico bengkaro ongkona-a hi*
 due.to INT-spring-3sPOS that trap as.soon.as-3sABS COMPL
tealo dahu lesa-'o ba-i to-tamua dahu.
 pass dog step.on-3sABS how-3sNOM RED-thrown.up dog
 'Due to the springiness of that trap, as soon as a dog passes and steps on it, the dog will be really strung up.' (dic:bosi) (NV;tVAb;NV)

4.4 Irrealis versus realis

Another factor which may influence the choice of verb form is realis versus irrealis. Unlike some other languages in Sulawesi, Moronene does not have a separate set of irrealis verb prefixes. Some examples suggest that an irrealis context, especially negation, may favour the use of a *moN-V* form rather than a *ABS-V* form. Note the following examples of negated *moN-V* clauses with a definite object.

- (83) *Poh, ka-i rua me-ngkae Maegani, mau mo-'onto*
 gee then-3sNOM down INT-stetch Maegani even AF/NF-see
tina-no na-i da'a, da-hoo tutuwi-o
 woman-3sPOS NEG-3sNOM NEG be-3sABS close-3sABS -3sPOS

mata-no, ka-i ala-a rent-a koie tina-no.
 eye then-3sNOM take-3sABS pull-3sABS that mother-3sPOS
 Well, Maegani stretched his hand down, he didn't even see his mother, he
 was covering his eyes, then he grabbed his mother and pulled her.'
 (Maegani 171) (NAuV;mV;AuAbVAb;NVAb;tVAb)

Note that in the above example the clause in question, *mo'onto tinano*, is followed by two V-ABS clauses with *tinano* as the topical object.

- (84) *Kicua karambau-'ute-no ta'ico da ari lako ni-dosa-ngku*
 if buffalo-small-3sPOS that.above REL finish go UF-debt-1sPOS
da mate, da-haku nta wada-'o, da-haku nta oli-o,
 REL die be-1sABS FUT pay-3sABS be-1sABS FUT buy-3sABS
na-hoo nangkua karambau-'ea-no, na-ku da'a nta
 be-3sABS because buffalo-big-3sPOS NEG-1sNOM NEG FUT
mo-'oli na-mo-nangi.
 AF/NF-buy because-INT-lose
 'If his small water buffalo which I had gone and borrowed was the one
 that died, I would pay for it, I would buy it, but because it was his big
 water buffalo, I wouldn't buy it because it lost.' (Maegani 250)
 (AuV;nVG;V;AuAbVAb; AuAbVAb;NAumV;V)

The above example gives a nice contrast between the use of the V-ABS form *oli-o* in a positive clause, followed by the use of the *moN-V* form *mo-'oli* in a negative clause. The objects of both clauses have the same definiteness and topicality. The following two clauses representing successive clauses in a dialog show the identical contrast:

- (85) *Da-hoo ndoka nta leu ro'ico mokole oli-o*
 be-3sABS please FUT come that.below king buy-3sABS
miano i cantete.
 person at latrine
 'That king down there is going to come to buy the person in the latrine.'
 (Maegani 161) (AuAbV;tVAb)
- (86) *Hee, osie mo-'oli miano i cantete; me'alu-o miano*
 hey don't AF/NF-buy person at latrine many-3sABS person
i laica.
 at house
 'Hey, don't buy the person in the latrine; there are many people in the house.'
 (Maegani 162) (AumV;VAb)

Some further examples of negated clauses with definite objects using *moN-V* forms:

- (87) *Na-ku paisa mo-'ala doi-'u.*
 NEG-1sNOM never AF/NF-take money-2sPOS
 'I never took/stole your money.' (dic:ala) (NAumV)

- (88) *Mo-hali to'u-o. Osie-mo, nda'a-mo nta mo-'ala.*
 STV-expensive very-3sABS don't-IMPV not-PRF will AF/NF-take
 'It is very expensive. No, I will not buy it.' (perc 4.e) (VAb;AumV)

In (88), the implicit definite object can be interpreted as a case of zero anaphora.

Other irrealis contexts include imperative clauses and future tense. Note the following imperative examples with definite objects using *moN-V* forms:

- (89) *Po-hule-mo na'a-u kia, mo-'al karambau-'u.*
 INT-return-IMPV also-2s friend AF/NF-take buffalo-2sPOS
 'Go home, friend, get your water buffalo.' (Maegani 204) (V;mV)
- (90) *Po-'ala-mo na'a-u kia, manu-'u.*
 AF-take-IMPV also-2s friend chicken-2sPOS
 'Friend, get your chicken.' (Maegani 258) (pV)
- (91) *Lako-mo mo-dampo pe'ico-'ira kinaa.*
 go-IMP AF/NF-cover yon-PL rice
 'Go cover those plates of rice.' (dic:dampo) (V;mV)

The following examples illustrate the difference between *mo-daga* versus *daga-'o* 'guard/mind'. The *moN-V* form is used in an irrealis future context, whereas the V-ABS form is used in a past realis context.

- (92) *Ico'o da nta mo-daga-'akita laica*
 you(sg) REL FUT AF/NF-guard-1piBEN house
 'You are the one who will guard the house for us.' (col 13) (mV)
- (93) *Ka-i totoro ka'asi koie i Siti daga-'o*
 then-3sNOM sit CMS that PI Siti guard-3sABS
tukaka-no.
 elder.sibling-3sPOS
 'Then poor Siti sat and watched over her elder brother.' (sio59) (NV;tVAb)

The following are further examples of *moN-V* forms being used in an irrealis future context with definite objects:

- (94) *Hapa-mo ka'asi co'o ka-u daa nta lako mo-ro-rongo*
 what-PRF poor you(sg) then-2sNOM be FUT go AF/NF-RED-carry
koie bio?
 that egg
 'What will you go and carry that egg with, poor thing?' (Maegani 280)
 (NAuV;mV)
- (95) *Mau hi-to teleu nta mo-'ita mo-'oli manu-no,*
 even if-1piNOM arrive FUT AF/NF-ask AF/NF-buy chicken-3sPOS
mo-'oli karambau-no, na-i ehe nta mo-wee-kita.
 AF/NF-buy buffalo-3sPOS NEG-3sNOM want FUT TR-give-1piABS
 'Even if we arrive and ask to buy his chicken, buy his water buffalo, he won't
 want to give them to us.' (Maegani 323) (NV;AumV;mV;mV;NAuVAb)

Note that it is also possible to find V-ABS clauses in irrealis contexts, whether negative, imperative, or future. Some examples:

- (96) *Na-i da'a-mo nta awa-a sandala-no da tuna.*
 NEG-3sNOM NEG-PRF FUT find-3sABS sandal-3sPOS REL fall
 'He's not going to find his sandal which fell down.' (dic:sempo) (NAuVAb;V)
- (97) *Dio-ho-mo isala koie ni-wawa-u.*
 put.away-3sABS-IMPV first that UF-carry-2sPOS
 'Put down what you are carrying first.' (dic:sakoi) (tVAb)
- (98) *O Dedi da-ko-si bara nta poko-sicu-'o koie dopi?*
 O Dedi be-2sABS-CTR Q FUT able-shift-3sABS that board
 'Dedi, will you be able to shift that board?' (dic:sicu) (AuAbVAb)

4.5 Non-individuated object versus individuated object

There are several factors which relate to the individuation of the object. Perhaps the most obvious is the definiteness or specificity of the object, and this factor accounts for the choice of verb form in the large majority of cases. Another factor influencing the choice of a *moN-V* with a definite object is animacy. An inanimate object is less individuated than an animate object. Whereas the V-ABS form occurs freely with objects of the highest level of animacy, namely human beings, the objects of *moN-V* forms have lower animacy. In almost all of the examples cited thus far, such objects are animals or inanimate objects. The following examples illustrate the difference in usage of *mon-toria* 'guard/mind' versus *toria-ho* 'guard it/mind it' with regard to animacy:

- (99) *Ma-ngku-mo i'aku-mo da mon-toria laica.*
 later-1s-PRF I-CTRS REL AF/NF-guard house
 'Later I will be the one who guards the house.' (co14) (mV)
- (100) *Po-'ia-mo nde'e na'a-u toria-ho ana-'u*
 INT-stay-IMPV indeed also-2s guard-3sABS child-2sPOS
 'Just stay and mind your child.' (lang42) (V;tVAb)

Although both objects are definite, and both clauses are irrealis, we hypothesise that the *moN-V* is used with less animate, less individuated, less topical objects, while the V-ABS form used with more topical individuated animate objects.

Only one example has been found so far in the texts examined of a *moN-V* form used with a definite (and highly topical) human object. It was given as example (83) in the previous subsection.

Another factor regarding individuation of the object is plurality. Plural objects or mass objects are less individuated than singular objects. When the V-ABS form is used, a distinction is made between a singular object marked with *-o* and a plural object marked with *-'ira*. The plural suffix, however, tends to be restricted to humans. If the object is plural and inanimate, one way of avoiding the choice between singular and plural absolutive clitics is to simply use the *moN-V* form which does not mark the object morphologically. The following examples illustrate the use of *moN-V* forms in clauses with definite plural inanimate objects:

- (101) *Ka-ndo ala-a nde'e sabe mo-'ala ko-'ira yo pera*
 then-3pNOM take-3sABS indeed go.up AF/NF-take that-PL ART all
sabu da me-mo-wondu
 soap REL PL-STV-fragrant
 'They went up and took all those fragrant soaps.' (Maegani 175)
 (NVAb;V;mV)
- (102) *Lako mo-'o-'ala pera nta po-gule-no koie*
 go AF/NF-RED-take all FUT NR-curry-3sPOS that
 'He went to gather all those spices he would use.' (Maegani 56) (V;mV)

It is of course also possible to find V-ABS clauses in with plural inanimate definite objects. This simply means that the primary 'volitional agent + affected undergoer' usage of V-ABS takes precedence over the possible secondary 'non-individuation of undergoer' usage of *moN-V*.

An example:

- (103) *Lako-mo poko-tinda-hira luwuo sabara-u.*
 go-IMPV CAUS-pack-3pABS all belongings-2sPOS
 'Go pack up all your belongings.' (dic:sabara) (V;tVAb)

4.6 Lexicalisation

It was mentioned above that in some verbs such as *mo-'ulea* 'serve/load' versus *ulea-ho* 'load it', there is a partial lexicalisation of the semantic distinction between the two forms. A thorough investigation of this phenomenon is beyond the scope of this paper, but we will mention several further examples.

One type of lexicalisation relates to the presence of the partly homophonous intransitive prefix *mo-/po-* or the stative prefix *moN-*. When this prefix occurs with a verb root which also is a transitive verb form, there can be an avoidance of the transitive *moN-V* form to avoid ambiguity. Instead the V-ABS form is always used for transitive. Some examples:

Intransitive		Transitive
<i>mo-gora</i>	'be noisy'	<i>gora-'o</i> 'make noisy'
<i>mo-nangi</i>	'lose'	<i>nangi-o</i> 'defeat'
<i>mom-pande</i>	'work at trade'	<i>pande-'o</i> 'know'
<i>mon-to'ori</i>	'adult'	<i>to'ori-o</i> 'know'

The theoretically possible transitive forms **mo-gora* 'make noisy' **mo-nangi* 'defeat', **mom-pande* 'know', and **mon-to'ori* 'know' are not attested.

The type of lexicalisation exemplified *mo-'ulea* 'serve/load' versus *ulea-ho* 'load it', is when both transitive forms exist, but certain meanings are restricted to one of the forms. For example, the verb forms *mo-wolohi* and *wolohi-o* can mean 'give in return, repay'. There is also a secondary meaning 'replace (lost or damaged item)', but this meaning is restricted to the V-ABS form *wolohi-o*. Similarly the verb forms *mon-timpa* and *timpa-a* mean 'find'. There is also an idiomatic expression *timpa-a laro* meaning 'regain consciousness'. This idiom is restricted to the V-ABS form.

4.7 Summary

From the above survey, the grammatical distinction of V-ABS versus *moN*-V seems to be used for grouping verbs into two contrastive categories with a semantic motivation. The semantic variants may be due to the action component or the more relational component of the verb, that is the object. The action component includes the temporal component (punctual as against durative) as well as volitional as against nonvolitional.

The relational component involves variation of the object in quantity or specificity. The semantic field is forced to be wide or narrow only because of the size of the object, or else the unclear boundary of the object, mingled among many, until one is singled out.

For the purposes of exposition, each of the illustrative examples has been listed under one particular extended meaning. But with some examples, several meanings are combined at once. The following clause, for example, comes simultaneously under the category of irrealis, durative and non-individuated object:

- (104) *Lako mo-dai-dai-'akita ko-'ira ica-'ate hai towo ni'i.*
 go AF/NF-RED-burn-1piBEN that-PL fish-small at shell coconut
 'Go grill those little fish for us in a coconut shell.' (dic:dai) (V;mV)

Many of the semantic factors discussed above can be related to the overarching category of transitivity. The seminal paper on this topic is Hopper and Thompson (1980). They define transitivity in terms of a set of ten independent features, each of which has a high transitivity value and a low transitivity value. The overall transitivity of a clause can be measured through the combination of these features. The individual features identified by Hopper and Thompson are as follows:

Table 4: Features of transitivity
(Hopper & Thompson 1980:252)

	HIGH	LOW
A. PARTICIPANTS	2 or more participants	1 participant
B. KINESIS	Action	Non-action
C. ASPECT	Telic	Atelic
D. PUNCTUALITY	Punctual	Non-punctual
E. VOLITIONALITY	Volitional	Non-volitional
F. AFFIRMATION	Affirmative	Negative
G. MODE	Realis	Irrealis
H. AGENCY	A high in potency	A low in potency
I. AFFECTEDNESS OF O	O totally affected	O not affected
J. INDIVIDUATION OF O	O highly individuated	O non-individuated

The last feature, individuation of object, can be further analysed according to different nominal features which correlate with individuation. These are set out below :

Table 5: Features of object individuation
(Hopper & Thompson 1980:253)

INDIVIDUATED	NON-INDIVIDUATED
Proper	Common
Human, animate	Inanimate
Concrete	Abstract
Singular	Plural
Count	Mass
Referential, definite	Non-referential

Among the factors affecting transitivity propounded by Hopper and Thompson, seven are mentioned in the previous sections: definiteness of object, durativity, volition, affirmative/negative, realis/irrealis, plurality and animacy. We can then encompass many of the semantic influences by stating the V-ABS forms are used in contexts of higher transitivity whereas *moN-V* forms are characterised by lower transitivity.

Having a widely open semantic field, the *moN-V* form seems to cover more variations of lexical meaning patterns than the V-ABS. It seems like a relatively unmarked or general verb form used for more general functions of low transitivity.

5 Photographic analogy

The data presented in the previous section can be explained by propounding a polysemy network for the various meanings of *moN-V*. The primary meaning would be antipassive; and there would be the following extended meanings: whole process, durative, non-volitional, irrealis, non-individuated object. One question that arises is: what is the point of contact semantically which sanctions such extensions of meaning?

One helpful analogy to explain the contrastive behavior of *moN-V* versus V-ABS is how we take a photograph. The implications of this analogy will shed light on some of the interconnections between the various meanings of *moN-V* forms.

In the following discussion we will examine a number of statements about the distinction between *moN-V* versus V-ABS forms and see how these statements can be elucidated using a photographic analogy.

1. *moN-V* portrays an action in an overall or general sense.

The *moN-V* form has its semantic field relatively widely open with regard to its semantic variants. To obtain an overall or general picture, one uses a simple camera, or a sophisticated camera, or even better a video camera with a wide angle lens to give a general impression to the viewer of the overall situation or scene. That is why including a specific object does not change the overall or general impression.

The *moN-V* form is more or less a default or general verb for an action. In many cases the *moN-V* form is the citation verb form. That is to say, if one elicits a verb from a Moronene speaker using the national language as a cue, the verb form elicited will usually be the *moN-V* form.

2. In certain instances, a non-specific object tends to involve a less dynamic action or a more general verb.

A picture giving an overall or general impression tends to be more characteristically inclusive and so the object tends to be inclusive or general and does not need to be made explicit. Such objects are highly predictable, stereotypical, non-referential. A general impression of the object is also obtained by looking at the object as a by-product of the process or in a generalised way, such as food being the object of eating or cooking or serving, water being the object of drinking, a house/building being the object of building, clothes being the object of washing. Such general predictable objects are often suppressed:

- (105) *Impia ka-i pom-po-'engka?*
 when then-3sNOM AF-CAUS-get.up
 'When will he build it?' (perc28) (NpV)

Such an object, however, may be made explicit:

- (106) *Tade-tade-no-mo nde'e mo-kea mo-ndo'u e'e.*
 suddenly-3sPOS-PRF indeed INT-want AF/NF-drink water
 'Suddenly she wanted to drink water.' (diu 13) (V;mV)

3. A dynamic action or specific verb requires a specific object, not necessarily a definite object.

For a dynamic action or specific verb, the action is more punctual, direct or concrete. The object can have some sort of boundaries to sharply direct the action, to achieve a dynamic action. In relation to the way a verb presents a situation, the V-ABS form can be compared to a zoomed focus. To achieve a zoomed picture, it is easier when there is something definite or referential in the speaker's and the hearer's knowledge. But in general a zoomed focus just requires a specific object.

- (107) *Ka-i awa-a olumpu-'ute.*
 then-3sNOM find-3sABS hut-little
 'Then he found a little hut.' (sio52) (NVAb)

However, even when using a wide-angle lense, it is still possible to focus on a particular object. In other words, it is still possible to focus on a definite object when using a *moN-V* form. One situation in which this is appropriate is where the object is a type or group specification (Langacker 1991a:53) rather than a specific individual. When one type is contrasted with another, this is sufficient reason to justify making a contrastive focus. The photographer focuses on one entity to show contrast with another entity which is left out of focus. Consider this example:

- (108) *Mo-tasu koie na'a-na pae ka-u rabusi-o.*
 AF/NF-dibble that also-3s rice then-2sNOM pull.up-3sABS
 'You dibble that rice, then you pull out weeds.' (roo10) (mV;NVAb)

From the context, the focus is not on one particular rice plant, but on rice as a type of plant in contrast to corn which was mentioned earlier in the text. This can fall under the category of non-individuated object.

When a V-ABS form is used, the zoomed focus may be on an object noun phrase which does not seem to be marked grammatically as specific, but is specific from the context:

- (109) *Ka-ndo daa si-simbu tungkale-e ewo, tungkale-e*
 then-3pNOM be RED-wander rummage-3sABS grass rummage-3sABS
saro.
 rubbish
 ‘They wandered around rummaging in the grass, rummaging in the rubbish.’
 (col 26) (NAuV;tVAb;tVAb)

In this case, the boundary of the *ewo* ‘grass’ is established by shared knowledge that it designates the grass around that particular house. Another example is an object which is one mingled among many being brought into individual identity or existence.

- (110) *Lau-lau-no-mo totoro naamo mo-'upu ka-i*
 RED-until-3sPOS-PRF sit and AF/NF-pick then-3sNOM
kuli-si-o ronga kaa-ho.
 skin-APPL-3sABS with eat-3sABS
 ‘he went and sat down and picked one, then he peeled it and ate it up.’
 (kol25) (V;mV;NVAb;tVAb)⁹

Even a mass noun with an unspecified boundary may gain a boundary due to a previous action and thus be available for zooming:

- (111) *Ka-i um-ahi ... mo-'o-'ala-akono tina-no*
 then-3sNOM INT-draw.water AF/NF-RED-take-3sBEN mother-3sPOS
e'e, da-hoo leu wowa-a, da-hoo lawa ndo'u-o.
 water be-3sABS come carry-3sABS be-3sABS receive drink-3sABS
 ‘Then she drew water ... and fetched it repeatedly for her mother, she would bring it, and she would receive it and drink it up.’ (diu15)
 (NV;mV;AuAbV;tVAb;AuAbV;tVAb)

In the above example, when *e'e* ‘water’ is first mentioned with the verb *mo'o'ala* ‘take’, it has no boundary. However when it is brought from the well *wowa-a*, its identity is fixed as the water which has just been fetched.

4. Less dynamic actions or general verbs need no object agreement on the verb, but dynamic actions or specific verbs do.

The most inner and final or end component of verb structure, which is the change of state component, may or may not be present. The speaker may not include the change of state component in his general or widely opened picture of the action. If it occurs, it is tightly bound to the particular dynamic action and the affected object, as seen more easily from the chart in §3 above (Figure 2). An overall action is not concerned for only one particular action or a particular object or the possible affectedness of the object. Consider the following examples:

- (112) *O kia po-'awa-aku, ari-aku mom-pando wawi.*
 O friend AF-get-1sABS finish-1sABS AF/NF-spear pig
 ‘Friend, I got it, I have speared a pig.’ (perc7) (pVAb;AuAbmV)

⁹ This example is discussed in Mead (1998:179).

- (113) *Inai da ari pando-o sapi-'u ka-i dungku*
 who REL finish spear-3sABS cow-2sPOS then-3sNOM until
me-luarako tariti-no.
 INT-go.out intestines-3sPOS
 'Who speared your cow so that its intestines spilled out?' (dic:tariti)
 (AuVAb;NV)

In the first example, *mompando wawi* 'speared a pig' is presented as an overall action. The object is indefinite. The verb *pando* 'spear' in both V-ABS and *moN-V* forms do not differ in terms of temporal component of the action, both have the same degree of directness of action and in both cases the object is specific. In the first example the speaker just wanted to present the overall scene of activity, so he was not concerned about presenting the end point of the activity and he did not specify the object. By itself, the *moN-V* form does not necessarily imply the success of the attack. The speaker gave a general picture of his state of success using the verb *po'awaaku* 'I got it' first, and only then mentioned the content of success.

The verb *mo-'awa* 'get/obtain' implies that the object is affected, but the affectedness is not as prominent as in the case of the V-ABS form *awa-a*. This is because the picture which is taken with *mo-'awa* has to cover the overall action, possibly all the physical and mental actions. So the prominence is on a general state of success rather than on the action as such.

Another set of interesting examples is:

- (114) *Da-hoo kolopua tangasa mo-aha pali-no.*
 be-3sABS turtle PROG AF/NF-sharpen axe-3sPOS
 'The turtle was sharpening his axe.' (kol20) (AuAbmV)
- (115) *Aha-'akita-'o ta'owu-nto.*
 sharpen-1piBEN-3sABS machete-1piPOS
 'Sharpen our machetes!' (perc24) (tVAb)
- (116) *Anto, aha-'akita-'o pali.*
 Anton sharpen-1piBEN-3sABS axe
 'Anton, sharpen the axe for us.' (perc31) (tVAb)

Note that both *mo-'aha/aha-'o* are used with a definite object. Based on the meaning of the verb *mo-'aha/aha-'o* 'sharpen', it is a durative action in both forms and both forms occur with a specific object. When one commands that one's machete should get sharpened, one would expect a result, namely the affectedness of one's machete. So the form V-ABS seems to match this expectation. In contrast, when one uses *mo'aha*, the affectedness is not in focus, but rather just describing the overall activity.

With a dynamic, specific verb, the object is a prototypical direct object, with salient affectedness, and so the object is more crucially involved in the action. The affected object together with the affectedness seem to be tightly bound and also close to the action which directly has contact with such an object. When the affectedness is in focus, automatically one zooms in on the object and the particular action next to it. So grammatically it is marked saliently with ABS clitic on the verb stem.

In relation to valency, Langacker explains how an object can be more tightly bound to the verb than the subject is, for objects are inner-layer participants. This is reflected in their

integration at the phonological pole (Langacker 1991b). This applies particularly, we suggest, in the case of a dynamic verb, because of the relation between the act causing the affectedness and the affected object.

5. It is misleading to label the *moN-V* form as actor or subject focus in the Philippinist sense (akin to ‘voice’).

The less dynamic verb has an object which may or may not be affected by the action and so it is involved, but not crucially, in the overall picture of the action. Grammatically it is hence not marked on the verb stem. The question is whether this *moN-V* form best described as subject pivot instead, which means the subject becomes a pivot. Some call it an actor focus or subject focus, as it is generally claimed that there are primary and secondary roles or pivots that are actor and undergoer or subject and object. According to Friberg (1996:143), in Konjo, ‘subject focus implies that there is no object or that the object is not relevant to the action at hand’; in other words the object is out of focus.

To explain why *moN-V* is not actor or subject focus, from the perception view, let us return to our photo-taking analogy again. In an overall picture the actor of an action is included, whereas in a zoomed picture only part of the actor is included, such as the hand or the mouth as the representative of the actor within the frame. To the viewer, seeing the hand or the mouth reminds him of the actor. With the general view provided by a *moN-V* form, the fact that one sees the whole body of the actor, but perhaps not so focused on one particular part, is not as important as the overall impression of the scene, with some other verb components also prominent. The primary role or pivot has not disappeared, yet it is not primarily prominent in the whole picture. It is not profiled. It is rather an action or activity which is in focus and profiled.

The use of the term ‘actor/subject focus’ in contrast to ‘undergoer/object focus’ gives the impression that each form zooms in and focuses in an equivalent way either on the actor subject or the undergoer object. But for Moronene, this would be a misleading impression. It might be fair enough to call the V-ABS form ‘object/undergoer focus’. The object is profiled. But a shift from a V-ABS form to a *moN-V* form does not imply that the photographer simply moves the camera a bit to focus on the actor rather than the undergoer. The actor is not profiled with a *moN-V* form. Rather it implies that the camera is not shifted, but instead the zoom is adjusted to a wider angle to take in the whole scene. Hence we suggest that some better terms to describe the *moN-V* form in Moronene are ‘action focus’¹⁰ or ‘activity focus’ or ‘scene focus’.

6 Conclusion

In many ways, the Moronene *moN-V* form may seem to have parallel behavior to similar forms in other Sulawesi languages. Yet some unexpected features of its usage, such as its occurrence with definite, specific objects in many Moronene texts, has been a stimulus for further investigation. Semantic analysis and a cognitive grammar approach have served well in clarifying this behavior. So instead of labelling it as actor or subject focus or as antipassive, it is labelled in several different ways to elucidate its behavior as understood from a cognitive-semantic approach.

¹⁰ Mead (1998:177) uses the term ‘action-focus’ to describe the Moronene *moN-V* form.

What are the best terms to use to describe the distinction between the V-ABS and *moN-V* forms? It depends which aspect of the distinction we want to emphasise. To reflect the semantic distinctions in verb behaviour, the terms ‘dynamic/less dynamic’ or ‘more transitive/less transitive’ can be used for V-ABS and *moN-V* forms respectively. The term ‘semitransitive’ is a helpful label for the *moN-V* forms to indicate that they are more transitive than intransitive verbs, but not as transitive as the V-ABS forms. To reflect perceptually-conditioned distinctions in the scope of predication, the terms ‘specific/general’ verb can be used for V-ABS and *moN-V* respectively. With regard to focus or profiling, the V-ABS forms can be characterised as ‘affectedness-focus’ and the *moN-V* forms as ‘action-focus’.

Himmelmann (2002:14) distinguishes two contrasting approaches to the analysis of voice phenomena in western Austronesian languages. In one approach, the so-called actor focus is analysed as an antipassive contrasting to an unmarked transitive construction. In the second approach, the different voices are regarded as ‘valency-neutral alternations’ so that in ‘actor voice the actor is the syntactic pivot, in undergoer voice the undergoer is the syntactic pivot, but both constructions share the same transitivity value’. The approach of this paper is perhaps more allied to the second approach than the first, but with some significant differences and refinements. The two Moronene verb forms have been presented as having the same valency, but differing in transitivity with regard to other factors. And the semantic difference is not merely a change of pivot, but involves a cluster of more subtle distinctions.

This paper represents a preliminary exploration of the semantics of the Moronene verbal prefix *moN-*. Various aspects of its meaning have been illustrated and explicated in terms various extended meanings and analogies. Further work is needed to refine the analysis, both in terms of examining a larger number of examples and in understanding how the various meanings interrelate to form a polysemy network.

It is likely that the semantics of similar verb prefixes in other languages in Sulawesi is different from that in Moronene. But investigation along the lines attempted in this paper may possibly reveal somewhat similar patterns of variation in meanings.

References.

- Andersen, T. David, 2001, Discourse functions of clause types in Moronene narrative poetry. PhD tutorial paper. Fuller Theological Seminary.
- Barr, Donald, 1988a, Da’a verbal affixes and clitics. In Steinhauer 1988:11–49.
- 1988b, Focus and mood in Da’a discourse. In Steinhauer 1988:77–129.
- Busenitz, Robert L., 1994, Marking focus in Balantak. In van den Berg 1994:1–15.
- Esser, S.J., 1927, *Klank- en vormleer van het Morisch* (Phonetics and morphology of Mori) part 2. *Verhandelingen van het Koninklijk Bataviaasch Genootschap van Kunsten en Wetenschappen* 67/4. Bandoeng: Nix
- Foley, W.A. and R.D. van Valin, 1984, *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.
- Forrest, Linda B., 1994, The de-transitive clauses in Bella Coola: passive versus inverse. In Talmy Givón, ed. 1994:147–168.

- Friberg, Barbara, 1991, Ergativity, focus and verb morphology in several South Sulawesi languages. In Ray Harlow, ed. *VICAL 2, Western Austronesian and contact languages, papers from the Fifth International Conference on Austronesian Linguistics*, 103–130. Auckland: Linguistic Society of New Zealand.
- 1996, Konjo's peripatetic person markers. In Steinhauer 1996:137–171.
- Givón, Talmy, 1990, *Syntax: a functional-typological introduction*, vol.2. Amsterdam/Philadelphia: John Benjamins.
- 1994, The pragmatics of de-transitive voice: Functional and typological aspects of inversion. In Talmy Givón, ed. 1994:3–44.
- Givón, Talmy, ed., 1994, *Voice and inversion*. Amsterdam: John Benjamins.
- Himmelman, Nikolaus P., 1996, Person-marking and grammatical relations in Sulawesi. In Steinhauer, 1996:115–136.
- 2002, Voice in western Austronesian: an update. In Fay Wouk and Malcolm Ross, eds *The history and typology of western Austronesian voice systems*, 7–16. Canberra: Pacific Linguistics.
- Hopper, P.J. and S.A. Thompson, 1980, Transitivity in grammar and discourse. *Language* 56:251–299.
- Kroeger, Paul R., 1996, The morphology of affectedness in Kimaragang Dusun. In Steinhauer 1996:33–50.
- Langacker, Ronald W., 1987, *Foundations of cognitive grammar*, vol. 1: *theoretical prerequisites*. Stanford: Stanford University Press.
- 1991a, *Foundations of cognitive grammar*, vol. 2: *descriptive application*. Stanford: Stanford University Press.
- 1991b, *Concept, image, and symbol: the cognitive basis of grammar*. Berlin and New York: Mouton de Gruyter.
- Lyons, John, 1977, *Semantics*. Cambridge: Cambridge University Press.
- Martens, Michael, 1988, Notes on Uma verbs. In Steinhauer 1988:167–237.
- Matti, David F., 1994, Mamasa pronoun sets. In van den Berg 1994:65–89.
- Mead, David, 1998, Proto-Bungku-Tolaki: reconstruction of its phonology and aspects of its morphosyntax. Doctoral dissertation, Rice University.
- Steinhauer, Hein, ed., 1988, *Papers in western Austronesian linguistics* No 4, 11–49. Canberra: Pacific Linguistics.
- Steinhauer, Hein, ed., 1996, *Papers in Austronesian linguistics* No. 3. Canberra: Pacific Linguistics.
- Strömme, Kari K., 1994, Person marking in the Mamuju language. In van den Berg 1994:91–113.
- van den Berg, René, ed., 1994. *Studies in Sulawesi Linguistics, Part 3, NUSA* 36.