



UNIVERSITY OF SYDNEY

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF

BACHELOR OF COMMERCE  
AND BACHELOR OF ARTS (HONOURS)

---

# Syntactic Derivations of Samoan Predicates

---

*Author:*  
James COLLINS

*Supervisor:*  
Prof. William FOLEY

1 November 2010

## **Acknowledgements**

I am indebted to the staff and students of the Department of Linguistics at the University of Sydney for their help during the course of 2010. In particular, my thanks to Professor William Foley, my supervisor, for his guidance and careful criticism during the writing of this thesis. I also wish to thank Dr. Jane Simpson and Dr. Toni Borowsky for their invaluable assistance and advice.

I am very grateful for the encouragement and support given to me by my fellow Honours students; Bethany Stanford, Yaegan Doran and Matthew Carroll.

By the nature of the work, the participation of Samoan speakers was vital. I wish to express my gratitude for the time and effort contributed by Samoan speakers in Australia and Samoa, including Sia Atonio, Tasi Leone, Rev. Iakopo Leleimalefaga, Bonnie Tolealoga, Freddie So'oalo, Leo Tanoi, Fitz Taua'aletoa, Leni Nauer, Poi Giglow and Luisa Mulinu'u. I reserve special thanks to Bob and Aniva Slater for their hospitality during my research in Samoa.

Finally, the research I have undertaken this year would not have been possible without the generous financial support of the Australian Linguistics Society and the Gerhardt Laves Scholarship.

## Table of Contents

1	Introduction	1
1.1	A Note on Objectives	2
2	Theoretical Preliminaries	
2.1	Introduction	3
2.2	Utterance Construction in Minimalism	3
2.3	Category Labels in Samoan	6
2.4	Theta Roles and Case	10
2.5	Conflation	14
2.6	Conclusion	17
3	Ergative Predicates	
3.1	Introduction	18
3.2	Modelling the Clause with Binary Branching	19
3.3	Theta Assigning Properties of Ergatives	40
3.4	Conclusion	54
4	Non-Ergative Predicates	
4.1	Introduction	55
4.2	Syntactic Behaviour of Non-Ergatives	55
4.3	Conclusion	74
5	The Samoan Case System	
5.1	Introduction	77
5.2	The Case System Defined	77
5.3	Formalising the Case System	79
5.4	A Note on Clitics	85
5.5	Conclusion	87

6	Transitivity Alternations	
5.1	Introduction	89
5.2	Causatives	89
5.3	The -Cia Suffix	104
5.4	Conclusion	112
7	Conclusion	113
	References	116

# 1 Introduction

This thesis gives a formal syntactic account of Samoan verbal classes. Samoan verbs may be divided into classes based on their observable syntactic behaviour (for example, case assignment, incorporation) or on their semantic properties (event structure, theta role assignment). The analysis aims to characterise these differences in terms of simple, lexically specified parameters. My objectives here are primarily theoretical, as opposed to descriptive. I intend to test the validity of certain linguistic assumptions using Samoan examples. My argument is informed by research conducted with Samoan speakers living in both Australia and Samoa.

A theoretical background for these propositions is outlined in Chapter 2. Many basic assumptions as to the architecture of the language faculty are derived from Chomskyan theories, which state that the syntax of a language is a simple program which generates sentences. Primarily, I appeal to Chomsky's Minimalist framework (Chomsky 1995, 2000, 2001, 2004), discussing in Chapter 2 the Minimalist hypotheses which are central to my analysis of Samoan. Other linguistic theoretical models are also examined for their applicability to the Samoan language.

Chapters 3 to 6 comprise my analysis of Samoan argument structure. In Chapter 3 I propose a structure modelling the basic Samoan clause and discuss the syntactic properties of transitive verbs. In Chapter 4, I extend the basic clause structure to account for the behaviour of intransitive verbs. I also discuss the grammatical distinction between the stative and dynamic aspects in Samoan. In Chapter 5, I discuss the Samoan case system which aligns to an Ergative-Absolutive pattern. The applicability of this system to an Optimality Theory analysis is also considered. The final chapter considers two morphological derivations of Samoan verbs and the appropriateness of the Lexicalist and syntactic models of morphology to each are analysed.

## 1.1 A Note on Objectives

My analysis of Samoan syntax must adhere to the following objectives, quoted from (Pylkkänen 2002 : 9)

- (1) (a) Define the nature of the primitive building blocks that enter into linguistic computation.
- (b) Characterize the manner in which the basic units combine into complex representations.

These objectives having been defined, they are carried out strictly with the intention of building a grammatical model. The analysis presented in the following sections does not present itself as a genuine picture of the actual, on-line linguistic computation of the Samoan speaker.

The analysis serves to respond or contribute to certain questions raised by theories of linguistics. To what extent can the success and failure of certain verbs to appear in certain grammatical constructions be attributed to purely structural relations, and to what extent can they be captured by other components of the human language faculty (most specifically, lexical semantics).

By its nature, this thesis does not attempt to be a comprehensive analysis of the language. But rather several aspects of interest uncovered during my field work and the applicability of various theoretical models to them are examined. The results of this analysis form a promising basis for future, more comprehensive study of the language.

## 2 Theoretical Preliminaries

### 2.1 Introduction

The following section describes the sentence building operations under a version of Minimalism, the grammatical model derived from hypotheses made by Chomsky (1995, 2000, 2001, 2004). The model here asserts the lexicon is a set of morphemes, each of which is "an articulated set of features" (Chomsky 1995 : 130). The features ascribe to each morpheme information concerning (at least) their phonological, semantic and syntactic characteristics.

Crucial to this analysis is the notion that *argument structure* is a lexically specified property of each morpheme. A morpheme's argument structure determines the syntactic configuration which holds between the morpheme and the other syntactic items it selects. This notion is expressed formally in the sections below.

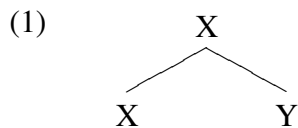
### 2.2 Utterance Construction in Minimalism

One of the fundamental Minimalist hypotheses is the simplification of the syntax module. Specifically, the syntax module is a simple operation which builds structures from lexical items. Further, the lexicon consists only of morphemes. Therefore the morpheme is the relevant "building block" for syntactic structures. The formation of morphologically complex words must occur syntactically (following Baker (1988 : 49-50)).

For each morpheme, a speaker must store at least three categories of lexical information, its pronunciation (or *Phonological Signature*), its meaning (or *Semantic Signature*) and information determining its behaviour with respect to other lexical items in an utterance (its *Argument Structure*) (Chomsky 1995 : 54-55).

In the Minimalist model, the argument structure of a morpheme determines its *category* and *combinatorial* information. Each morpheme must have a *category label* (corresponding to traditional notions of *parts of speech* such as Noun, Verb, Determiner). Further, a morpheme's argument structure determines with which other morphemes it must necessarily combine. If the argument structure determines that a morpheme has the category label X and must combine with a morpheme with the category label Y, X *selects* Y and Y is an *argument* of X. Category and combinatorial information determines the legal and illegal configurations of morphemes in the formation of an utterance.

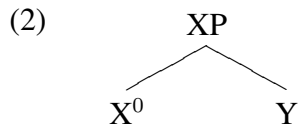
In Minimalism, syntactic objects larger than a morpheme are represented graphically by a binary branching tree. All the nodes of the tree are labelled. A parent node takes its label from one of its children. The child which selects its sibling (by the requirements of its argument structure) will pass its label to its parent. In the structure (1), the node X selects Y, and therefore the parent of X and Y takes the label X.



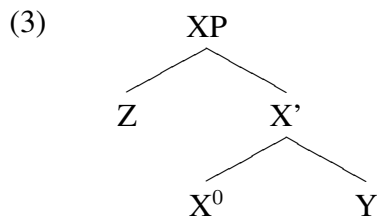
When a morpheme forms part of a syntactic structure, it occupies a terminal node of the tree. A terminal node occupied by a morpheme with the category label X is labelled  $X^0$ . All terminal nodes are *heads*. If the morpheme X selects Y as its first argument, Y is a sibling to  $X^0$ . As X selects Y, the parent of Y of  $X^0$  takes the label X. If X selects more arguments, higher generations of nodes take the label X. The last ancestor node with the label X (that is, either where X has been selected by another morpheme or the X is the label of the root node of the structure) is labelled XP and is termed the *maximal projection* of X. It is the point where X projects no more arguments.



If Y is a sibling to X<sup>0</sup>, Y is the *complement* of X, and X<sup>0</sup> and Y enter into a head-complement relationship. If X<sup>0</sup> takes Y as its complement and selects no other arguments, XP is represented in (2).

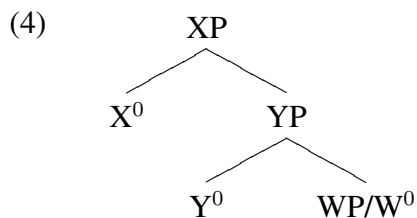


Say, after selecting Y, the morpheme X must combine with a second argument, Z. The parent of X<sup>0</sup> and Y is labelled X' and the item Z is the sister of X'. XP is the mother of X' and Z. The argument which is the sister of the morphemes first projection (X') is the *specifier*. XP is represented in (3), where Z is the specifier of X<sup>0</sup> and Y is its complement. Z and X<sup>0</sup> enter into a head-specifier relationship.



If the morpheme X does not select any arguments, then the terminal node is the maximal projection and the labels X<sup>0</sup> and XP occupy the same terminal node.

For example, let Y in (2) represent the morpheme Y<sup>0</sup> and its complement W (which does not select any arguments). (2) is re-drawn in (4)



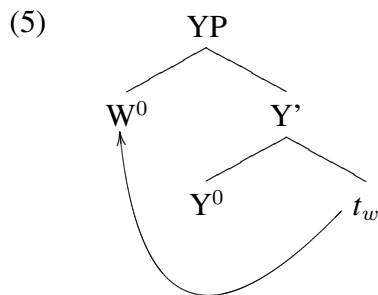
Following Hale & Keyser (2002 : 1), a maximum of two arguments are projected by each morpheme. That is, a morpheme may select a complement, a

specifier, both or neither.

The objective of Chomsky's Minimalist model is to reduce the sentence building component of the language faculty to a single operation, one that takes two items from the lexicon (say X and Y) and replaces them with a new, combined syntactic object, [X Y] (Chomsky, 2000 : 133). This operation is termed *Merge*. It corresponds structurally to a parent node with two child nodes, thus a tree generated by this operation is binary branching.

Note that Chomskyan theories of syntax (1981, 1995) are by nature derivational. A morpheme can change its position in the tree structure after being Merged. This process is name  $Move_{\alpha}$ , Internal Merge or simply movement by various Chomskyan theorists.

Say  $W^0$  is generated as the complement of  $Y^0$ . It then moves to occupy the specifier position of  $Y^0$ . The original position of  $W^0$  is marked with "*t*", symbolising a trace element.



### 2.3 Category Labels in Samoan

The Samoan lexicon may be divided into two categories of morphemes (see Mosel & Hovdhaugen, 1992 : 71) based on a simple syntactic principles.

Samoan contains a small number of *tense particles*, which mark the tense of the clause. One type of morpheme can cluster with a tense particle (and “head” the clause, using the terminology of Mosel & Hovdhaugen (1992)), the other type cannot.

The latter category of morphemes (which cannot head a clause) is the category of *particles*. It can be further divided into a large number of categories based on semantics and syntactic distribution, including tense particles, determiners, adverbs, prepositions, interrogative particles and many others.

The former category of morphemes (which can head a clause) are *full words*, using the terminology of Mosel & Hovdhaugen (1992 : 71-72). They state that the category accounts for the vast majority of the Samoan lexicon. I agree with Mosel & Hovdhaugen (1992 : 75-79) who state that the Samoan language lacks a defined Noun-Verb distinction, thereby requiring a single category which is neither verbal nor nominal.

Mosel & Hovdhaugen generalise that full words may serve as clausal heads, marked for tense, aspect and mood (behaving as “verbs” in a traditional sense) *and* they may serve as arguments of the clause, clustering with determiners and case markers (“nouns” in a traditional sense). To capture the dual function of these words, I will refrain from the traditional labels V and N and use a global label R, for “root word”.

Compare the usage of the lexical item *fafine* in (6), where it has a verbal function, and (7), where it has a nominal function.

- (6) ‘o le fōma‘i e fafine  
TOP the doctor PRES woman  
‘The doctor is a woman.’

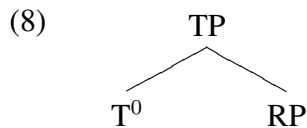
- (7) *e* 'umi le *fafine*  
 PRES tall the woman  
 'The woman is tall.'

A root word's usage as nominal or verbal is determined by the particles with which it clusters. The division of the Samoan lexicon into "root words" and "particles" aligns well with Chomsky's (1995 : 54) division of a language's morphemes into two categories, items which are "substantive" in semantic content, and items which are not substantive.

Particles exist to modify root words, semantically and/or syntactically. The category label of the particle determines whether the root word it modifies is used verbally or nominally.

The verbal usage of *fafine* in (6) is signalled by its clustering with the Tense particle *e*, marking present or habitual tense. Particles of this type mark the tense, aspect and mood of a clause and have the category label T. These particles also serve as complementisers, introducing subordinate clauses.

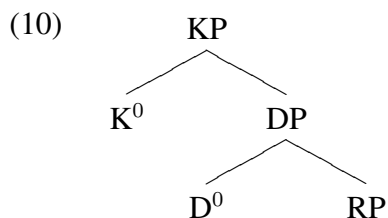
Structurally, the tense particle ( $T^0$ ) selects the root word (R) as a complement, as in (8).



Certain other particles determine that the root word is being used nominally. They are determiners (labelled D) and case heads (labelled K). Determiners mark specificity and the grammatical number of a root word. Case heads signal the theta role of the root word. In example (9), the word *togālā'au*, "garden", is marked with the determiner *le* to signal it is specific and singular, and the case marker *i* to signal it was the location of the event. The cluster  $K^0$ - $D^0$ - $R^0$  is a nominal expression, forming an argument of the clause.

- (9) *sā siva le teine i le togālā‘au*  
 PAST dance  $D_{Spec}$  girl  $K_{Loc}$   $D_{Spec}$  garden  
 ‘The girl danced in the garden.’

The  $K^0$ - $D^0$ -RP cluster (a nominal argument) is represented structurally in (10).



Note that *le teine* in (9) is a  $D^0$ - $R^0$  cluster, without a  $K^0$ , but is still an argument of *siva*. Evidence suggests that arguments of this type (Absolutive arguments) include a non-overt case marker.

Samoan speakers allow any Absolutive argument to be marked with the case marker *o*. The example in (9) may be paraphrased as in (11). Although the unmarked form was far more common in data, speakers allowed any Absolutive argument to be marked with *o*, seemingly without restriction.

- (11) *sā siva o le teine i le togālā‘au*  
 IMP dance  $K_{Abs}$   $D_{Spec}$  girl  $K_{Loc}$   $D_{Spec}$  garden  
 ‘The girl was dancing in the garden.’

Chapin (1970 : 367) notes that pronominals which are Absolutive (and not clitics) must be marked with the case marker *o*. Further, Yu (2008) proposes that where an Absolutive argument is not marked with *o*, as in (9), it is accompanied by a high boundary tone.

Throughout the thesis I will assume that the Absolutive argument is a KP, rather than a DP.

## 2.4 Theta Roles and Case

Having laid out a model of argument structure, where morphemes are heads which select other syntactic items, we can examine further how argument structure interacts with semantics. Does a morpheme's lexically specified "meaning" (its S-signature) significantly influence the configurations and categories of the morpheme's arguments? example Levin & Rappaport Hovav (2005)

Considering this question, I will examine specifically the semantic relations which hold between predicates and their arguments, that is, thematic roles, or theta roles. Theta roles specify the participation of an argument in the denotation of the clause.

Take the Samoan sentence in (12).

- (12) *na sasa e le teine le maile*  
PAST smack  $K_{Erg}$  D girl D dog  
'The girl smacked the dog.'

There are three root words, the predicate *sasa* and its arguments *teine* and *maile*. Payne (1989) stresses that theta roles are relations which exist between "real" world events and participants. Just as the concepts "dog", "girl" and "hitting" exist in the real world and must be represented by linguistic symbols, the relation between the argument "dog" and the predicate "hitting" (one of reception) and between "girl" and "hitting" (one of causation), must be represented in the clause by linguistic symbols.

Many linguists propose there is a finite set of theta roles. These theta roles include, for example, AGENT (the volitional *causer* of an event), GOAL (the *final* locus of an event), SOURCE (the *initial* locus of an event), and so on. Theta roles by this model are often unable to be analyzed in component parts (see Levin & Rappaport, 2005 : 35-48).

In contrast, Hale & Keyser (1993 : 65-67) state that that theta roles are individually specified by each predicate and thus, not drawn from a finite set. That is to say that the predicate in (12), *sasa*, assigns the role of "object being hit" to its Absolutive argument, while the predicate *fufulu*, "to wash", assigns the role of "object being washed", and *vali*, "to paint", assigns the role "object being painted", and so on.

The theta role of an argument is by this model specifically determined by the predicate which selects the argument. Theta roles are therefore lexically specified by each predicate and should be represented in the lexical entry of the predicate. As theta roles are components of a predicate's meaning, it follows that they are encoded into the predicate's S-signature (its semantic component).

Say the predicate in (13), *si'o* - "surround", selects two arguments, an "object being surrounded" and "an object which surrounds".

- (13) *e si'o le fale 'i ni niu*  
 PRES surround the house  $K_{Obl}$  some coconut tree  
 'The house is surrounded by coconut trees.'

The S-signature of *si'o* is such that it encodes a state, "surround", which is a two-place predicate, in that it expresses a relation between two entities. The S-signature of *si'o* is represented in (14), using English as a metalanguage.

- (14) **si'o**  
 S-signature:  $x$  surrounds  $y$

The S-signature encodes a concept, represented by the English word "surrounds" in (14), a concept which I will take as an unanalysable unit following (Fodor 1998). The concept requires predication of two items,  $x$ , "the object which surrounds" and  $y$ , "the object surrounded".  $x$  and  $y$  are variables which must be assigned semantic values by other lexical items in the syntactic tree (in Samoan,

by root words accompanied by determiners and case markers).

In terms of sentence structure, when the root *si'o* appears in a clause, it must necessarily appear with two arguments. This syntactic feature of *si'o* is represented in its argument structure. It selects the two arguments in its complement and specifier projections. Further, we can observe that these two arguments fulfill the roles defined by x and y in (14). The theta roles represented by x and y in (14) are mapped onto the structural projections defined *si'o*'s argument structure.

Following (Hale & Keyser 1993), the syntactic object which Merges into a position projected by a predicate assumes the theta role associated with that position.

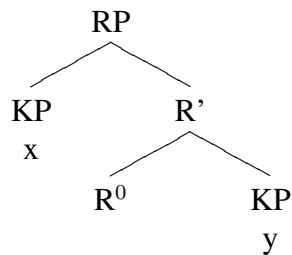
Say the lexical entry of *si'o* is structured as in (15).

(15) **si'o**

P-signature: /si'o/

S-signature: x surrounds y

Argument Structure:



The case system of a language marks arguments with a finite set of symbols. The symbol signals the argument's theta role. In Samoan, case is marked by the  $K^0$  (case marker) in a nominal expression. The case markers are drawn from a small finite set, yet their purpose is to indicate a potentially boundless set of theta roles. Some of these are outlined in the table in (16). Many case markers have multiple uses.



(16)

e	<i>ergative</i>
'o	<i>topic</i>
o	<i>absolutive, genitive</i>
∅	<i>absolutive</i>
i	<i>locative</i>
'i	<i>allative, instrumental</i>
a	<i>agentive genitive</i>

The theta roles specified by one predicate could potentially be represented by almost all of the cases in (16), depending on syntactic considerations, for example, whether the root is verbal or nominal or whether an argument is topicalised or incorporated. Below are some examples of different constructions of the root 'ai, meaning "to eat", with different cases assigned to its arguments. In each example, *teine* takes the role of "eater", and *i'a*, "object being eaten".

(17) *na 'ai e le teine le i'a*  
PAST eat  $K_{Erg}$  the girl the fish  
'The girl ate the fish.'

(18) *na 'ai le teine i le i'a*  
PAST eat the girl  $K_{Obl}$  the fish  
'The girl ate some of the fish.'

(19) *e 'ai i'a le teine*  
PRES eat fish the girl  
'The girl eats fish.'

(20) *le 'ai o le i'a e le teine*  
the eat  $K_{Gen}$  D fish  $K_{Erg}$  D girl  
'The eating of the fish by the girl.'

(21) *le 'ai a le teine*  
the eat  $K_{AGen}$  the girl  
'The eating of the girl.'

Chomsky (1980) asserts that the rules which govern the distribution of phonological case markers on core arguments (as in the Samoan system), also govern the structural positioning of arguments. He proposes the Case Filter, whereby, all nominal arguments must have case (if they are phonological). This requirement is satisfied by each argument occupying a position in the tree which licenses case assignment.

## 2.5 Conflation

This model of syntax proposes a lexicon which is a set of morphemes and nothing larger. Morphemes are the only units which enter into syntactic operations to form larger items. Therefore, syntactic operations proposed by non-Minimalist theories which form larger items *outside* of sentence construction are prohibited. These operations include Word Formation Rules, that is, the building of morphologically complex words before inserting the word into a syntactic structure.

Samoan contains a wealth of morphologically complex words. For example, the word *tosoga*, meaning ‘pulling’ or ‘tug-of-war’, is composed of two morphemes, the free morpheme *toso*, meaning ‘to pull’, and the bound morpheme *ga*, a generic nominal suffix.

By the model detailed in this chapter, both *ga* and *toso* must enter into the syntactic structure separately. The question arises as to how the morphologically complex word *tosoga* is derived in the Phonological Form.

Hale & Keyser (2002) propose the model termed Conflation, which is itself a revision of Baker’s (1988) model of Incorporation and a borrowing of a term introduced by Talmy (1985).

The lexical entry of all morphemes specifies a phonological signature (P-

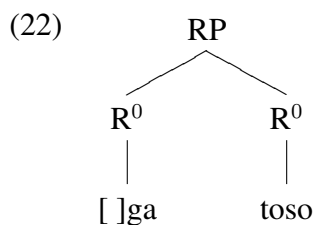
signature) which governs the morpheme's pronunciation. Hale & Keyser (2002 : 60-69) assert that the morphemes in a language's lexicon are either phonologically "defective" or phonologically "substantial". Phonologically defective and substantial morphemes correspond respectively to bound and free morphemes in traditional linguistics.

If a morpheme is phonologically defective, it includes a null phonological feature. Its P-signature is either entirely null or partially null. Partially null P-signatures include a null element along with other non-null features.

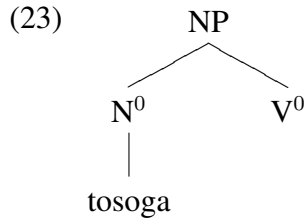
Null phonological features must be eliminated to derive a valid Phonological Form. If a phonologically substantial morpheme Merges with a defective morpheme, the P-signature of the substantial morpheme will replace the null element of the defective morpheme. This replacement is the process Conflation.

Let the Samoan bound morpheme *ga*, in *tosoga*, be a phonologically defective morpheme, whose P-signature consists of a phonological element "ga" and a null element. Its P-signature may be represented as something like "[ ]ga". The free morpheme *toso* is phonologically substantial. When the two morphemes are combined in sentential syntax by Merge the P-signature of the free morpheme, "toso" replaces the null element of "[ ]ga", forming the morphologically complex word "tosoga".

The two morphemes *toso* and *-ga* both take the category label R. When they Merge, they form the structure in (22)



By the automatic process of Conflation, the substantial P-signature of *tosog* replaces the null element of *ga*, forming the complex word *tosoga*. The P-signature of the phonologically substantial morpheme is deleted at the source.

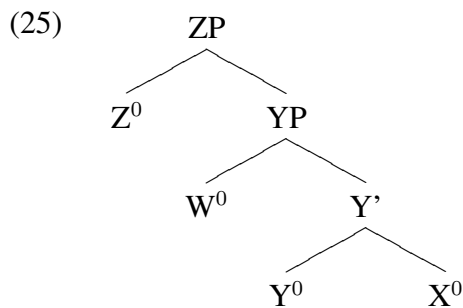


Conflation is crucially a relation which holds between heads and complements. A substantial head may conflate with a defective selecting head if and only if the substantial morpheme is the *strict complement* of the defective head. Strict complementation is defined by Hale & Keyser (2002 : 59) as in (24)

(24) STRICT COMPLEMENTATION

A head X is the strict complement of a head Y iff Y is in a mutual ... sister relation with the maximal categorial projection of X

Observe the toy structure in (25), where  $W^0$  is a specifier of  $Y^0$ .



If  $X^0$  (which is non-branching) is phonologically substantial, it may conflate with  $Y^0$  as it is the strict complement of  $Y^0$  (its maximal projection  $XP$  is the sister of  $Y^0$ ). Likewise, if  $Y^0$  is phonologically substantial, it may conflate with  $Z^0$ ,

as YP is the sister of  $Z^0$ .  $W^0$  may not conflate with any other head in (25) as its maximal projection is not a sister of any head. This model states in simple terms that conflation does not interact with Specifiers.

## **2.6 Conclusion**

Having discussed these theoretical preliminaries, I propose to test their applicability to various syntactic processes of the Samoan language. These assumptions are fundamentally disputed by conflicting theories of linguistics. Where relevant, I will also test the applicability of other theories at different times. Otherwise, I implicitly assume the notions of the models proposed above.

## 3 Ergative Predicates

### 3.1 Introduction

In this chapter, I use the framework defined in Chapter 2 to build an analysis of Samoan roots. I will analyse specifically the class of roots termed *Ergatives*. These are roots which can freely appear with an Ergative argument. In the next chapter I will focus on Non-Ergative roots. These are roots which may only appear with an Ergative argument if they assume additional morphology.

At this point, I will make clear the distinction between the terms “root” and “predicate”. I use root to refer to the entire class of morphemes in Samoan which fall under the category R. In Chapter 2, I stated that morphemes belonging to the category R may head a clause, taking tense. They also may serve as arguments of the clause, taking articles and/or case markers. In the former instance they are used predicatively. Therefore a predicate is a root R which clusters with a tense morpheme.

The root *vali*, “to paint” falls into the Ergative class. It assigns the theta role of ‘object being painted to an argument. This argument most often appears in Absolutive case (but other options which I will examine later are available). The example in (1) demonstrates a predicative use of *vali* with an Absolutive argument.

- (1) *na vali le fale*  
PAST paint the house  
‘The house was painted.’

This clause may freely include an Ergative argument. An Ergative Argument is always the Agent of the predicate. It takes the case marker *e*. The example in (2) demonstrates (2) with an explicit agent encoded by an Ergative argument.

- (2) *na vali e le tamaloa le fale*  
PAST paint  $K_{Erg}$  the man the house

‘The house was painted by the man.’ or ‘The man painted the house.

This chapter examines in detail the grammatical behaviour of the Ergative class of roots, making generalisations and predications about the Samoan case system and the assignation of theta roles to arguments. The analysis requires exploration into the interaction of two syntactic sub-systems, the *case system* and the *theta system*.

### 3.2 Modelling the Clause with Binary Branching

In this section, I will consider the Samoan clause and how it is represented structurally under Minimalist assumptions. In particular it questions the assumption of binary branching. How useful is a binary branching model in describing a verb initial language like Samoan? What are the inefficiencies of the model? I will consider weaknesses of the theory, particularly in relation to the artifices which are required to generate the correct word order. I will also consider advantages of the theory, detailing the correct predictions made by the structure.

The basic Samoan clause can be modelled with the structure in (3)

(3) NUCLEUS + [ARG<sub>1</sub>, ARG<sub>2</sub>,...ARG<sub>n</sub>]

The object labelled NUCLEUS is itself complex, consisting of a tense particle (a morpheme denoting the tense, aspect and mood of the clause), a predicate (denoting an event or state) and any pronominal clitics. The nucleus is schematically represented in (4)

(4) NUCLEUS → [TENSE-PARTICLE-(*Clitic*) PREDICATE-(*Clitic*)]

Clitics are adjoined to the right of the tense marker or right of the predicate, depending on their thematic role. Clitics denoting the Agents of transitive clauses appear to the right of the tense marker. Clitics serving as the sole core argument of intransitive clauses similarly adjoin to the right of the tense marker. In this respect,

the pronominal clitic system follows a Nominative-Accusative case marking pattern. Other particles, such as the negative particle and adverbs appear between the pronominal clitic and predicate.

The system is demonstrated in (5-6).

(5) *na 'ou alu*  
 PAST 1.SG go  
 'I went.'

(6) *na 'ou sasa le teine*  
 PAST 1.SG hit the girl  
 'I hit the girl.'

Arguments assigned Dative, Instrumental, Locative and other Oblique cases adjoin to the right of the predicate (see Chapin 1970 and Pizzini 1971 for a fuller discussion) .

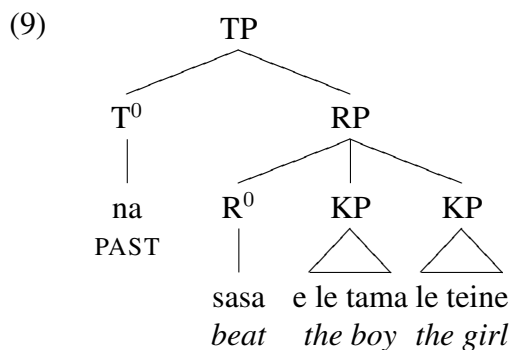
If the predicate is Ergative, it is transitive. It takes two arguments, the Agent-like argument is marked with Ergative case and the Patient-like argument is marked with Absolutive case. Both of the arguments follow the predicate. They may appear in either order, demonstrated in (7) and (8). The ordering in (7) is far more common. Some Samoan speakers in Australia did not allow the ordering in (8), but all speakers in Samoa allowed both (7) and (8). Ochs (1988 : 113) notes that the argument immediately following the predicate receives greater pragmatic focus.

(7) *na sasa e le tama le teine*  
 PAST hit  $K_{Erg}$  the boy the girl  
 'The boy hit the girl.'

(8) *NA sasa le teine e le tama*  
 PAST hit the girl  $K_{Erg}$  the boy  
 'The boy hit the girl.'



Chapin (1970) and Pizzini (1971) use Fillmore's case grammar (1968) to model the Samoan Ergative clause. The predicate and its arguments have a flat, ternary structure. This is permissible as the case grammar model does not assume binary branching. A flat structure for the Samoan clause in (7) is presented in (9)



The structure in (9) is impossible under Minimalist assumptions as it contains ternary branching. However, the structure is theoretically advantageous for at least one reason. It accounts for certain syntactic properties of Samoan reflexive constructions.

Mosel & Hovdhaugen (1992 : 726) note a marked lack of reflexive constructions in Samoan, that is, constructions where the Agent and Patient refer to the same participant and both are expressed with distinct lexical items. They state that concepts which are expressed by reflexive clauses in English are expressed in Samoan by specific, Intransitive lexical items such as *tā'ele*, “to wash oneself”, or *tōa'i*, “to kill oneself”. Otherwise, actions where an Agent performs an act on his or her own body can be expressed by placing the body part in the Patient position.

However, Chapin (1970) gives a description of the properties of Samoan reflexives. Chapin's examples were considered grammatical by speakers. It is important to note, however, that such reflexive constructions were never given as spontaneous descriptions of reflexive action.

Samoan reflexive pronouns are constructed by attaching the particle *lava* to the right of a regular pronoun. Take the example (10), where the reflexive is marked with Absolutive case.

- (10) *sā sogi e Ioane<sub>i</sub> ia-lava<sub>i</sub>*  
 IMP cut  $K_{Erg}$  Ioane 3.SG-self  
 ‘Ioane cut himself.’

Samoan also allows the Ergative argument to be a reflexive pronoun, as in (11).

- (11) *sā sogi Ioane<sub>i</sub> e ia-lava<sub>i</sub>*  
 IMP cut Ioane  $K_{Erg}$  3.SG-self  
 ‘Himself cut Ioane.’

Note that if two arguments have the same denotation (the person “Ioane” in these examples), they are marked with a subscript “*i*”. This is *co-indexation*.

The grammaticality of (11) is problematic for Chomskyan theories of Government and Binding (1981). Chomsky (1981) proposes that the English translation of (11) is ungrammatical because in English, the Agent *binds* the Patient, but the Patient does not *bind* the Agent. Lexical items like the Samoan *ia lava* and the English *himself* are anaphors, and therefore constrained by Principle A, defined by Chomsky (1995).

- (12) PRINCIPLE A:

An anaphor must be bound in its local domain.

The term “local domain” means the smallest clause which the anaphor occupies. As we are only dealing with mono-clausal examples, this concept is not immediately relevant.

The term “binding” is defined by Chomsky (1995) with the following stipulations.

(13) BINDING

A binds B if and only if

- i) A c-commands B
- ii) A and B are co-indexed.

C-command is a structural relation. Informally, a node c-commands its sibling, and all the children and descendants of its sibling. It is given the following formal definition by Chomsky (1986 : 8).

(14) C-COMMAND

A c-commands B if and only if

- i) A does not dominate B
- ii) Every node X which dominates A also dominates B.

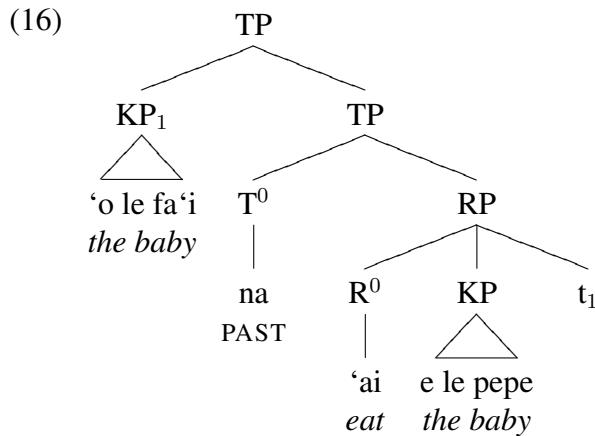
“Domination” in (14) is defined by parenthood. A node dominates its children and their children and so-on. The definition of X in (14) is the subject of some debate. Chomsky (1986, 1995) states it is any branching node, regardless of category.

An anaphor like *ia-lava* must be bound. It is theoretically impossible assuming *Principle A* for *ia-lava* to occupy a position in the tree where it is structurally higher than the argument with which it is co-indexed.

In Samoan, any argument can topicalise. This is achieved by replacing the argument’s case marker with the Topic case marker ‘*o*’. The argument is also “fronted”. It appears in the clause to the left of the tense marker.

- (15) ‘*o le fa’i na ‘ai e le pepe*  
K<sub>Top</sub> the banana PAST eat K<sub>Erg</sub> the baby  
‘As for the banana, the baby ate it.’

As the T<sup>0</sup> is the left most edge of the clause, the topicalised argument must occupy a position above the TP. The clause in (15) is sketched in (16), using the flat structure.



If *ia-lava* is a true anaphor, subject to Principle A, it should not be able to occupy this position. This prediction is accurate, as (17), a paraphrase of the grammatical (10), is ungrammatical.

- (17) \**o ia-lava s̄a sogi e Ioane*  
 $K_{Top}$  3sg-self IMP cut  $K^0$  Ioane  
 ‘As for himself, John cut.’

Likewise, the anaphor *ia-lava* is unable to take its antecedent from within a relative clause, as in (18).

- (18) \**s̄a sogi e le teine e va'ai i Ioane<sub>i</sub> ia-lava<sub>i</sub>*  
 IMP cut  $K_{Erg}$  the girl COMP see  $K_{Obl}$  Ioane 3.SG-self  
 ‘The girl who saw Ioane cut himself.’

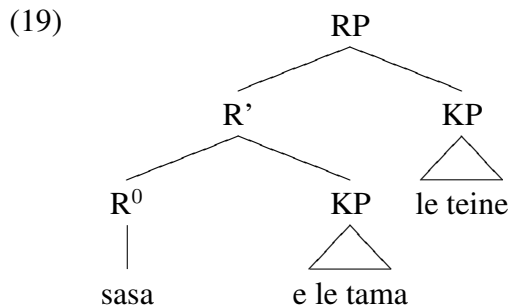
It is therefore reasonable to assume that *ia-lava* adheres to Principle A. Principle A determines that if *ia-lava* is an anaphor, it must be c-commanded by an antecedent. The results where *ia-lava* is clearly not c-commanded by an antecedent are ungrammatical.

It is clear that in the flat structure, the Absolutive argument and the Ergative argument c-command each other. They are sister nodes, therefore neither argument dominates the other. They therefore meet the requirements of mutual c-command

defined in (14). If the two arguments are co-indexed, an anaphor in either position is bound by an antecedent. The flat structure in (9) therefore accurately predicts that either the Ergative or Absolutive may be an anaphor.

How can a binary branching tree capture the facts, satisfying both the requirements of word order and mutual c-command?

Under the model of Hale & Keyser (2002 : 1), a predicate may select a maximum of two arguments, one complement and one specifier. The only possible argument structure which generates the correct word order is displayed in (19), again a structural representation of (7).



This tree fails to capture the mutual c-command between the Ergative and Absolutive arguments. Further, it fails to describe phenomena involving Incorporation.

Incorporation is a process where transitive predicates become intransitive. This is achieved by one argument appearing as a bare nominal, adjoined immediately to the right of the predicate. Incorporated predicates give a habitual, atelic reading of the event. The incorporated object never has a unique or highly specified referent, but is usually generic and plural (see Mithun 1984 and Collins 2010 for my own notes on Samoan incorporation). The incorporated paraphrase of (7) is demonstrated in (20).

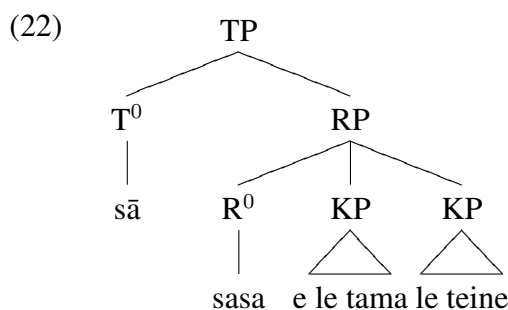
- (20) *e sasa teine le tama*  
 PRES hit girl the boy  
 ‘The boy hits girls (is a girl-hitter).’

In (Collins 2010) I argued for a restriction on Incorporation. The incorporated object must be a complement of the incorporating predicate. Massam (2004) also argues that Incorporation in the Polynesian language Niuean is subject to a complement condition. This hypothesis has theoretical backing in Chomskyan linguistics. Baker (1988) states that as an incorporated object has “moved” into its new position, it must c-command its original position.

If these findings are correct, the Patient (*teine* in (20)) must be generated as a complement of *sasa*. This is a weakness of the binary structure in (19) where *teine* is a specifier. Furthermore, the Agent argument is the complement. This falsely predicts that the Agent argument may incorporate.

- (21) *e sasa teine le tama*  
 PRES hit girl the boy  
 ‘The boy hits girls (is a girl-hitter).’  
 but not ‘\*The girl hits boys (is a boy-hitter).’

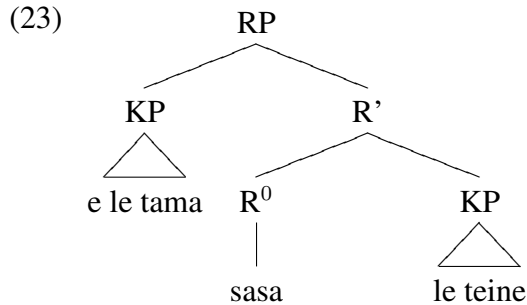
The flat structure in (9), repeated below, places both the Agent and Patient as complements of the predicate.



This structure also falsely predicts that both readings of (21) are grammatical, where the Agent incorporates. As the Agent and Patient are both complements,

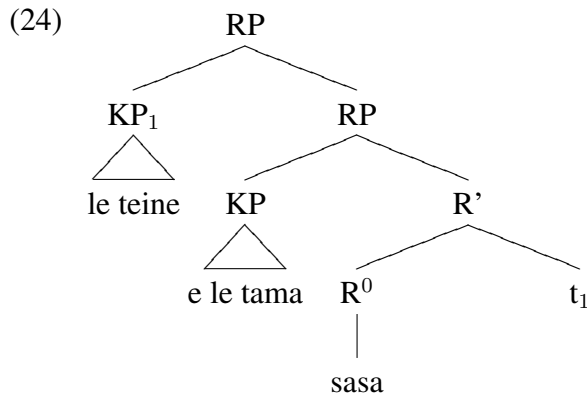
they should have the same capability of incorporating.

A revised version of the binary tree follows in (23). The object is the complement so it may legally incorporate. The Agent is positioned in the specifier.



This structure predicts neither the correct word order nor the mutual c-command relationship between the Agent and Patient. Let us leave aside the (large) question of verb placement and discuss the ordering of the Agent and Patient. This structure places the Agent before the Patient, correctly predicting the unmarked AG-PAT ordering (Agent before Patient) as in (7). How can the more marked, but grammatical, PAT-AG ordering in (8) be generated?

Word order variations in Minimalism are resolved via Scrambling. The alternate ordering where the Agent follows the Patient, as in (8), can be derived via Scrambling. Scrambling is defined by (Boskovic 1998) as simple movement, where an item leaves its original position and adjoins to a higher position. Scrambling can derive the structure in (24).



An analysis with Scrambling makes strong predictions about c-command relationships. The Agent only c-commands the Patient in (23), with no Scrambling. The Patient only c-commands the Agent in (24), after it has Scrambled to a higher position.

This ensures that c-command is determined by word ordering. The item to the left will c-command the item on its right. The Agent c-commands the Patient when the ordering is AG-PAT; the Patient c-commands the Agent when the ordering is PAT-AG. This makes the prediction that the Agent may be a reflexive pronoun if and only if the ordering is PAT-AG. This prediction is accurate.

(25) *sā sogi Ioane<sub>i</sub> e ia-lava<sub>i</sub>*  
 PAST cut Ioane  $K_{Erg}$  3sg-self  
 ‘Himself cut Ioane.’

(26) \**sā sogi e ia-lava<sub>i</sub> Ioane<sub>i</sub>*  
 PAST cut  $K_{Erg}$  3sg-self Ioane  
 ‘Himself cut Ioane.’

Likewise, the Patient may not be a reflexive pronoun if it precedes the Agent.

(27) \**sā sogi ia-lava<sub>i</sub> e Ioane<sub>i</sub>*  
 PAST cut 3SG-self  $K_{Erg}$  Ioane  
 ‘Ioane cut himself.’



The ungrammaticality of (26) and (27) is predicted by the structures in (23) and (24). The Patient only c-commands the Agent if it moves to a higher position, thus appearing to its left in word ordering.

The flat structure in (22) does not make this claim. In this structure, the Ergative and Absolutive arguments are structural sisters and therefore have a symmetrical relationship. This falsely predicts that (26) and (27) are grammatical.

A simple constituency test provides further evidence for a binary structure over a flat structure. The flat structure predicts that the entire RP is a constituent. The sequence [R, AG, PAT] is a constituent and there are no intermediary constituents. That is, neither [R, PAT] nor [R, AG] are constituents. The binary structure predicts that [R, PAT] is an intermediary constituent and [R, AG] is not a constituent.

This is supported by a simple replacement test. The entire predicate of a clause may be replaced by a light verb in a subsequent clause, including the Agent, the object and the predicate. This suggests the sequence [R, PAT, AG] is a constituent.

- (28) *sā tofi e Leni popo i Aso Tofi ma toe fai ananafi*  
 PERF split K Leni coconut K Thursday and again do yesterday  
 ‘Leni split coconuts on Thursday and did so again yesterday.’

The predicate and the object can be specifically targetted for replacement, suggesting the [R, PAT] sequence is a constituent.

- (29) *sā tofi e Leni popo i Aso Tofi ma fai e Tasi ananafi*  
 IMP split K Leni coconut K Thursday and do K Tasi yesterday  
 ‘Leni split coconuts on Thursday and Tasi did so yesterday.’

The predicate and the Agent cannot be specifically targetted for replacement. This suggests that [R, AG] is not a constituent.

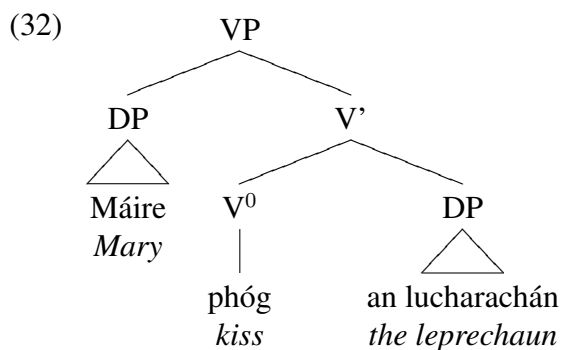
- (30) *\*sā tofi e Leni popo i Aso Tofi ma fai lālā ananafi*  
 IMP split K Leni coconut K Thursday and do branch yesterday  
 ‘\*Leni split coconuts on Thursday and did so branches yesterday.’

### 3.2.1 Verb Initial Word Ordering

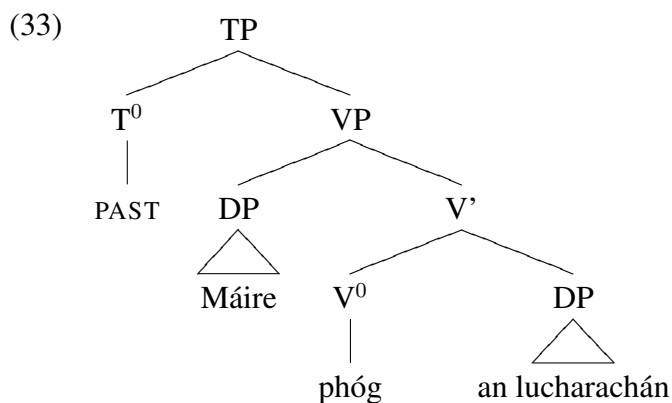
The binary branching model is structurally advantageous with respect to both c-command relationships and constituency. How then can the issue of verb initial ordering be resolved? If the structure in (23) is correct (generating AG-R-PAT word ordering), the R-AG-PAT word ordering of the Samoan clause must be derived somehow. The question of verb initial languages are a topic of consternation among linguists of the Chomskyan tradition. Intuitively, verb initial word ordering is a feature of natural language which binary branching handles poorly.

Carnie & Guilfoyle propose the following solution for the verb initial language Irish (Carnie & Guilfoyle 2000). They suggest a VP for transitive verbs in Irish which resembles my proposition in (23). The underlying structure for the Irish sentence in (31) is modelled in (32).

- (31) *phóg Máire an lucharachán*  
 kiss.PAST Mary the leprechaun  
 ‘Mary kissed the leprechaun.’



When the structure in (32) takes an abstract tense morpheme, the tree in (33) is generated.



They suggest the verb *phóg* moves to the  $T^0$  in order to inherit its tense (in Chomskyan terms, “check its tense features”).

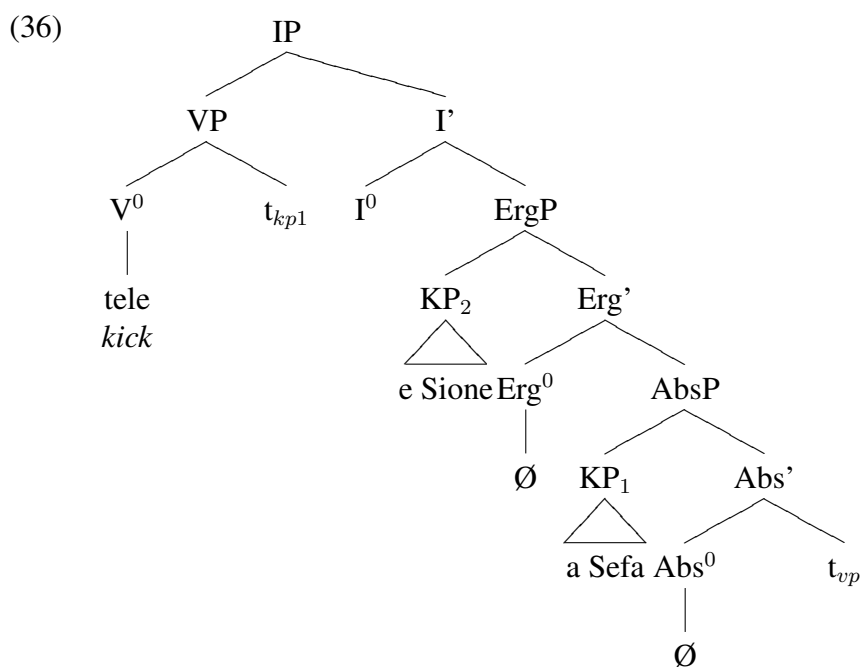
While V-to-T movement might be tenable for Irish, analogous “R-to-T” movement is impossible in Samoan. Intervening negative particles, preverbal adverbs and clitics suggest that the tense particle and predicate are morphologically distinct and cannot possibly occupy the same node. Consider the clause in (34), where multiple items intervene between the tense particle and predicate.

- (34) *sā* ‘*oe lē toe fufulu le ipu*  
 PERF 2SG not again wash D dish  
 ‘You didn’t wash the dish again.’

Massam (2001) analyses the Polynesian language Niuean, which has very similar syntactic properties to Samoan. She suggests that the derivation of verb initial word ordering occurs in three steps.

She suggests that the VP is a [V + PAT] sequence. The Agent argument is generated higher in the structure. The Patient argument moves from out of the VP into the specifier position of a higher functional head (which checks case). After this has occurred, the entire VP moves to the specifier of the functional head I, giving the final structure in (36)

- (35) *ko e tele e Sione a Sefa*  
 PRES kick  $K_{Erg}$  Sione  $K_{Abs}$  Sefa  
 ‘Sione is kicking Sefa.’



There are numerous theoretical problems with this analysis. The first is its complexity, including two movements and numerous null elements expected to be present in every transitive clause. Further, the head  $I^0$  seems to serve no purpose other than to provide a specifier position into which the VP can raise. It also has a theory internal fault. The trace of the Absolutive argument is not governed by the Absolutive argument. This is predicted to be impossible in Minimalism (Chomsky 1995 : 36-37).

Empirical problems also arise when the analysis is applied to Samoan, specifically in regards to adjunct arguments. Consider the clause in (37).

- (37) *e sogi e Leni ufi 'i le naifi*  
 PRES cut  $K_{Erg}$  Leni yam  $K_{Ins}$  the knife  
 ‘Leni cuts the yams with a knife.’

The predicate *sogi* is modified by the Oblique phrase *'i le naifi*. Structurally, the Oblique phrase is an adjunct. It is an optional constituent of the phrase projected by the predicate (the RP). If Massam's (2001) analysis is applied, the entire RP moves to a higher position, along with its projections. This implies that the Oblique argument must move along with the RP, falsely predicting that (37) is ungrammatical.

I will propose an analysis which bears a certain weakness. It assumes the existence of a single functional head, existing to resolve a difficulty in a binary branching system. It nevertheless accounts for the distinct syntactic positioning of the predicate and tense marker (unlike Carnie & Guilfoyle's analysis) and allows arguments to remain in their generated positions (unlike Massam's analysis).

### 3.2.2 Verbalisers and Nominalisers

In Chapter 2 I suggested the lack of a Noun-Verb distinction in Samoan. Morphemes denoting events, states and entities are all given the global label R. This hypothesis requires a few amendments based on the following evidence.

In Chapter 2 I gave the following examples. The root *fafine* is used verbally in (38) as it takes a tense particle. It is used nominally in (39) as it takes an article.

(38) *'o le fōma'i e fafine*  
*K<sub>Top</sub> the doctor PRES woman*  
 'The doctor is a woman.'

(39) *e 'umi le fafine*  
*PRES tall the woman*  
 'The woman is tall.'

Likewise, morphemes which denote events can be used nominally or verbally.

(40) *e 'uma le tapē o le pua'a*  
*PRES finished the kill of the pig*

‘The killing of the pig is finished.’

- (41) *sā tapē e le tama le pua‘a*  
 PRES kill  $K_{Erg}$  the boy D pig  
 ‘The boy killed the pig.’

There are some words which may only be used verbally and never nominally. They are uniformly morphologically complex. The examples below are grouped by their affix.

Firstly there are root morphemes affixed with the *Ornative* suffix, *-a*. This suffix attaches to root words which are semantically nominal, denoting physical objects. It creates a stative predicate, denoting being in ample supply of the object.

(42) **Ornative Suffix**

<i>‘ele‘ele</i>	DIRT, SOIL	<i>‘ele‘ele-a</i>	X IS DIRTY
<i>fale</i>	HOUSE	<i>fale-a</i>	X IS PROVIDED WITH A HOUSE
<i>lōi</i>	ANT	<i>lōi-a</i>	X IS OVERRUN WITH ANTS
<i>namu</i>	MOSQUITO	<i>namu-a</i>	X IS OVERRUN WITH MOSQUITOS

None of the affixed predicates in (42) can take an article and be used nominally. None of the following forms are possible. They are always used verbally.

- (43) \**le ‘ele‘elea* - “the being dirty”  
 \**le falea* - “the being provided with a house”  
 \**le lōia* - “the being overrun with ants”  
 \**le namua* - “the being overrun with mosquitos”

The same principle applies to the *De-ergative* prefix *ma-*. This prefix takes roots which denote inherently dynamic events and creates non-Agentive predicates. The predicate affixed with *ma-* denotes a spontaneously occurring event.

(44) **Inchoative Prefix**

<i>fa'i</i>	x BREAKS y	<i>ma-fa'i</i>	x IS BROKEN
<i>goto</i>	x SINKS	<i>ma-goto</i>	x IS SUNK
<i>pu'e</i>	x RISES, GOES UP	<i>ma-pu'e</i>	x IS SWOLLEN
<i>fasi</i>	x WOUNDS y	<i>ma-fasi</i>	x IS WOUNDED

Again, these forms are unable to appear in DPs.

- (45) \**le mafa'i* - “the broken one, the being broken”  
 \**le magoto* - “the sunken one, the being sunken”  
 \**le mapu'e* - “the swollen one, the being swollen”  
 \**le fasi* - “the wounded one, the being wounded”

Finally, the *Desiderative* prefix *fia-* also inhibits the root from appearing in a DP. Semantically, *fia-* denotes that the event is ‘wanted’ or ‘needed’. The participant who is the ‘wanter’ is usually the Agent.

(46) **Desiderative Prefix**

<i>fufulu</i>	x WASHES y	<i>fia-fufului</i>	x WANTS TO WASH y
<i>tipi</i>	x CUTS y	<i>fia-tipi</i>	x WANTS TO CUT y
<i>sasa</i>	x BEATS y	<i>fia-sasa</i>	x WANTS TO BEAT y
<i>alu</i>	x LEAVES	<i>fia-alu</i>	x WANTS TO LEAVE

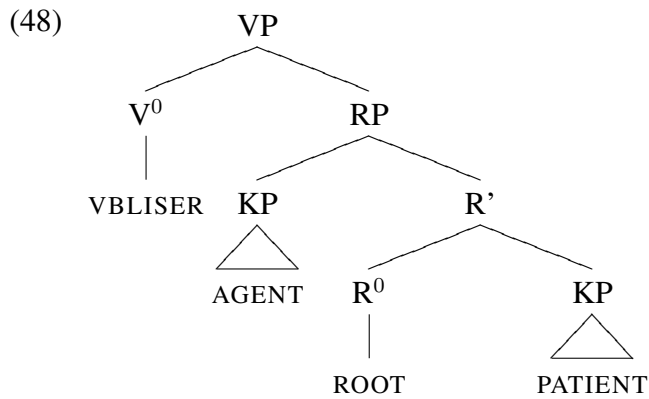
- (47) \**le fia-fufulu* - “the wanting to wash”  
 \**le fia-tipi* - “the wanting to cut”  
 \**le fia-sasa* - “the wanting to beat”  
 \**le fia-alu* - “the wanting to leave”

No word affixed with these morphemes may appear in a DP. This might suggest that the morphemes themselves are intrinsically verbal in category. Roots which are affixed with these morphemes are no longer labile in terms of verbal and nominal usage. They may only take a verbal usage. For this reason I will classify these morphemes (and others) as “verbalisers”. They are symbolised with

the category label V.

In terms of structure, the affixes *-a*, *ma-* and *fia* are bound morphemes. I discussed bound morphemes in Chapter 2, stating that under the model of Hale & Keyser (2002) they are phonologically defective. This means that they must select a phonologically complete head as their complement. The phonologically complete head “Conflates” with the bound morpheme, forming a complex word. The complete morpheme replaces the phonologically null element of the bound morpheme. This process is intended to replace any morphological theory which places word formation outside of syntax.

If conflation is assumed, these morphemes must select the root as their complement. If they select an Ergative root, the structure in (48) is generated (assuming my proposed structure in (23) for Ergative roots).

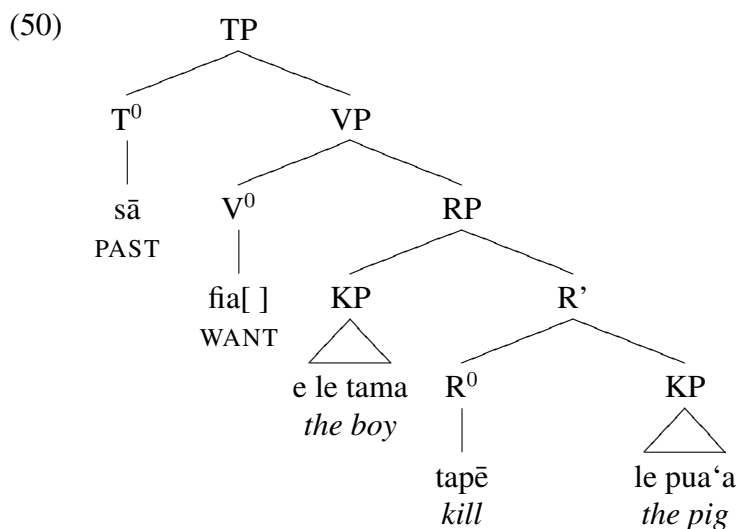


Say the verbaliser *fia-* takes the ergative root *sogi*, giving a desiderative version of the clause in (41).

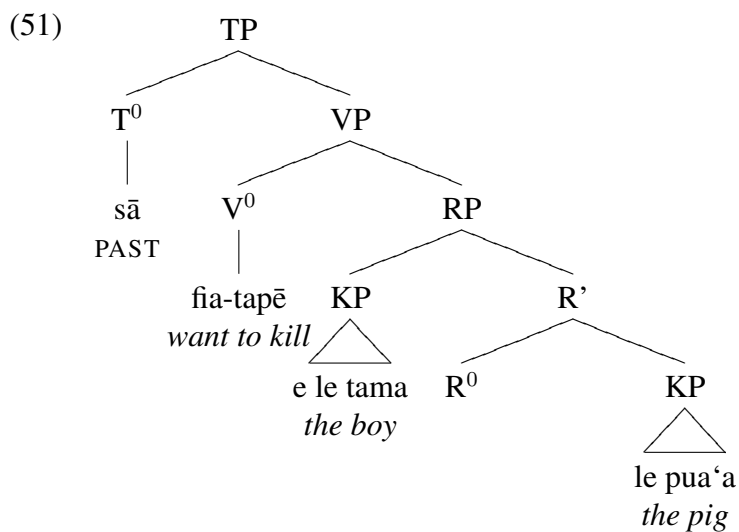
- (49) *sā fia-tapē e le tama le pua'a*  
 PRES want-kill K<sup>0</sup> D boy D pig  
 ‘The boy wanted to kill the pig.’

The clause is represented in (50)



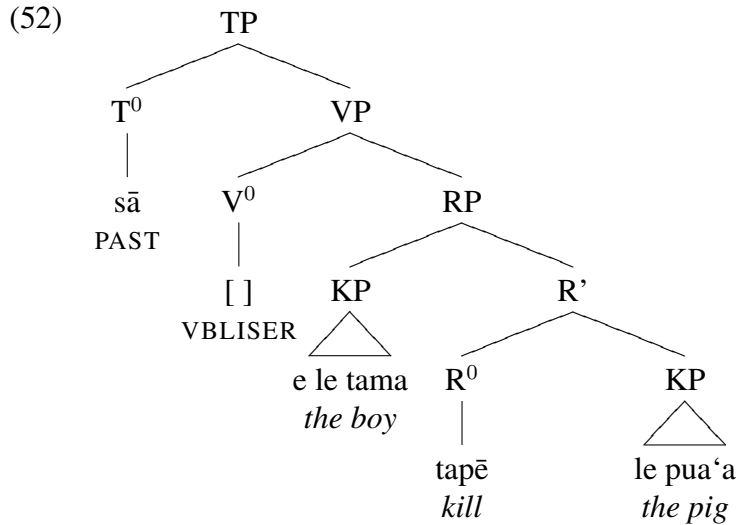


The phonologically complete head *tapē* conflates with the bound morpheme *fia-* to form the structure in (51)



The structure in (51) correctly predicts the word ordering of (49). It also maintains the strengths of the binary structure in (23); it predicts the correct c-commanding relationship between the Agent and Patient, licences the incorporation of the Patient and excludes the incorporation of the Agent.

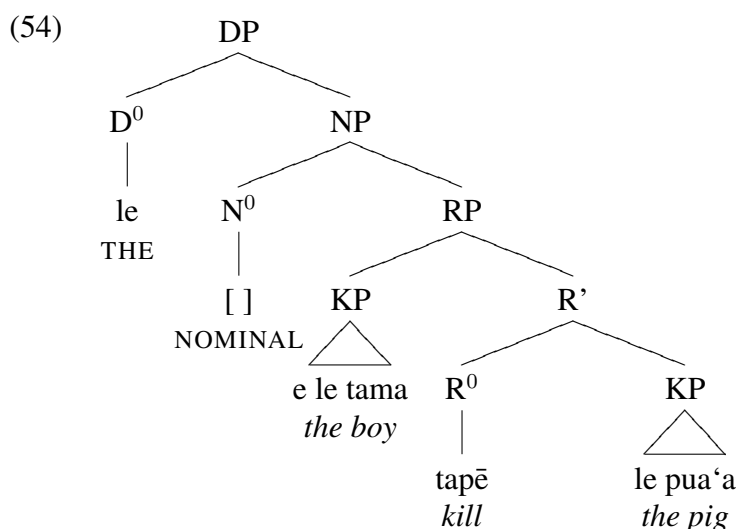
What if the clause structure in (48) is the regular clause structure for all Ergative roots? That is, all Ergative clauses contain a phonologically null verbaliser. The class of verbalisers in Samoan includes the morphemes *-a*, *ma-* and *fia-* and a phonologically null variant. The clause in (41) is represented structurally in (52).



Presumably an analogous hypothesis must be made for roots which are used nominally.

The DP in (53) can be represented in the tree (54). Note that the case assignment in DPs can be identical to the case assignment in TPs. The Agent can take Ergative case and the Patient can take Absolutive case.

- (53) *le tapē e le tama le pua'a*  
 the kill  $K_{Erg}$  the boy the pig  
 'The killing of the pig by the boy.'



The phonologically null verbaliser alternates with phonological verbalisers (such as *-a*, *ma-* and *fia*). The phonologically null nominaliser similarly alternates with phonological nominalisers. Observe the examples below. The table in (55) demonstrates examples with the nominalising *-ga* suffix. Note that although all morphologically simple predicates may be used nominally, those affixed with *-ga* generally derive more time stable concepts than the nominal usage of a root without any morphology.

(55) **-ga Nominaliser**

<i>teu</i>	x DECORATES y	<i>teu-ga</i>	DECORATION
<i>tuli</i>	x HUNTS, CHASES y	<i>tuli-ga</i>	HUNTER, HUNTING PARTY
<i>tu'u</i>	x PUTS, LEAVES y	<i>tu'u-ga</i>	SHARE (of food)
<i>tali</i>	x ANSWERS, RECEIVES y	<i>tali-ga</i>	RECEPTION

This hypothesis states that roots must conflate with a verbaliser or nominaliser. It correctly predicts Samoan verb-initial word ordering in both DPs and TPs. It gives a hypothesis where the verb-initial status of Samoan is a by-product of the root words being ambiguously verbal or nominal. They must be specified as verbal or nominal by a functional head.

The theory circumvents the problems of Carnie & Guilfoyle’s (2000) hypothesis where the root adjoins to  $T^0$ . In Samoan, this theory is untenable as adverbs and negative particles can intervene between the root and tense particle. Adverbs and negative particles can now legally adjoin to the maximal projection VP in (52) and appear in the correct position structurally. It similarly avoids the problems of Massam’s (2001) structure. It significantly reduces her number of movements and null elements. Further, it does not make any false predictions entailed by the movement of the entire RP. Using the same assumptions as Massam (2001), including binary branching, my hypothesis is a significant simplification.

### 3.3 Theta Assigning Properties of Ergatives

#### 3.3.1 The Projection Principle

As I outlined in Chapter 2, I take the theta system to be a syntactic reflection of lexical semantics. That is, a root, like *vali*, bears certain semantic properties. It encodes an event, ‘painting’. It also encodes certain requirements for participants in that event. If an Absolutive argument, like *le fale* in (1-2), is an argument of *vali*, it might inherit the theta role of ‘object being painted’. As stated in Chapter 2, I take this role, ‘object being painted’, to be lexically encoded in the semantics of *vali*. Further, the role is also represented in the argument structure of *vali*. That is, one position projected by the morpheme *vali* will encode the role of ‘object being painted’ to whatever syntact object occupies it.

As I appeal to a Minimalist framework, the analysis of the Samoan theta system is strongly informed by Chomsky’s (1981 : 29) hypothesis stated in (56)

(56) PROJECTION PRINCIPLE

Representations at each syntactic level are projected from the lexicon, in that they observe the subcategorization properties of lexical items.

Say the theta role “object being painted” is a feature F which is specified by the lexical information of some lexical item (or ‘projected’ by a lexical item in the

sense of Chomsky (1981)). Following the Projection Principle in (56), it must be projected at each syntactic level of representation, including the tree structure.

F may be projected by the argument *fale*, ‘house’, as it is this argument which bears the theta role ‘object being painted’. However, if this hypothesis were true, it would contradict the Projection Principle. The feature F must be projected by the lexical item *fale* at each level of representation. F is not projected by *fale* in the Lexicon. There is nothing intrinsic about *fale* which suggests it is lexically specified to bear the theta role of ‘object being painted’. I must therefore conclude that F is projected by the root *vali*. In the derivation of the syntactic structure, *vali* projects F onto *fale*.

### 3.3.2 Defining Semantic Properties of Ergatives

What is the motivation for a distinct class of Ergative roots as opposed to Non-Ergative roots? That is, why do some roots freely take an Ergative argument, while others require additional morphology to take an Ergative argument? I propose that the determining factor rests on the theta roles assigned by Ergative roots. They are lexically specified to select two arguments. One argument is an initiating force of the dynamic event. The event necessarily involves another participant in some way. The relation between the event and the non-Agentive argument is difficult to generalise over the entire class of Ergatives.

The root *fa’i*, ‘to break off’ or ‘to pull out’ is an Ergative root. As such it may freely combine with an Ergative argument.

(57) *na fa’i le lālā*  
 PAST break off the branch  
 ‘The branch was broken off.’

(58) *na fa’i e le tamaloa le lālā*  
 PAST break off  $K_{Erg}$  the man the branch  
 ‘The man broke off the branch.’

The root *gau*, also meaning “to break” is Unaccusative. An Unaccusative root is a type of Non-Ergative root. It assigns the theta role with the approximate value of “object breaking” onto its sole core argument.

- (59) *na gau le lālā*  
 PAST break the branch  
 ‘The branch broke.’

The clause cannot take an Ergative argument unless the root takes the prefix *fa‘a-*.

- (60) \**na gau e le tamaloa le lālā*  
 PAST break  $K_{Erg}$  the man the branch  
 ‘The branch broke by the man.’
- (61) *na fa‘a-gau e le tamaloa le lālā*  
 PAST CAUSE-break  $K_{Erg}$  the man the branch  
 ‘The man broke the branch.’

What then, is the fundamental difference between a predicate like *fa‘i* and a predicate like *gau*? Is it enough to state that the two roots project different structures, accounting for their distinct syntactic behaviour? The question remains as to why they are assigned different structures.

Levin & Rappaport Hovav (2005) state that roots which form predicates can be classified by their *ontological type*. While they give a large set of types, including *state*, *place* and *manner*, they state that the syntactic behaviour of roots cannot be predicted from the ontological type alone (2005 : 73). Hale & Keyser (1993 : 53, 2002 : 31-34) give an example, where two English verbs, *splash* and *smear*, behave differently syntactically, despite having the same ontological type. They both encode events of a surface being annointed with a non-solid substance. The verb *splash* is able to participate in a causative-inchoative alternation, while the verb *smear* is not.

(62) We splashed mud on the wall.

Mud splashed on the wall.

(63) We smeared mud on the wall.

\*Mud smeared on the wall.

Hale & Keyser (2002 : 31-34) state that there is a crucial semantic difference between *splash* and *smear*. The verb *smear* encodes certain particularities about the manner in which the event is perpetrated by an agent. It is a controlled motion by an agent spreading the material onto a surface. It is an *agent-manner* verb. The verb *splash* encodes particularities about how the material is distributed on a surface, from the projection of diffuse, airborne liquid. It is a *patient-manner* verb.

This distinction might characterise the difference between *fa'i* and *gau*. The root *fa'i* encodes certain restrictions on how the Agent manipulates the Internal Argument to change its state, an *agent-manner* root. Speakers describe it as a “pulling motion”, to stress and twist something until the material integrity is ruptured. The root *gau*, in contrast, encodes restrictions as to how the “breaking” Argument changes state, its material composition before and after the event. It is a *patient-manner* root. It specifies a splintering type of breaking, especially of wooden or plant material.

There is a further semantic distinction between *fa'i* and *gau*, and by extension, a distinction between Ergatives and Unaccusatives. The root *fa'i* is a *necessarily initiated* event. It requires that an explicit or implicit participant causes some change which immediately controls the initiation of the event. The root *gau* has no such requirement, it is *inchoative* and therefore may occur spontaneously with no intervention or initiation by an Agent.

The distinction between events which are and are not necessarily initiated appears to align with the *agent-manner* and *patient-manner* distinction, in that they both demarkate Ergative and Unaccusative classhood. It is possible that if a root

bears semantic features assigning it to the *agent manner* class, it therefore must be *externally caused*. Its status as an *agent manner* root necessarily implies the presence of an initiating Agent.

### 3.3.3 The Agent Theta Role and Internal Arguments

An Ergative root is necessarily transitive. It encodes an event involving two participants and therefore it selects two syntactic arguments. One argument is the initiating Agent and the other is not the initiating Agent.

The theta role of the non-Agentive argument relies on the semantic characteristics of the event type. For this reason I term this argument the *Internal Argument*, rather than terms such as *Theme* or *Patient*, to be deliberately vague about the argument's participation in the event denoted by the verb.

Very little is able to be generalised in relation to global characteristics of the Internal Arguments of each predicate.

They may be highly affected, as in (64).

- (64) *na isi e le fafine le popo*  
 PAST split  $K_{Erg}$  the woman the coconut  
 'The woman split the coconut.'

In contrast, they may be barely affected at all, as in (65).

- (65) *na su'e e le fafine le popo*  
 PAST search.for  $K_{Erg}$  the woman the coconut  
 'The woman searched for the coconut.'

They may undergo a change of state, as in (66).

- (66) *sā fufulu e le teine le ipu*  
 IMP wash  $K_{Erg}$  the girl the dish  
 'The girl washed the dish.'



They may undergo a change of location, as in (67).

- (67) *sā tūlei e tamaloa le ta'avale*  
IMP push  $K_{Erg}$  man the car  
'The men pushed the car.'

They may be entirely created by the event of the predicate, as in (68)

- (68) *sā lalaga e le fafine se lavalava*  
IMP weave  $K_{Erg}$  the woman a lavalava  
'The woman wove a lavalava.'

The argument may even be a causing force of the event. The Internal Argument in (69) "causes" the event, by inciting the Ergative argument to laugh.

- (69) *sā ula e le fafine le ulaga*  
IMP laugh at  $K_{Erg}$  the woman the joke  
'The woman laughed at the joke.'

The argument which is marked with Ergative case in regular transitive clauses always receives the Agent theta role. This role has certain global values for all Ergative predicates. Regardless of the event denoted by the predicate and its event structure or semantic specificities, the argument marked with Ergative case always has a small set of constant characteristics.

Primarily, the argument is the initiating force of the event denoted by the root. It necessarily participates in the action of the event. The participation is not necessarily volitional. Reinhart (2002 : 5) notates this characteristic with the diacritic [+C] (for "causation"), a practice adopted here.

However, the sole factor of causation is insufficient in characterising the role of the argument taking Ergative case in all Samoan clauses. Examples (70) and (71) demonstrate that Instruments and Natural Forces make poor Ergative arguments of the predicates 'ai, "eat", and *moto*, "punch".

(70) *sā* 'ai e le naife le fa'i  
 IMP eat K<sub>Erg</sub> D knife D banana  
 ‘\*The banana was eaten with a knife.’

(71) *na moto e le ita Iakopo*  
 PAST punch K<sub>Erg</sub> the anger Iakopo  
 ‘\*The anger punched Iakopo.’

In the example (70), the instrument *le naifi* is a direct participant in the causation of the event denoted by ‘*ai*, “eating”, and therefore bears the [+C] feature. However, the construction is unsuccessful unless the sentence takes the rather unlikely one of an animate knife eating a banana. Likewise, the clause in (71) is only possible if *ita* is some kind of animate agent.

Perhaps unexpectedly, Instruments and Natural Forces are successful Ergative arguments with other predicates.

(72) *'olo'o sasa lo'u tino e le savili*  
 PROG beat my body K<sub>Erg</sub> the breeze  
 ‘The breeze beat my body.’

(73) *na vili e le masini le ogala'au*  
 PAST turn over K<sub>Erg</sub> the machine the log  
 ‘The machine turned over the log.’

What then is the crucial difference which separates predicates like ‘*ai*, “eat” and *moto*, “punch” from predicates like *sasa*, “beat” and *vili*, “turn something over”. They are all canonically Ergative verbs where the Internal Argument is highly affected. I suggest the difference could lie in the richness of their semantics.

The predicates *moto* and *sasa* encode very similar events, yet *moto* restricts heavily the manner in which the agent participates. *moto* requires the Agent to

make a certain shape with his or her hand and strike the Internal Argument. In contrast, *sasa* is very lax about the manner of striking, only encoding a generic impact on the Internal Argument. In my data, it encodes events such as “smacking a dog”, “slapping a table”, “beating a drum” and the example in (72).

Compare the two predicates, *fa‘apāgotā*, meaning “to put under arrest” and *pao*, “to stop s.o.”. Note that *pau* especially relates to the stopping of criminals, as opposed to the predicate *tāofi* which is a broader sense of “to stop”.

(74) *sā fa‘apāgotā e le ta‘ita‘i le faomea*  
 PERF arrest  $K_{Erg}$  the leader the thief  
 ‘The leader arrested the thief.’ theft.’

(75) *sā pao e le ta‘ita‘i le faomea*  
 PERF stop  $K_{Erg}$  the leader the thief  
 ‘The leader stopped the thief (e.g. with his/her new policies on security)’  
 or ‘The leader stopped the thief (e.g. manually)’

The predicate can restrict the directness of causation by the Agent. The predicate *fa‘apāgotā* is semantically rich. It entails that an Agent physically stops and restrains an Internal Argument. In contrast, the predicate *pao* is far more general, encoding any kind of halting an Internal Argument’s action. As such, the predicate *pao* does not necessarily require the physical presence of the Ergative argument. The Ergative argument may be an indirect cause, as in the first reading of (75).

To generalise, some predicates, such as *‘ai* and *moto* bear entailments which require that the event is performed by a human. Other predicates do not entail this requirement. Some predicates such as *fa‘apāgotā* require the Agent to be physically present in the scene of the event. To generalise, Ergatives as a class necessarily impose the feature of initiation onto the argument marked with Ergative case. This is not enough to characterise a whole class. The Agent argument of each predicate has certain other requirements imposed upon it, such as volitionality, physicality or animacy, but none of these apply to the whole class of

predicates. They are specifically determined for each predicate.

Just as I use the term Internal Argument to remain vague about the characteristics of the Absolutive argument's theta role, I will use the term Agent in referring to the Ergative argument's theta role in a very broad sense. I do not claim that an Agent necessarily bears connotations of volitionality or animacy.

### 3.3.4 Ergativity and Selection

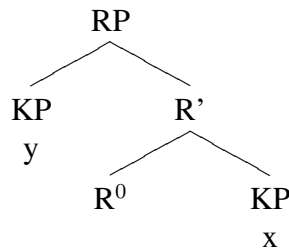
The model in (76) demonstrates a possible lexical entry for each member of the class of Ergative roots, to the extent that generalisations may be made about the whole class.

(76) **Ergative R<sup>0</sup> (Generalised)**

P-signature: *Lexically Specified*

S-signature: *y INITIATES [EVENT INVOLVING x]*

Argument Structure:



The P-signature (pronunciation) for each member of the class of Ergative roots obviously cannot be generalised, except to state that all Ergative roots are free morphemes and therefore, by the model of Conflation defined by Hale & Keyser (2002), not “phonologically defective”.

It is probably too difficult to generalise an S-signature (meaning) for all members of the Ergative class. The attempt in (76) is a rough sketch. It is represented using English as a metalanguage. Two theta roles are defined for two core arguments. They are defined using the variables “x” and the initiating Agent “y”, both

of which must be satisfied by syntactic items elsewhere in the utterance. The exact values of the theta roles are lexically specified, other than the fact that “y” is the Agent and “x” is not the Agent. By modelling theta role assignment as lexically specified (as opposed to approaches defining a bounded, generalised set of theta roles as in (Jackendoff 1990, Dowty 1991)) the subtle variances in the arguments’ participation in different event types can be captured formally. The stipulation to this requirement is that the arguments be generated within the RP.

The Argument Structure under this model determines that the root projects the two arguments in its complement and specifier positions. The choice of which position respond to my earlier argument based on the ability of the Internal Argument to incorporate, and that the Agent appears to c-command the Internal Argument in unmarked clauses.

This model raises a theoretical issue. If this model assumes that the Agent and Internal arguments are projections of the root, how then are examples like (77) and (78) possible? The examples are repeated below.

(77) *na vali le fale*  
PAST paint D house  
‘The house was painted.’

(78) *na ‘ai le fa’i*  
PAST eat the banana  
‘The banana was eaten.’

In these examples, the Agent argument is not present in the phonological form. All Ergative predicates in Samoan can freely appear with or without their Agent. In examples where the predicate appears without the Agent, such as (77), the Agent neither has a syntactic nor semantic value. The construction in (77) may be used when the speaker is downplaying the relevance of the Agent or does not know the identity of the Agent.

Consider the S-signatures of the following Ergative roots conforming to the model defined in (76).

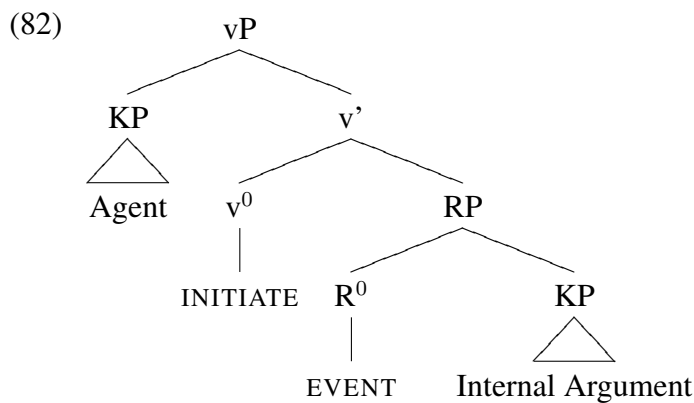
(79) **S-Signature** (*vali*):  
 $x$  INITIATES [PAINT  $y$ ]

(80) **S-Signature** (*sali*):  
 $x$  INITIATES [SCOOP OUT  $y$ ]

(81) **S-Signature** (*sogi*):  
 $x$  INITIATES [CUT  $y$ ]

Numerous linguists working within the Minimalist framework propose a similar semantic structure for transitive predicates. They characterise the INITIATE element as a wholly separate lexical item. This lexical item is termed “VOICE” by Kratzer (1996), “little  $v$ ” by Chomsky (2000, 2004), Collins & Thrainsson (1996) and Stechow (1996), “EXT $\theta$ ” by Pylkkanen (2002) and “INIT” by Ramchand (2008). I will notate this hypothetical item as “ $v^0$ ”.

Under this theory, the  $R^0$  only selects the Internal Argument (in a complement position). The Agent is selected by the  $v^0$  (in a specifier position). The  $v^0$  takes the RP as its complement. The structure of a regular transitive predicate is therefore construed as below.



With this model, Ergative case may be licensed by the specifier position of  $v^0$ . That is, whichever argument appears in the specifier of  $v^0$  receives Ergative case (see Hale & Bittner, 1996a and 1996b).

How applicable is this theory to the Samoan language? It has advantages and disadvantages. If the Agent argument is generated as an argument of a distinct head, the optionality of the Agent in Samoan has a structural motivation.

Compare the examples in (83) and (84).

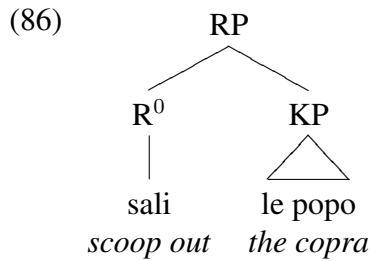
(83) *na sali le popo*  
PAST scoop.out the copra  
'The copra was scooped out.'

(84) *na sali e le tamaloa le popo*  
PAST scoop.out  $K_{Erg}$  the man the copra  
'The man scooped out the copra.'

In contrast, the Internal Argument must have a semantic or syntactic value. It is not optional. If the Internal Argument is not present in the clause, it cannot have a generic, unknown or unstated interpretation. In this respect (85) is not analogous to (83)

(85) *na sali e le tamaloa*  
PAST scoop.out  $K_{Erg}$  the man  
'\*The man scooped things out.'

The structure in (82) predicts this phenomenon. The clause in (84), where both the Agent and the Internal Argument are represented, will take the structure in (82). The clause in (83), where the Agent is deleted, can be structured by simply omitting the  $v^0$ .



While there is a syntactic motivation for the vP analysis, the analysis makes incorrect semantic predictions in two instances.

Firstly, this does not cohere with the model of theta role assignment in this thesis. By my assumptions, theta roles are assigned by lexical items to their projections. Therefore, the R<sup>0</sup> assigns the Internal Argument theta role, and the v<sup>0</sup> assigns the Agent theta role. This notion is argued by Ramchand (2008), Pylkkänen (2002) and Kratzer (1996) and others.

Subsequently, the Agent theta role must be assigned by the same lexical item in every transitive clause. This predicts that there is a single theta role assigned to the Agent in every transitive clause, regardless of the root word's semantics. The participation of the Agent argument in the event must be semantically equivalent for every Transitive clause.

I have established in the previous section that the participation of the Agent depends significantly on the semantics of the root. If the root is a complex action, explicitly detailing the action of an agent, it will preclude Instruments and Natural Forces from taking Ergative case. If the root encodes a semantically vague action, Instruments and Natural Forces are permissible in Ergative case. This accounts for the ability of *le savili*, “the breeze”, to serve as the Ergative argument of *sasa*, “beat”, but its inability to serve as the Ergative argument of *moto*, “punch”. See examples (70-74).



Furthermore, the model separates the semantics of INITIATION from the root word. The  $v^0$  encodes the causation of the event by the Agent. The  $R^0$  encodes the actual processes and/or results of the event (Ramchand, 2008). This presumes that the root word, on its own, does not encode an initiated event. It inherits its status as *initiated* by Merging with  $v^0$ .

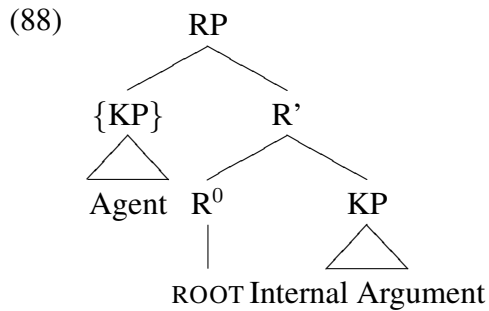
Separating the semantics of INITIATES from the root word predicts that the structure in (86) encodes an inchoative event. This is incorrect. Regardless of whether the Agent is stated or unstated, the event of “scooping out” is necessarily performed by an Agent. It is clearly distinct from actual inchoative events such as *pa'ū*, “to fall” or *mū*, “to burn”. These events can occur without any intervention by an external party.

For these reasons I assume the S-signature of an Ergative root like *sali* is structured as in (80), repeated below. The S-signature is an atomic unit, that is, it is not derived in sentential syntax by Merging with functional heads. Rather, the root *sali* specifies the following semantics in its lexical entry. The element INITIATE is an inherent feature of the semantics of *sali* and all other Ergative roots.

(87) **S-Signature** (*sali*):

$x$  INITIATES [SCOOP OUT  $y$ ]

The structure of Ergative roots therefore follows in (88). I return to the single  $R^0$  analysis projecting both core arguments. The optionality of the Agents is represented simply by stating that the Specifier is an optional branch. Note that Chomsky assumes specifier positions are optional (1995 : 53).



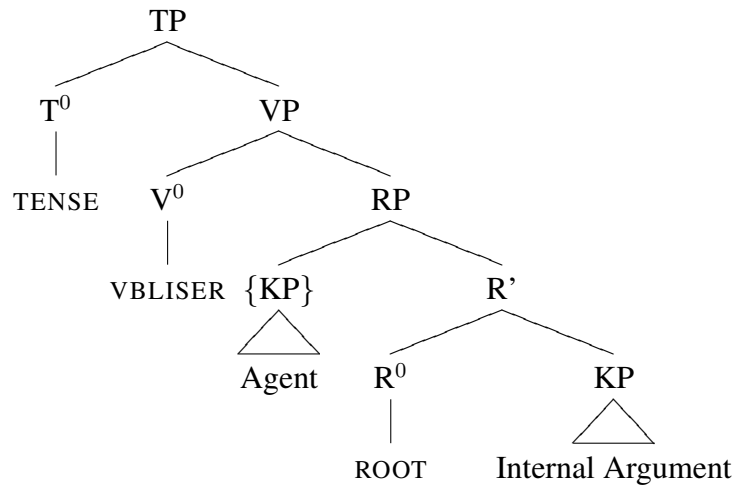
### 3.4 Conclusion

Throughout this Chapter, I have described certain semantic and syntactic properties of Samoan Ergative predicates. These included reflexives, Agent deletion, verbalising affixes and theta role assignment. Considering these properties I have questioned the validity of Minimalist assumptions in representing them.

I also noted the difficulties in generalising the theta assigning properties of Ergative roots. One generalisation is that Ergative roots denote events which must be initiated by an Agent, regardless of whether the Agent is implicit or explicit. Further, as they are transitive, they involve another participant, termed the Internal Argument. The Internal Argument is always syntactically and semantically present in every clause.

In responding to these theoretical issues, I propose a hypothetical structure for a regular Samoan transitive clause, stated below.

(89)



## 4 Non-Ergative Predicates

### 4.1 Introduction

In this chapter, I look at the syntactic structure and behaviour of Samoan predicates which are not Ergative, that is, they cannot freely combine with an argument with Ergative case marking unless they assume additional morphology. Falling under this category are so-called *Unaccusatives*. These are both predicates which assign a theta role to an Internal Argument. I will argue that sub-classes of Unaccusatives may take a stative aspect in the sense of (Vendler 1957). Other Unaccusatives are non-stative, in that denote dynamic events. Another class of Non-Ergatives are the *unergative* predicates. These are dynamic events initiated by a participant that do not involve any other participants.

I reiterate the objective of this analysis, to explore the extent to which certain theoretical models of syntax are most applicable to the Samoan language. Of primary interest in this chapter are models which delimit classes of roots based on their abilities to appear in certain syntactic constructions.

### 4.2 Syntactic Behaviour of Non-Ergatives

The following Samoan sentences have identical constructions in one sense, in that they are finite clauses with a sole Absolutive argument, being characterised by the frame [T R DP]. I take them all to be fundamentally distinct. I propose that the distinction lies within the lexical entries of the predicate the argument structure it projects.

(1) **Ergative**

*na tipi le vao*  
PAST cut the grass  
'The grass was cut.'

(2) **Unaccusative**

*na pa'u le la'au*  
PAST fall the tree  
'The tree fell.'

(3) **Unergative**

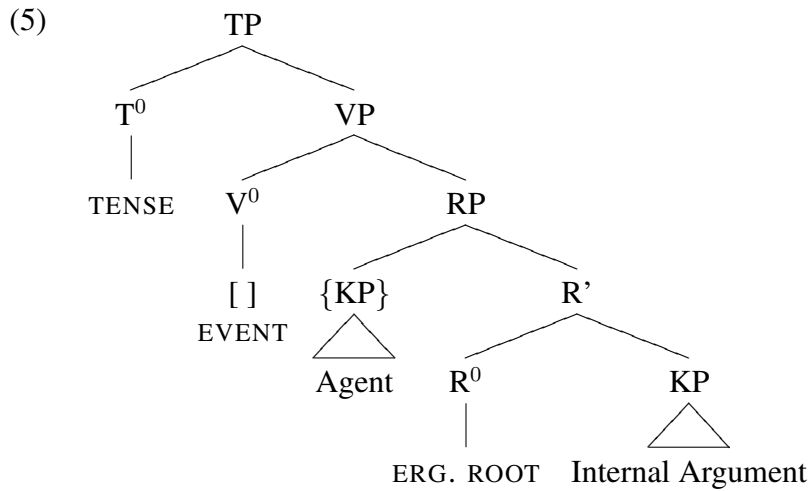
*na alu le fafine*  
PAST go the woman  
'The woman left.'

There are a number of disparities between the predicates in (1-3), especially in relation to the types of constructions in which each predicate may appear. For example, I have already discussed the transitive variant of the predicate in (1), demonstrated in (4).

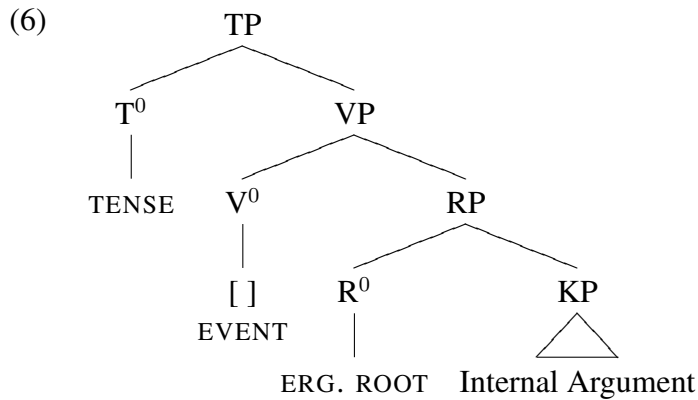
(4) *na tipi e le tama le vao*  
PAST cut  $K_{Erg}$  the boy the grass  
'The boy cut the grass.'

This variant is achieved by simply adding an Ergative KP to the clause with no morphological marking of the verb. In structural terms I have argued that Ergative predicates like *tipi* project two arguments, an Internal Argument and an Agent. The Agent I take to be projected in an optional specifier, accounting for the Transitive-Intransitive variation.

I have further argued that Samoan roots are neither Nouns nor Verbs. To predicate a clause, a root must be "verbalised". This is achieved by the root Conflating with a verbal head ( $V^0$ ). This element is essentially an artifice to ensure the verb initial word ordering and preserve binary branching. As such, it is a marked inefficiency in the model. Regardless, it does correctly predict outcomes relating to structural relations and constituency. The structure for Ergative predicates is displayed in (5).



The Agent argument is bracketed with { } to notate that it is optional. If it is not projected, the clause is rendered as in (6). The agent remains implicit due to the lexical semantics of the root. Ergative events must necessarily be initiated by an Agent.



We may observe that the predicates in (2-3) disallow this kind of "transitivisation". Examples are ungrammatical where an Ergative argument is inserted into the clause without any morphological change to the predicate.

- (7) \**na pa'u e le ta'ita'i le niu*  
 PAST fall  $K_{Erg}$  the chief the coconut tree  
 'The chief made the coconut tree fall.'

- (8) \**na alu e le ta'ita'i le fafine*  
 PAST go  $K_{Erg}$  the chief the woman  
 'The chief made the woman leave.'

To generate causative sentences, where an Ergative argument has caused the state of affairs indicated by the predicate, the predicate must be affixed with the prefix *fa'a-*. Affixing *fa'a-* to the predicates in (7-8) renders them grammatical.

- (9) *na fa'a-pa'u e le ta'ita'i le niu*  
 PAST CAUSE-fall  $K_{Erg}$  the chief the coconut tree  
 'The chief made the coconut tree fall.'

- (10) *na fa'a-alu e le ta'ita'i le fafine*  
 PAST CAUSE-go  $K_{Erg}$  the chief the woman  
 'The chief made the woman leave.'

The roots in (2-3) are intransitive, that is, they take one core argument. The grammatical function of *fa'a-* is transitivity. It is a "valence increasing" operator (by the terminology of Payne, 1997 : 175-196). It introduces a new core argument into the clause.

Non-Ergative roots take one core argument due to the nature of the events they encode semantically. Event types which intrinsically involve one participant are intransitive in Samoan.

An interesting question that I will pursue relates to the existence of a class of Statives. Vendler (1967) categorises verbs into types based on "aktionsart", or "lexical aspect". Verbs are classified by temporal semantic properties, for example, whether they are durative or punctual, telic or atelic, stative or dynamic. Verbs which are stative are labelled States. States as predicates which are by nature attributive. They do not encode events involving any change, motion or a

sequence of internal phases. They are simple inherent characteristics of items.

In Samoan, the question whether there is a lexically distinguished class of Statives is an interesting one. Consider the predicates in (11), which seem to encode states.

- (11) *ma'ai*, “sharp”  
*lelei*, “good”  
*oti*, “dead”  
*'uma*, “finished”  
*'umi*, “tall/long”  
*fou*, “new”  
*pala*, ‘rotten”  
*mago*, “dry”

Mosel & Hovdhaugen (1992 : 338) claim that Samoan does not make a complete lexical distinction between “dynamic and stative events”. They state that certain predicates, such as *alu*, “to go/leave” and *sasa*, “beat”, only encode dynamic events and cannot be interpreted as states. Otherwise, all predicates which can take a stative interpretation can also take a dynamic interpretation.

Lakoff (1966) defines certain tests for determining whether a verb is “Stative”. Dowty (1975) and others argue that at least some of his tests actually test verbs for Agentivity (for example, in English, if the verb cannot grammatically follow the verbs *command* or *persuade*). If a predicate fails an Agentivity test, it is not necessarily Stative, as the class of Stative predicates is a proper subset of the class of Non-Agentive predicates. Nevertheless, at least one of Lakoff’s (1966) tests holds weight, specifically, whether a verb can take a progressive tense/aspect.

The tense particle category, which I have labelled T, strongly determines whether predicates take a stative or dynamic reading.



If the predicates in (11) appear in a clause headed by the present/habitual tense marker *e*, they receive a stative interpretation.

- (12) *e ma'ai le naifi*  
PRES sharp the knife  
'The knife is sharp.'

The tense particle '*ua* appears to indicate that a dynamic event occurred in the past, and the change is relevant at the present time, analogous somewhat to the English *have -en* tense. If the stative predicates in (11) appear with this tense marker, they receive a dynamic, eventive interpretation.

- (13) '*ua ma'ai le naifi*  
PERF sharp the knife  
'The knife has become sharp.'

With the present progressive tense marker, '*ole'ā*, the predicates in (11) indicate that the Absolutive argument is receiving that attribute at the time of speaking.

- (14) '*olo'o pala le apu*  
PROG rotten the apple  
'The apple is becoming rotten.'

The tense particles *sā* and *na* both indicate past tense. Mosel & Hovdhaugen (1992 : 340) state the difference between the particles is aspect. The particle *na* entails that the event is bounded temporally. That is, an event with the *na* necessarily involves a transition from one state of affairs to another.

- (15) *na pala le apu*  
PAST rotten the apple  
'The apple became rotten.'

The tense marker *sā* denotes that the event took place in the past over a static period. If the event type is punctual, the event is interpreted as habitual. When marked by tense particle *sā*, the predicates in (11) receives a stative interpretation.

- (16) *sā pala le apu*  
PAST<sub>2</sub> rotten the apple  
'The apple was rotten.'

These predicates are ambiguous as to whether they represent a State, or a change of State. Presumably, their lexical semantics (S-signature) is neither intrinsically dynamic nor stative. As such it can be dually represented as in (17) and (18).

- (17) **S-Signature** (*ma'ai*):

*x* IS SHARP  
*x* BECOMES SHARP

- (18) **S-Signature** (*pala*):

*x* IS ROTTEN  
*x* ROTS

If the S-signature of a predicate like *pala* was encoded solely as a stative aspect ("x IS ROTTEN"), the example in (14) should be ungrammatical when evaluated by Lakoff's (1966) test for stativity.

Alternatively, the S-signature of a predicate like *pala* could just be a simple change of State ("x ROTS"). The aspectual content of the tense particle determining whether the change of state occurred in the past (and therefore the predicate receives a stative interpretation), or is occurring in the present (and therefore the predicate is dynamic). This hypothesis holds weight when considering the above examples. However, it would predict that the interpretations of predicates like *pala* are in complementary distribution, determined by the tense particle. It would suggest that no tense particles generate ambiguous readings. This is proven false as the present tense marker *e* generates ambiguous readings. The clauses below demonstrate.

- (19) *e ma'ai le naifi o Tasi*  
 PRES sharp the knife of Tasi  
 'Tasi's knife is sharp.'
- (20) *ina e vevela, e ma'ai lea naifi*  
 CONJ PRES hot, PRES sharp this knife  
 'When it is hot, this knife gets sharper.'

The predicate can therefore take a dynamic reading (depending on the context) and is therefore not strictly a Stative. With regards to these predicates, Mosel & Hovdhaugen's (1992) assertion that Samoan lacks true Statives is supported. I will label the predicates in (11) Labile Statives, indicating that they generate both stative and dynamic interpretations.

Note the *ma-* prefix which I discussed in Chapter 3, which transforms Ergative verbs into intransitive Statives. These predicates allow progressive aspect, suggesting they are Labile Statives. They may take a dynamic or stative reading.

- (21) *'ole'ā ma-fa'i le nifo*  
 PROG STAT-break the tooth  
 'The tooth is being broken/breaking.'
- (22) *na ma-fa'i le nifo*  
 PAST DEERG-break the tooth  
 'The tooth was broken.'

Mosel & Hovdhaugen fail to note that there are some intransitive predicates which demonstrate distinct morphology when expressing a stative or dynamic event. The Samoan terms for colours demonstrate this alternation.

(23) **Stative/Dynamic Variation**

uliuli	<i>x is black</i>	tauga-uli	<i>x becomes black</i>
sinasina	<i>x is white</i>	sina	<i>x becomes white</i>
mūmū	<i>x is red</i>	mū	<i>x becomes red</i>
samasama	<i>x is yellow</i>	sasama	<i>x becomes yellow</i>

In these examples, the stative and dynamic version are in complementary distribution. Whenever a dynamic, change of state interpretation is intended, the appropriate morphological version is used. The Stative variants in (23) are unable to take a progressive tense, thereby fulfilling Lakoff's (1966) test for stativity.

- (24) \**'olo'o mūmū le fale*  
 PROG red the house  
 'The house is being/becoming red.'

In my data, I did not find examples of predicates demonstrating morphological change to alternate between dynamic and stative outside the realm of colour, suggesting it may be a very semantically restricted feature of the language.

The stative examples in (23) are not Labile. They only allow a stative interpretation. For this reason I will label them "True" Statives. Other potential members of the class of True Statives are the qualitative predicates *lelei*, "good" and *leaga*, "bad" which also give undesirable results when marked with a progressive tense.

Predicates such as *pala*, "rotten" and *ma'ai*, "sharp" denote ambiguously dynamic and stative events which are both intransitive and Non-Agentive. In this respect they align with the class of Unaccusatives, in the sense of Perlmutter (1978). These are predicates which encode an event which is not necessarily initiated by an Agent-like participant. These events can occur inchoatively or spontaneously. Some other potential members of this class follow in (25).

- (25) *pa'ū*, "fall"  
*goto*, "sink"  
*opeopea*, "float"  
*tafe*, "flow"  
*se'e*, "slide"

These predicates fit the description of an Intransitive event which may occur spontaneously. Therefore they might belong to the Labile Statives class demonstrated in (11). However, the predicates in (25) do not allow the same stative

interpretations generated by Labile Statives. As such I will label them “True” Unaccusatives.

Like Labile Statives, True Unaccusatives take a dynamic reading with the past tense marker *na*.

- (26) *na pa'ū le niu*  
PAST fall the coconut tree  
'The coconut tree fell over.'

When appearing with the *sā* tense marker, the Labile Statives take a stative reading, as in (13), repeated below.

- (27) *sā pala le apu*  
IMP rotten the apple  
'The apple was rotten.'

When True Unaccusatives take the *sā* tense marker, they take a habitual, dynamic reading.

- (28) *sā pa'ū lea 'apefa'i*  
IMP fall this ladder  
'This ladder used to fall over.'

Where Labile Statives take the present tense marker *e*, they are ambiguous as to whether they encode a state or dynamic event. True Unaccusatives simply encode a habitual dynamic event.

- (29) *e pa'ū lea 'apefa'i*  
PRES fall this ladder  
'This ladder falls over (all the time).'  
*but not* '\*This ladder is fallen.'

What semantic feature determines a predicate's classification as a True Unaccusative as opposed to a Labile Stative? That is, what semantic feature excludes the stative interpretation? The predicates in (25) each encode an event where an Internal Argument undergoes a (potentially involuntary) change of location. They do not encode changes of state. The body of the Internal Argument can be totally identical after having undergone the events denoted by True Unaccusatives. I therefore tentatively propose that the semantic distinction between Labile Stative events and True Unaccusative events is that Labile Stative events involve a significant change in the composure of the Internal Argument. It remains to be seen whether other semantic factors determine a predicates membership as a True Unaccusative, True Stative or Labile Stative.

For reasons which I will discuss below, True Statives, True Unaccusatives and Labile Statives share enough syntactic features that it is useful to label them as one class. They are therefore simply labelled Unaccusatives.

The fundamental factors determining a predicates classification as Unaccusatives are that they are intransitive (involving one participant) and are not necessarily initiated by an Agent. They may seemingly occur spontaneously. This places them in contrast with another type of intransitive. This class is termed Unergative (by Perlmutter (1978)).

The predicates in (30) are all bodily movements and functions.

- (30) *moe* "sleep"  
*sola* "run away"  
*fo'i* "return"  
*māfatua* "sneeze"  
*tū* "stand up"  
*oso* "jump"  
*tale* "cough"

*siva* “dance”  
*mapu* “whistle”  
*alu* - “go”

These predicates encode events which are intransitive, that is they only involve one participant. They differ from the Unaccusatives in that the core argument is the initiating force of the event. As such, the predicates in (30) are canonical examples of Unergatives. Other types of events which are Unergative include predicates which encode events of light or sound emission.

(31) *ū* - “make a thundering sound”  
*susulu* - “shine (of the sun and moon)”  
*pulapula* - “shine (of surfaces)”  
*‘emo‘emo* - “twinkle”  
*gaolo* - “rattle”  
*e‘ē* - “squeal”

Although the core arguments of the predicates in (31) are definitely not volitional or animate, they are initiating forces or causers of an event. They align with the Ergative case marked argument of Ergative predicates, which also may be non-volitional and non-animate so long as they are initiators. For this reason, I label the initiators of Unergative arguments Agents. As is the case with the Agents of Ergative predicates, my usage of the term Agent does not imply any agentivity or volitionality.

#### 4.2.1 The Structure of Intransitives

(Levin, 1983) and (Marantz 1984) argue that the sole core argument of an Unaccusative is a Patient at the level of “Deep Structure”, and that the sole core argument of Unergatives is an Agent at the level of “Deep Structure”. Note that their term Patient equates to my concept of Internal Argument. The term “Deep Structure” is a concept from pre-Minimalist Chomskyan linguistics. It refers to

the position in the tree structure where an argument is generated. The intuitions of Levin and Marantz simply state that the argument of an Unaccusative behaves syntactically like the Internal Argument of a transitive predicate, meaning it is generated in the same position. The argument of an Unergative behaves syntactically like the Agent of a transitive predicate and, likewise, is generated in the same position.

Baker (1988 : 46) states this intuition with his Uniformity of Theta Assignment Hypothesis (UTAH).

(32) THE UNIFORMITY OF THETA ASSIGNMENT HYPOTHESIS (Baker, 1988 : 46)

Identical thematic relationships between items are represented by identical structural relationship between those items at the level of D-structure.

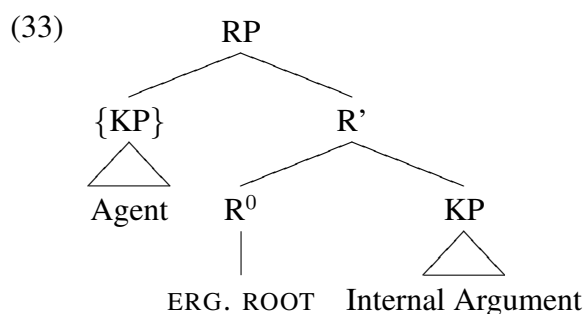
Again I clarify that “D-structure” refers to the position where an argument is generated. The terms of UTAH state that arguments bearing the same theta role are generated in the same structural position. If the Patient of an transitive verb is generated in the complement position of the root, it will always be generated in that position, regardless of whether it is incorporated, topicalised, relativised or in any way re-generated in a second position. Baker (1996) extends this model to state that the theta role of an argument specifically determines its structural position, regardless of its selecting predicate. If an argument is an Agent, it will always be generated in a specifier position.

Throughout this thesis I have argued against the need for generalised theta roles (such as PATIENT, THEME, GOAL, BENEFICIARY). I have instead argued that predicates individually assign specified theta roles to their arguments (‘object being painted’, ‘object being scooped out’ and such). However, I have argued that there does exist (at least one) thematic distinction between the two core arguments of Ergative predicates, Agents and Internal Arguments.



The crucial distinction is that Agents necessarily initiate the performance of an event. Any other core argument (an Internal Argument) has no such requirement. This simple distinction presumes two (very broad) categories of thematic roles. If UTAH (in (32)) is assumed, this distinction is enough to warrant the presumption that the categories of thematic roles are consistently generated in identical structural positions.

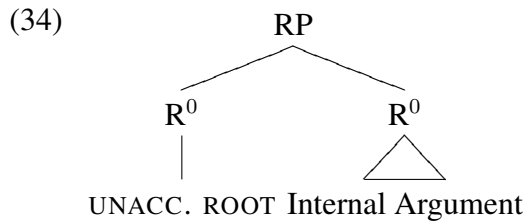
Recall the structure I have argued (in Chapter 3) for the Ergative root and the projection of its two arguments.



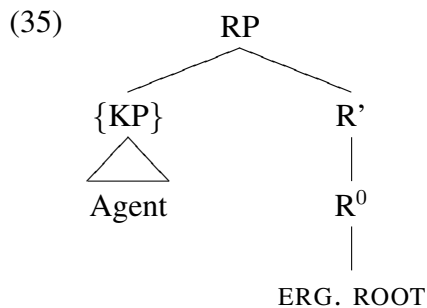
The structure is a generalised Argument Structure for all Ergative roots. That is to say, all members of the Ergative root class are lexically specified to possess this structure, where the argument which initiates the event is generated in the (optional) specifier position and the other argument is generated as the complement.

This generalisation correctly predicts that there is no Ergative root whose Agent follows the Internal Argument in unmarked word ordering. It also correctly predicts that there is no Ergative root whose Internal Argument is optional while their Agent is obligatory.

By extension of this hypothesised structure and the UTAH, Unaccusative predicates should project a structure as in (34). They select a non-initiating argument (an Internal Argument) and therefore it should be generated in the complement position.



Likewise, the sole argument of an Unergative is an Agent. It therefore should be generated in an (optional) specifier position, as in (35).



At this point we encounter a theoretical issue, one that is noted by (Chomsky 1995 : 247). If binary branching is assumed, how is an intransitive predicate with a specifier argument generated. The structure in is predicted to be impossible as it contains a unary branch (between R' and R<sup>0</sup>).

Hale & Keyser (1993) specify the principle of *unambiguous projection* which requires that nodes in a tree be defined as distinct from their children or parents. There is no distinction between R<sup>0</sup> and R' in (35) - they bear the same phonological and semantic features. In contrast, the node RP (which is binary branching) is distinct from R' as it includes in its scope the contents of KP as well as R'.

The model of Merge set out by Chomsky (1995) states that lexical items are selected from the lexicon and build the syntactic structure one piece at a time. He sets out the terms of Merge by stipulating the following requirements on syntactic objects (or nodes in a tree).

(36) TYPES OF SYNTACTIC OBJECTS (Chomsky, 1995 : 243)

Syntactic Objects are of the following types

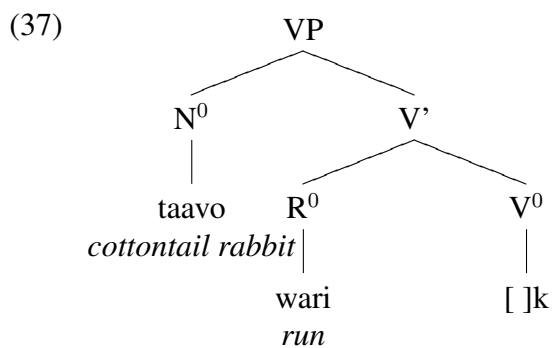
(a) Lexical items

(b)  $K = \{\gamma, \{\alpha, \beta\}\}$ , where  $\alpha, \beta$  are objects and  $\gamma$  is the label of  $K$

This most basically states that all nodes in a tree are either lexical items (terminal) or combinations of two other syntactic objects other binary branching nodes or terminal nodes. This makes a strong hypothesis with regards to binary branching.

How is (35) possible considering these assumptions? Hale & Keyser (1993, 2002) offer a theory whereby Unergative verbs are underlyingly transitive, in that they project two core argument positions, a specifier and a complement. The specifier contains the Agent argument as expected. The head of the structure is a verbalising head and the complement is the Unergative root (by their terms, a nominal or indeterminate category).

The structure in (37) demonstrates Hale & Keyser's (2002 : 53) hypothesis applied to a Hopi example.



The Unergative root *wari*, “to run” conflates with a verbalising morpheme *-k*, deriving the Unergative verb form *warik-*, which must itself be further inflected

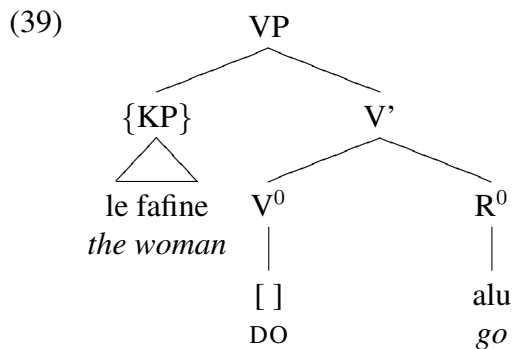
with tense and other particles. This maintains binary branching as it assumes the Unergative root is underlyingly transitive.

Hale & Keyser (2002 : 117) also note the language Basque which necessarily take a light verb *egin* (which has a meaning “do”) in order to compose Unergative verb forms. Some examples of Basque Unergative forms are listed in (38). They suggest that Unergatives are composed of a [V N] head-complement structure.

- (38) *negar egin* - “(do a) cry”  
*eztul egin* - “(do a) cough”  
*barre egin* - “(do a) laugh”  
*jolas egin* - “(do a) play”  
*oihu egin* - “(do a) shout”

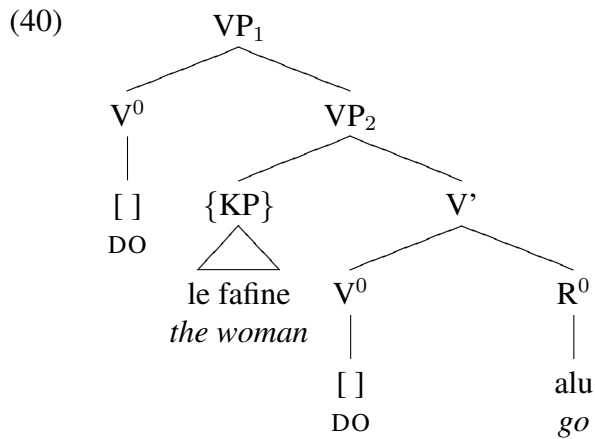
They also propose that this theory accounts for the preponderance of Unergative predicates in English which are ambiguously verbal or nominal (*cough, yawn, laugh, cry, shout, run, walk*). They are of an indeterminate category and conflate with a verbal head, satisfying the requirements of binary branching.

An analogous structure in Samoan (for the Unergative root *alu*) follows in (39).



However, this proposition is highly problematic when considering my analysis of Samoan up until this point. I have *already* proposed the existence of a

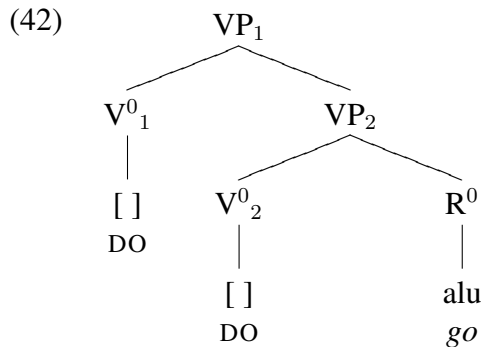
(sometimes) phonologically null verbaliser in the Samoan clause, accounting for Samoan's verb-initial word ordering. If the structure in (39) is embedded as the complement of the (higher) null verbaliser, the tree in (40) is generated.



This structure generates the correct word ordering, however, it is semantically implausible. It contains two (phonologically null) morphemes which serve identical functions. Especially considering that the Agent is projected in an optional specifier. Note the example in (41), where the Agent of an Unergative is omitted. This leaves Absolutive case unassigned. Note that the Allative argument (*'i le ōfisa*) maintains an Oblique case.

- (41) *e alu 'i le ōfisa*  
 PRES go K<sub>All</sub> the office  
 'The office is visited.' / 'People go to the office.'

If the structure in (40) omits the specifier, the highly problematic structure in (42) is generated.



Therefore I will concede that binary branching is untenable in this instance, given my prior assumption of an already present null verbaliser. To maintain UTAH, the Agent of an Unergative must be generated in a specifier position. To maintain binary branching, the Unergative Argument Structure must be a transitive configuration underlyingly. If this is true in Samoan, the head of this “transitive” configuration must be non-overt. Given my established clause structure, this doubles the need for null, verbalising heads - leading to repeated nodes in the structure. When the specifier is omitted there is no option but to generate two consecutive phonologically null heads with the same semantic function, as in (42).

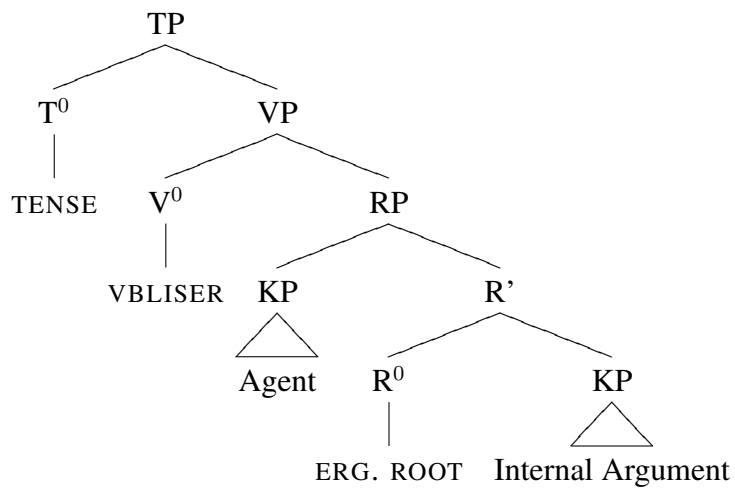
Despite the relaxation on binary branching, I will maintain the analysis using the higher verbalising head, proposed in Chapter 3. The verbalising head correctly generates Samoan’s verb initial word ordering. It provides a structural location for the phonological verbalising heads *fia-*, *ma-* and *-a*. It resolves the ambiguity between a roots nominal and verbal usage and it maintains the correct c-command relationship between the Agent and Internal Argument.

### 4.3 Conclusion

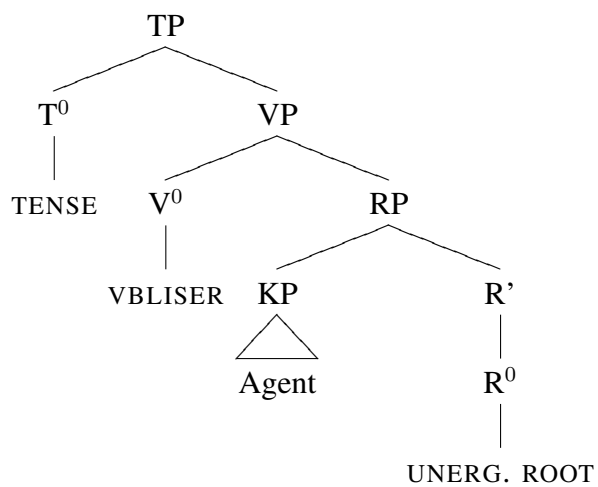
Therefore, at this point, I define three basic clause types in Samoan; Ergative, Unergative and Unaccusative. They are configured with the tree diagrams below. The structure in (45) represents three established subcategories of Unaccusatives; True Unaccusatives (intrinsically dynamic), True Statives (intrinsically stative)

and Labile Statives (neither intrinsically stative nor dynamic).

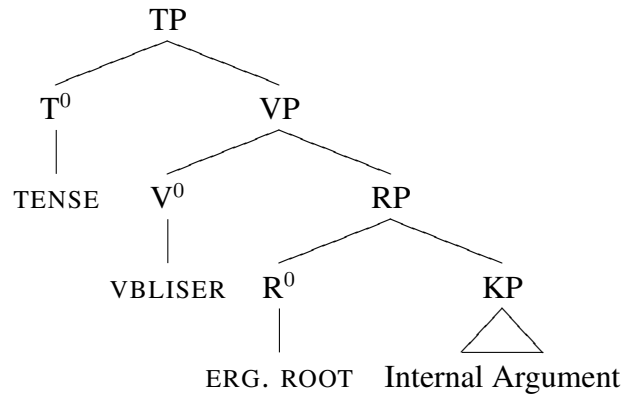
(43) **Ergative**



(44) **Unergative**



(45) **Unaccusative**





## 5 The Samoan Case System

### 5.1 Introduction

In this section I will account for the Samoan system of marking case on core arguments. By core arguments I mean arguments which are generated in the complement or specifier positions projected by the root. This deliberately excludes Oblique arguments. These arguments are adjunct to the root.

Samoan displays an Ergative-Absolutive case marking pattern. The Agents of transitive predicates is marked as distinct from other core arguments. This differs from languages such as English, Italian and Japanese which display a Nominative-Accusative case marking pattern. The Internal Argument of transitive predicates is marked as distinct from other core arguments. Interestingly, the Samoan system of pronominal clitics aligns to a Nominative-Accusative case marking pattern. These notions are considered while I test the applicability of an Optimality Theory analysis in describing the Samoan case system.

### 5.2 The Case System Defined

With a defined clause structure for three types of predicates, I can now give some brief notes about the assignation of Absolutive and Ergative case in Samoan.

The predicate *'ai*, meaning “to eat”, belongs to the Ergative class. It takes two arguments, an Agent and an Internal Argument. The Agent is marked with Ergative case, signaled by the case marker *e* and the Internal Argument with Absolutive case, signaled by no case marker or the case marker *o*.

- (1) *na 'ai e le pepe (o) le fa'i*  
PAST eat  $K_{Erg}$  the baby ( $K_{Abs}$ ) the banana  
'The baby ate the banana.'

Ergative predicates like *'ai* may also participate in alternations where the Agent argument receives Absolutive case rather than Ergative case. These include Incorporation constructions and Partitive constructions, exemplified in (2) and (3) respectively.

(2) *e 'ai fa'i le pepe*  
 PRES eat banana the baby  
 'The baby eats bananas.'

(3) *sā 'ai le pepe i le fa'i*  
 IMP eat the baby *K<sub>Obl</sub>* the banana  
 'The baby was eating away at the banana.'

In (2), the Incorporation construction, the Internal Argument is rendered as a caseless, bare root - directly affixed to the predicate. The partitive construction in (3), where the Internal Argument receives an Oblique case, signifies that the event is atelic. That is, the event does not have a natural temporal conclusion.

Observable in all the above examples is a crucial feature of Samoan syntax, its Ergative-Absolutive case assignment pattern.

A clause is intransitive if it has exactly one argument which is not oblique, that is, one core argument. Unaccusatives and Unergatives are intransitive. The incorporating predicate in (2) and the partitive predicate in (3) are also intransitive. The Internal Argument loses its argument status by incorporating or by being assigned an oblique case. Comrie (1978) refers to the sole core argument of Intransitives as "S". They are always marked with Absolutive case.

If a clause is transitive, it takes two core arguments, the Agent and Internal Argument. Samoan marks the Agent argument of transitives with the *e* case marker. The Internal Argument of a transitive receives Absolutive case. As "S" receives the same case as the Internal Argument of a transitive, the Samoan case system

fits the definition of Ergative-Absolutive.

### 5.3 Formalising the Case System

How is it possible to formalise this case system? We must derive an analysis where the Agent gets Ergative case if and only if the Internal Argument can receive Absolutive case. If the Internal Argument is prevented from receiving Absolutive case (by incorporating or taking Oblique case) the Agent must receive Absolutive case. We can start with the informal generalisation in (4) which states that Ergative case is semantically determined, being attracted to the core argument which initiates or causes the event denoted by the predicate (that is, the Agent).

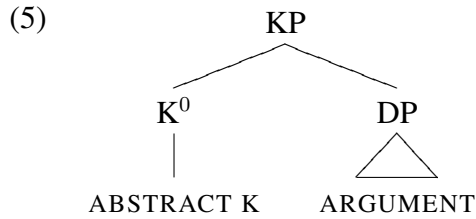
(4) ERGATIVE AS A SEMANTIC CASE

If a core argument is the Agent, it receives Ergative case.

This rule can be formalised by appealing to the model of Distributed Morphology (Halle & Marantz, 1993). This model states that the phonological forms of functional morphemes (at least), such as case markers, tense markers, prepositions and so on, are inserted late in the derivation of a surface structure. That is, these morphemes do not enter into a tree structure with a pre-defined phonological signature. Rather, they are abstractions. The phonological signature is assigned to them based on their relevant syntactic and semantic characteristics.

I will now apply the model to the case marker of core arguments. Let us assume that the  $K^0$  of core arguments is an abstraction, which may take various values based on the syntactic and semantic characteristics of the argument. It is this notion which allows the Absolutive and Ergative cases to be assigned to the Agent KP.

If a KP is generated in a core argument position, it is represented by the generalised structure in (5).



Where the KP in (5) occupies the specifier position of the RP, it receives an Agent theta role. The rule in (4) can therefore be formalised with Distributed Morphology. The rule in (6) is lexically specified for ABSTRACT K.

- (6) ERGATIVE SPELLOUT OF ABSTRACT K  
 ABSTRACT K = /e/ → KP = [+C]

The rule states that ABSTRACT K receives the phonological value of /e/ if the KP is an Agent (represented with Reinhart’s (2002) diacritic [+C]). The KP receives this diacritic by virtue of being generated in the specifier of RP. The phonological value marking Absolutive case (either /Ø/ or /o/) can be assigned to the core argument not subject to the rule in (6).

This rule correctly predicts that the Internal Argument of Ergatives and Unaccusatives receives Absolutive case and the Agent of Ergatives receives Ergative case.

- (7) *sā* ‘ai e le pepe (o) le fa’i  
 IMP eat  $K_{Erg}$  the baby ( $K_{Abs}$ ) the banana  
 ‘The baby was eating the banana.’

- (8) *na pa’ū (o) le niu*  
 PAST fall  $K_{Abs}$  the coconut tree  
 ‘The coconut tree fell over.’

Also note that as ABSTRACT K only heads core argument KPs, the model correctly predicts that Absolutive case is not assigned in example (9), where the Agent of an Unergative is deleted.

- (9) *e alu 'i le ōfisa*  
 PRES go  $K_{All}$  the office  
 ‘The office is visited.’ / ‘People go to the office.’

However, the rule in (6) incorrectly predicts that the Agent of an Unergative receives Ergative case. The ABSTRACT  $K$  in (10) heads a KP which is an Agent, and therefore is spelled out as */e/*.

- (10) *\*na tale e le fafine*  
 PAST cough  $K_{Erg}$  the woman  
 ‘The woman coughed.’

It further incorrectly predicts the assignment of Ergative case to the Agent in incorporation and partitive constructions.

- (11) *\*e 'ai fa'i e le pepe*  
 PRES eat banana  $K_{Erg}$  the baby  
 ‘The baby eats bananas.’

- (12) *\*sā 'ai e le pepe i le fa'i*  
 IMP eat  $K_{Erg}$  the baby  $K_{Obl}$  the banana  
 ‘The baby was eating away at the banana.’

It is difficult to model a rule which assigns Absolutive case to ABSTRACT  $K$ . Ergative case is a “Semantic Case”. It marks arguments which bear a common semantic feature, in that they are initiating forces of events. There is a definable property which can govern the assignment of case. In contrast, Absolutive case is a “Grammatical Case”. The rules governing its appearance are based on the syntactic construction. The semantic properties of Absolutive arguments are unable to be generalised. That is, they may variously be Agents or Internal Arguments. There is no semantic property (such as [+C]) which can govern its appearance.

It appears, in informal terms, that exactly one core argument of a predicate must be assigned Absolutive case.

Let us then maintain the rule in (6) as governing the assignment of Ergative case but appeal to other theories characterising the assignment of Absolutive case. Prince & Smolensky’s Optimality Theory (2004) might provide a solution, using the version extended to syntax (Woolford 2007) and (Grimshaw 2008). Optimality Theory (OT) models syntax as a set of constraints on possible outputs. If a structure has multiple potential outputs, these outputs are compared against a set of violable constraints. The output (or “candidate”) which violates more important or more highly ranked constraints “loses” and is ungrammatical.

The rule in (6), governing the assignation of Ergative case must be restated as a violable constraint. This constraint is termed MAX-ERGATIVE.

(13) MAX-ERGATIVE

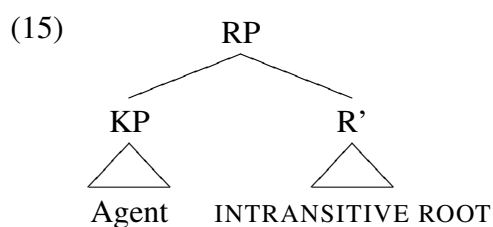
If a KP is [+C] and core it receives Ergative case

The rule stating that Absolutive case must be assigned to one core argument is similarly a constraint. It is restated in (14).

(14) MAX-ABSOLUTIVE

If a KP is core it receives Absolutive case.

The figure in (15) generalises a structure for intransitive predicates with an Agent. This intends to cover Unergatives, incorporation constructions and participative constructions.



The structure in (15) may logically generate two outputs given there are two cases for core arguments in Samoan. An intransitive predicate with an Ergative

agent or an intransitive predicate with an Absolutive agent. They are represented schematically in the ungrammatical (16) and the grammatical (17).

(16) **Intransitive with Ergative Agent** (Incorrect)

\*T INTR. PRED. [*e* ARGUMENT]

(17) **Intransitive with Absolutive Agent** (Correct)

T INTR. PRED. [*o* ARGUMENT]

Constraints in OT must be ranked. That is, one of (13) and (14) must outrank the other. Multiple outputs can therefore be compared based on which constraints they violate. An output which violates more important constraints will “lose” and thus be ungrammatical.

Let us state then that MAX-ABSOLUTIVE outranks MAX-ERGATIVE. A tableau follows in (18). It evaluates the possible outputs (16) and (17) based on the two defined constraints.

(18)

	MAX-ABSOLUTIVE	MAX-ERGATIVE
a. $\text{☞ T PRED } [o \text{ DP}_{Ag}]$		*
b. $\text{T PRED } [e \text{ DP}_{Ag}]$	!*	

Each output violates the constraint MAX-ABSOLUTIVE if it contains a core argument which does not take Absolutive case. Each output violates the constraint MAX-ERGATIVE if an Agent argument does not take Ergative case. The structure in (16) violates a more important constraint while the structure in (17) violates a less important constraint. As such, (17) “wins” and is grammatical.

The correct result is also predicted when evaluating Intransitive predicates with an Internal Argument, that is, Unaccusatives. The Internal Argument must receive Absolutive case. Where the KP receives Ergative case, it violates MAX-ABSOLUTIVE. There are no violations of MAX-ERGATIVE as there are no Agentive arguments.

(19)

	MAX-ABSOLUTE	MAX-ERGATIVE
a. $\text{T PRED } [o \text{ DP}_{IA}]$		
b. $\text{T PRED } [e \text{ DP}_{IA}]$	!*	

With transitive predicates, there are four logically possible outputs. That is, the Agent and the Internal Argument can receive either Absolutive or Ergative case. This warrants the addition of a third constraint. Informally speaking, Absolutive and Ergative cases are never multiply defined in a single clause in Samoan. Only one argument will ever receive either Absolutive or Ergative case. Under OT formalisms, Samoan demonstrates the following constraint in (20).

(20)  $*_{\alpha}C_{\alpha}C$

Core cases may not be multiply defined.

The Constraint in (20) is never broken in Samoan. Therefore it is an undominated constraint. No constraint outranks it. Any output which violates  $*_{\alpha}C_{\alpha}C$  immediately crashes or “loses” and is ungrammatical.

The evaluation of the four possible outputs for transitive predicates is demonstrated in (21).

(21)

	$*_{\alpha}C_{\alpha}C$	MAX-ABSOLUTE	MAX-ERGATIVE
a. $\text{T PRED } [o \text{ DP}_{Ag}] [o \text{ DP}_{IA}]$	!*		*
b. $\text{T PRED } [e \text{ DP}_{Ag}] [o \text{ DP}_{IA}]$		*	
c. $\text{T PRED } [o \text{ DP}_{Ag}] [e \text{ DP}_{IA}]$		*	!*
d. $\text{T PRED } [e \text{ DP}_{Ag}] [e \text{ DP}_{IA}]$	!*	**	

The two options with multiply defined cases are eliminated immediately. The correct clause is selected from the remaining two candidates as it correctly assigns Ergative case to an Agent, unlike the other candidate. The winning candidate only violates one constraint (one argument does not take Absolutive case).



## 5.4 A Note on Clitics

I take the Ergative-Absolutive case system of Samoan to be a product of two factors. It contains a phonological case marking Agentive arguments (Ergative) and a global constraint where Absolutive case must be assigned. Therefore, non-Agentive arguments and the sole argument of Intransitives receives Absolutive case. For this reason, Samoan meets the definition of an Ergative-Absolutive language.

A language like English is Nominative-Accusative. The sole argument of intransitives is marked the same way as the Agentive argument of transitives, as in (22) and (23). They are both pre-verbal. The bold argument in (22) and (23) is Nominative.

(22) **The woman** sleeps.

(23) **The woman** builds airplanes.

Chomsky's proposes the Extended Projection Principle (1981). It is a constraint (active in English) which determines that Nominative case must be assigned in every clause. This is an equivalent to the constraint active in Samoan (which I have formalised as MAX-ABSOLUTIVE) which determines that Absolutive case must be assigned to one core argument. Chomskyan theory defines the Extended Projection Principle structurally. The highest argument in the tree will be assigned Nominative case (by moving to the specifier of T).

Samoan pronominal clitics show a Nominative-Accusative case pattern. If the sole argument of an intransitive is a pronominal, it will cliticise to the right of the tense particle, as in (24). If the tense particle is the present tense *e*, the pronominal cliticises to the left of the tense particle.

(24) *sā*      *'ou alu*  
PAST 1.SG go  
'I went.'

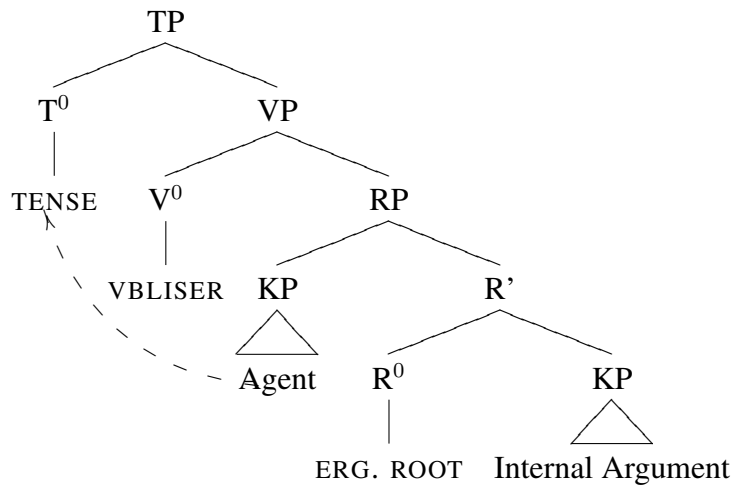
The Agent argument transitives similarly cliticises to the T<sup>0</sup>. The Internal Argument of transitives does not.

- (25) *sā* 'ou *sasa* le *teine*  
 PAST 1.SG hit the girl  
 'I hit the girl.'

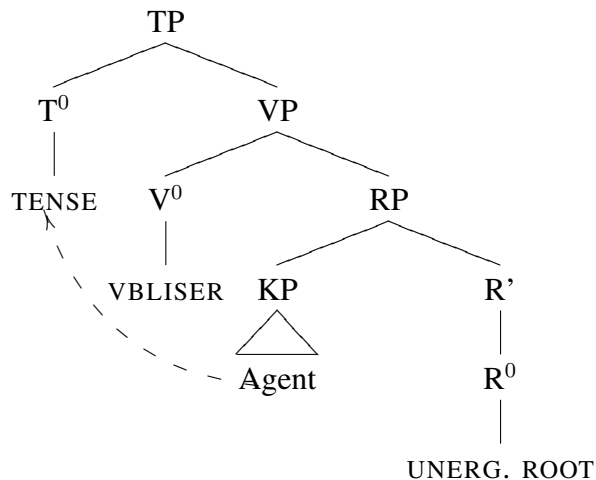
Therefore, Samoan pronominal clitics fit the definition of a Nominative-Accusative case pattern. The Internal Argument of transitives is marked differently from the sole argument of Transitives and the transitive Agent.

Given our defined clause types in Samoan, Chomsky's (1981) account for the assignment of Nominative case applies successfully. The structurally highest argument in the clause is the one which receives Nominative case (if it is a pronominal). From this position it can adjoin to T<sup>0</sup>.

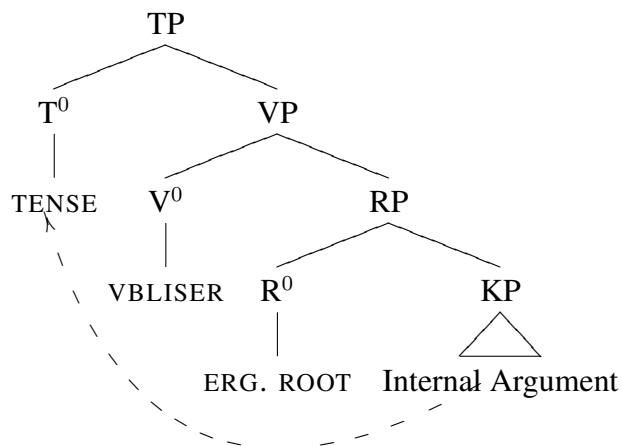
(26) **Ergative**



(27) **Unergative**



(28) **Unaccusative/Stative**



## 5.5 Conclusion

In this chapter I demonstrated the active constraints in the Samoan language which govern the assignment of case in a clause, having established the argument structure of predicates in the previous chapters. In the interests of using Samoan as a means of testing the applicability of various linguistic theories, I

have demonstrated that Optimality Theory is a viable method of determining the characteristics of the Samoan case system.

Furthermore, Chomsky's (1981) analysis of the assignment of Nominative case in English (and other languages) is successfully applied to the Nominative-Accusative system of Samoan pronominal clitics. The structurally highest argument can raise to adjoin to  $T^0$  (its position in the phonological form). If height in the structure is the determining factor of Nominative case as Chomsky assumes, the structure I have proposed for Samoan correctly account for the assignation of Nominative case to transitive agents and the sole core argument of intransitives.

## 6 Transitivity Alternations

### 6.1 Introduction

In this section I discuss two morphological processes whereby Non-Ergative roots are transformed into Ergative roots. Therefore, they are transitivising processes. First, there is the *fa'a-* causative, which introduce into the clause an Agent which causes the event denoted by the root. Secondly, I look at one usage of the *-Cia* suffix which creates Ergative roots from Statives.

In analysing these affixes I discuss two opposingly different perspectives on the formation of morphologically complex words. The first model is the Syntactic model of morphology, which I have been assuming throughout this thesis. This theory states that morphemes (free or bound) enter into syntactic structures separately. They are fused together by regular syntactic processes. In contrast, the Lexicalist hypothesis states that affixes join to roots outside of syntax. That is, roots take affixes and form morphologically complex words in a distinct module. The morphologically complex word is then generated in the syntactic structure.

I assess both theories and their applicability in accounting for the syntactic phenomena associated with the *fa'a-* and *-Cia* affixes.

### 6.2 Causatives

In this section I will discuss the *fa'a-* prefix. This is termed the Causative prefix. It licenses the ability of Unergative and Unaccusative roots to take an Ergative argument. The Ergative argument denotes the causer of the event denoted by the root. Where intransitives take the *fa'a-* prefix, they are “transitivised”. They behave to a large extent like regular Ergative predicates.

The Unaccusative root *pa'ū*, meaning “to fall”, is Non-Ergative. It cannot

combine with an Ergative argument.

- (1) \**na pa'u le niu*  
PAST fall the coconut tree  
'The chief made the coconut tree fall.'
- (2) \**na pa'u e le ta'ita'i le niu*  
PAST fall  $K_{Erg}$  the chief the coconut tree  
'The chief made the coconut tree fall.'

It may only combine with an Ergative argument if it is affixed with the causative prefix *fa'a-*.

- (3) *na fa'a-pa'u e le ta'ita'i le niu*  
PAST CAUSE-fall  $K_{Erg}$  the chief the coconut tree  
'The chief made the coconut tree fall.'

The affix *fa'a-* is a transitivity prefix. It creates Ergative roots from Non-Ergative roots. With the affixation of *fa'a-*, a Non-Ergative root takes two core arguments, an Agent and an Internal Argument.

Observe the differences in meaning between Non-Ergative roots and their counterpart affixed with *fa'a-* in these tables. The tables are divided by the Non-Ergative predicate types Unaccusative and Unergative.

The affixation of *fa'a-* onto Unaccusative roots is simple. If the root predicate encodes an inchoative event, the event occurs without necessarily being initiated by an external participant. When affixed with *fa'a-*, the event is necessarily initiated by an external participant (implicit or explicit). The *fa'a-* affixed variant is analogous to an Ergative predicate. It is a dynamic event involving two participants, necessarily initiated by one participant.

Consistent with the rest of the thesis, I have notated the meaning (S-signatures) of the predicates using English as a metalanguage. The theta roles defined by the

predicates are referenced using the variables “x” and “y”, indicating these are “open slots” to be valued by information from elsewhere in the clause.

(4) **Unaccusative Roots (Dynamic)**

mū	<i>x burns</i>	fa'amū	<i>y makes x burn</i>
mago	<i>x dries out</i>	fa'amago	<i>y makes x dry out</i>
pa'ū	<i>x falls</i>	fa'apa'ū	<i>y makes x fall</i>
goto	<i>x sinks</i>	fa'agoto	<i>y makes x sink</i>
liusuāvai	<i>x melts</i>	fa'aliusuāvai	<i>y makes x melt</i>
opeopea	<i>x floats</i>	fa'aopeopea	<i>y makes x float</i>
se'e	<i>x slides</i>	fa'ase'e	<i>y makes x slide</i>

The affixation of *fa'a-* onto Unaccusative roots with a stative reading is similar. The root encodes an attribute ascribed to an Internal Argument. They, by definition, do not encode dynamic events but characteristics. When affixed with *fa'a-*, the predicate denotes a dynamic event. The stative lexical aspect of the root is cancelled by the *fa'a-* prefix. The *fa'a-* affixed Statives are also transitive. They are events where one participant causes the Internal Argument to bear the state encoded by the root.

(5) **Unaccusative Roots (Stative)**

fou	<i>x is new</i>	fa'afou	<i>y makes x new</i>
leaga	<i>x is bad</i>	fa'amago	<i>y makes x bad</i>
'umi	<i>x is tall</i>	fa'a'umi'umi	<i>y makes x long</i>
malū	<i>x is soft</i>	fa'amalū	<i>y makes x soft</i>
sipa	<i>x is slanted</i>	fa'asipa	<i>y makes x slanted</i>
pala	<i>x is rotten</i>	fa'apala	<i>y makes x rot</i>
paolo	<i>x is shady</i>	fa'apaolo	<i>y makes x shady</i>

The next list displays Unergative roots and their *fa'a-* affixed counterparts. These examples differ from the Unaccusative examples in that both arguments of the *fa'a-* affixed version are “Agentive”. Unergative roots encode intransitive

events which are perpetrated or initiated by some force. Typically the event involves the physical body of the participant. When the root is affixed with *fa'a-*, the Agent of the root (the Caused Agent) is convinced, forced or incited to perform the action by the external party, the Agent of *fa'a-* (the Causing Agent).

(6) **Unergative Roots**

moe	<i>sleep</i>	fa'amoe	<i>y makes x sleep</i>
alu	<i>go</i>	fa'aalu	<i>y makes x go</i>
siva	<i>dance</i>	fa'asiva	<i>y makes x dance</i>
tagi	<i>cry</i>	fa'atagi	<i>y makes x cry, sound out</i>
tale	<i>cough</i>	fa'atale	<i>y makes x cough</i>
māfatua	<i>sneeze</i>	fa'amāfatua	<i>y makes x sneeze</i>

The affixation of *fa'a-* onto Ergative predicates has mixed success. Ergative roots are mostly unable to be affixed with *fa'a*. A very small number of examples of Ergative verbs taking the *fa'a* suffix were found in my research. In these examples, the Internal Argument of the verb is marked with an oblique case or incorporated. The examples in (7-11) show causative-Ergative alternations with the predicates '*ai*, "eat", and *susu*, "suck". Examples where the Internal Argument is incorporated are also demonstrated.

(7) *na 'ai e le pepe le fa'i*  
 PAST eat  $K_{Erg}$  the baby the banana  
 'The baby ate the banana.'

(8) *na fa'a-'ai e le fafine le pepe 'i le fa'i*  
 PAST CAUSE-eat  $K_{Erg}$  the woman the baby  $K_{Obl}$  the banana  
 'The woman fed the baby the banana.'

(9) *na fa'a-'ai fa'i e le fafine le pepe*  
 PAST CAUSE-eat banana  $K_{Erg}$  the woman the baby  
 'The woman fed the baby bananas.'



- (10) *e susu e le pepe le fagu*  
 PRES suck  $K_{Erg}$  the baby the bottle  
 ‘The baby sucks the bottle.’
- (11) *e fa’a-susu e le fafine le pepe ‘i le fagu*  
 PRES CAUSE-suck  $K_{Erg}$  the woman the baby  $K_{Obl}$  the bottle  
 ‘The woman feeds the baby the bottle.’
- (12) *e fa’a-susu fagu e le fafine le pepe*  
 PRES CAUSE-suck bottle  $K_{Erg}$  the woman the baby  
 ‘The woman bottle feeds the baby.’

How productive is the affixation of *fa’a-* onto Ergative roots? The existence of examples suggests that the construction is not syntactically problematic. The rarity of examples might suggest that it requires very specific semantic or syntactic circumstances. These circumstances can only be extrapolated from the existing data in (7-11). What about these examples licenses a causative construction?

Two hypotheses are formulated. The first relates to the volitionality of the Agent. In the examples above, the Agent of the caused event (“the baby”) is mentally incapable of being volitional. It is being forced to undertake the action denoted by the predicate. Potentially, the causativisation of Ergative roots might relate to the volitionality of the Agent of the caused event (or the “Causee”, using the terminology of Levin & Rappaport Hovav (2005)).

The hypothesis states that a non-volitional Causee (who is forced to perform the action) licenses the causative prefix. The rarity of causativised Ergatives is explained by the fact that Agents in regular Ergative clauses are prototypically (but not necessarily) volitional. This analysis is attractive but runs into problems when considering examples such as those below. This example and others allow an interpretation where Causee can be volitional.

- (13) *na fa’a-alu e le ta’ita’i le fafine*  
 PAST CAUSE-go  $K_{Erg}$  the chief the woman

‘The chief made the woman leave.’

- (14) *sā fa'a-siva e le ta'ita'i le fafine*  
IMP CAUSE-dance  $K_{Erg}$  the chief the woman  
‘The chief made the woman dance.’

- (15) *sā fa'a-moe e le ta'ita'i le fafine*  
IMP CAUSE-sleep  $K_{Erg}$  the chief the woman  
‘The chief made the woman sleep.’

A second hypothesis explaining the lack of causativised Ergatives relates to the case of the Internal Argument. In the non-causative variants (7), repeated below, the Internal Argument (*le fa'i*) takes Absolutive case. When the predicate is causativised, the Causee (*le pepe*) takes Absolutive case. It effectively “steals” Absolutive case from the Internal Argument.

- (16) *sā 'ai e le pepe le fa'i*  
IMP eat  $K_{Erg}$  the baby the banana  
‘The baby ate the banana.’

- (17) *sā fa'a-'ai e le fafine le pepe 'i le fa'i*  
IMP CAUSE-eat  $K_{Erg}$  the woman the baby  $K_{Obl}$  the banana  
‘The woman fed the baby the banana.’

- (18) *sā fa'a-'ai fa'i e le fafine le pepe*  
IMP CAUSE-eat banana  $K_{Erg}$  the woman the baby  
‘The woman fed the baby bananas.’

It is a global rule of Samoan that Absolutive case may only be assigned once per predicate (a constraint I characterised in an Optimality Theory format in the previous chapter). All examples with more than one argument taking Absolutive case are immediately excluded by speakers.

Therefore, the assignation of case to the Internal Argument in examples (7-11) poses a serious problem. The Internal Argument must take case via alternative

strategies. It may incorporate, as in (18). Incorporated objects relinquish their requirement to take case. It also may take an Oblique case, as in (17).

Potentially, the rarity of causativised Ergatives might be expected. The Internal Argument must bear certain semantics to license incorporation or Oblique case. If the Internal Argument is incorporated, it must be interpreted as non-specific and plural (or even non-referential). The incorporating predicate must encode a habitual action. If the Internal Argument receives Oblique case, its semantic role must cohere with a particular Oblique case. In the examples above, the Internal Argument can validly take an Instrumental interpretation (and thus receive the Instrumental case marker 'i). It is only in these semantic circumstances that the affixation of *fa'a-* onto Ergative roots is licensed.

A crucial question is whether *fa'a-* can productively affix to Ergative roots if the Internal Argument is incorporated. The hypothesis above predicts this will be the case. These hypotheses raise questions which might prove an interesting starting point for fruitful future research.

The remainder of this section discusses formal concerns relating to the construction of morphologically complex words, such as *fa'a-* predicates. At this point, I wish to discuss conflicting assumptions in theoretical linguistics with regards to morphology. The behaviour of *fa'a-* will serve as a problem for conflicting models, comparing their capabilities. The particular conflict in question relates to complex word formation.

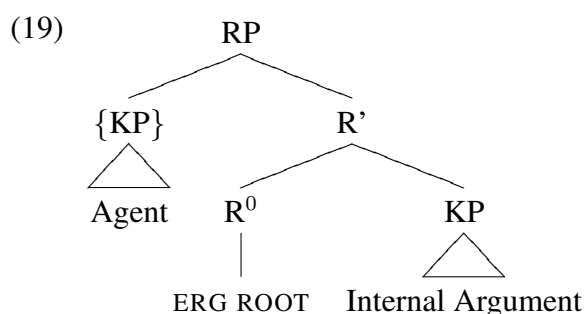
The formalisation of complex word formation is a somewhat controversial issue in linguistics. The debate hinges on whether a distinct system of "Word Formation" in a grammatical model is justified. The Lexicalist hypothesis (Scalise & Guevara 2005) states that complex words (such as *fa'a* predicates, consisting of more than one morpheme) are derived by a set of Lexical Rules. These rules are

independent of the syntactic rules of the language's grammar. The rules operate in a pre-syntactic environment and thus, when a tree structure is built, morphologically complex words are the "building blocks". In structural terms, the items which occupy terminal nodes of a tree are morphologically complex words.

In opposition is the approach which I have assumed in this thesis. This approach treats roots and affixes as lexically equivalent and eschews the need for a distinct, Word Formation module. The hypothesis determines that morphemes are the relevant unit which build syntactic structures. Morphologically complex words are built within the syntax. This approach is typified by (Baker, 1988) who argues that affixation is comparable to incorporation. Chomsky (1995 : 183-185) references Baker's findings and builds his model based on their assumption. Hale & Keyser (2002) extend the model with the notion of Conflation. Their hypothesis (outlined here in Chapter 2) states that affixation is an automatic process where a root is generated syntactically as the complement of an affix.

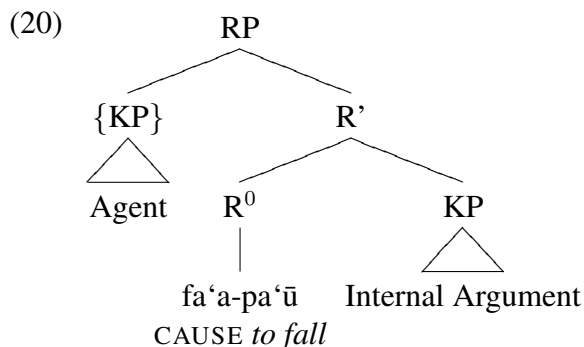
Let us assume that either argument holds weight and explore their possibilities.

Recall my hypothesised Argument Structure for Ergative roots, demonstrated in (19). It assigns an Agent theta role to its optional specifier and an Internal Argument theta role to its complement.



If *fa'a-* predicates are Ergative, their underlying structure must be in some way equivalent. If the Lexicalist hypothesis is assumed, a *fa'a-* predicate enters

the syntax fully formed. The Argument structure of *fa'a-pa'u* is represented in (20)



This predicts that the Agent argument of *fa'a-* predicates is optional. This prediction is accurate.

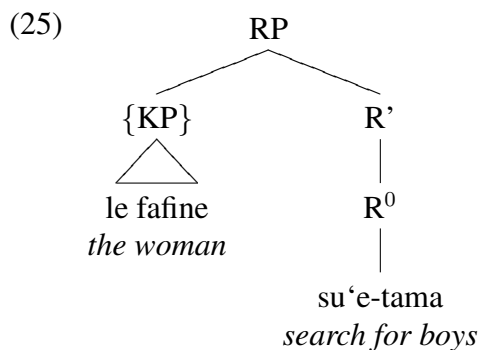
(21) *sā fa'a-pa'u le niu*  
 IMP CAUSE-fall the coconut tree  
 'The coconut tree was made to fall.'

(22) *sā fa'a-alu le fafine*  
 IMP CAUSE-go the woman  
 'The woman was made to leave.'

(23) *e fa'a-leaga le i'a*  
 PRES CAUSE-bad the fish  
 'The fish was spoiled.'

However, problems arise for the Lexicalist hypothesis when considering incorporation. In (Collins, 2010) I discussed the ability of Samoan predicates to incorporate syntactically complex items. This process evades description using a Lexicalist framework. The Lexicalist theory (Di Sciullo & Williams, 1987) proposes that sequences of a predicate and its incorporated object are formed pre-syntactically. The complex predicate *su'e tama* in (24) is therefore formed independently of syntactic rules and inserted into the syntax as a terminal node.

- (24) *e su'e tama le fafine*  
 PRES search boy the woman  
 'The woman is searching for boys.'



I noted that the Samoan language (as well as the related languages Niuean (Massam, 2001) and Tongan (Ball, 2005)) can incorporate syntactically complex objects. Consider the examples in (26) and (27).

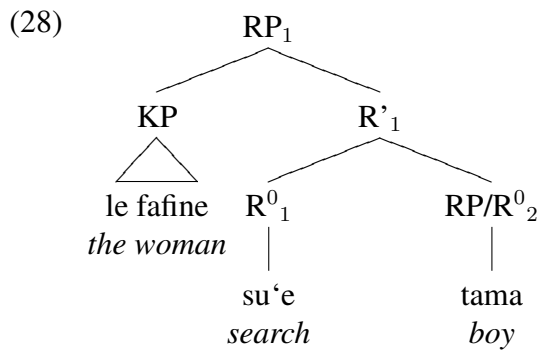
- (26) *e su'e ta'apaepae 'ia 'a'ai le ta'ifau*  
 PRES search chicken SUBJ eat.PL the dog  
 'The dog searches for chickens to eat.'

- (27) *e su'e ma'a e togi 'i ai le atigipūpū le teine*  
 PRES search rock PRES throw ANAPH the seashell the girl  
 'The girl searches for rocks to throw the seashell against.'

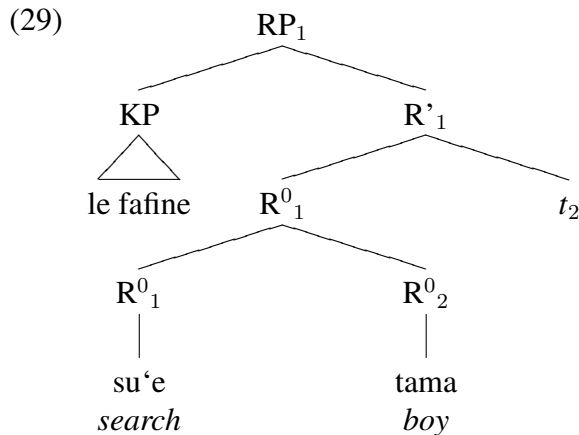
It is clear that the complex predicates *su'e ta'apaepae 'ia 'a'ai* in (26) and *su'e ma'a e togi 'i ai le atigipūpū* cannot be formed external to syntax and be generated in a single terminal node. They demonstrate syntactic phenomena. In (26), the nominal *ta'apaepae*, "chickens", is plural. As such it triggers plural agreement on its selecting predicate *'a'ai*. In (27), the nominal *ma'a*, "rocks", binds an anaphor within a relative clause. These processes necessarily indicate the operation of syntax and therefore the complex predicates in (26) and (27) cannot be formed "outside" of syntax.

I instead maintain an analysis consistent with Baker (1988), where the incorporate object is generated in a core argument position as a bare root (without a  $D^0$  or  $K^0$ ). It cannot receive case without a  $K^0$  and therefore moves to adjoin to the selecting  $R^0$ .

The tree is generated as in (28).



The movement of incorporation derives the tree in (29).



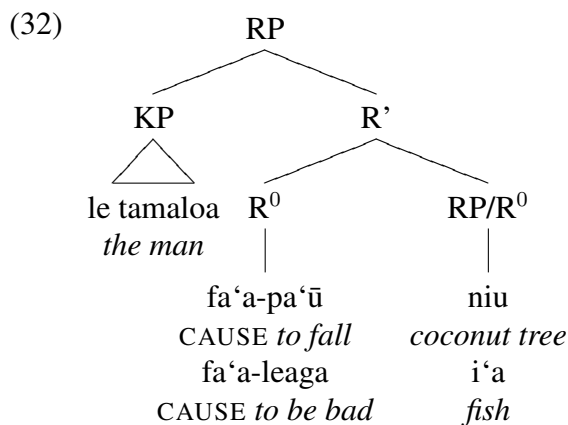
If *fa'a-* predicates are Ergative, they should be able to incorporate their Internal Argument - an pivotal characteristic of Samoan Ergative roots. This is only true for certain classes of *fa'a-* predicates.

*fa'a-* predicates with Unaccusative and Stative roots are able to incorporate their Internal Arguments.

(30) *e fa'a-pa'ū niu le tamaloa*  
 PRES CAUSE-bad coconut tree the man  
 'The man fells coconut trees.'

(31) *e fa'a-leaga i'a le tamaloa*  
 PRES CAUSE-bad fish the man  
 'The man spoils fish.'

This is expected given the model in (20). The Internal Arguments should be generated in a complement position and therefore can legally incorporate.

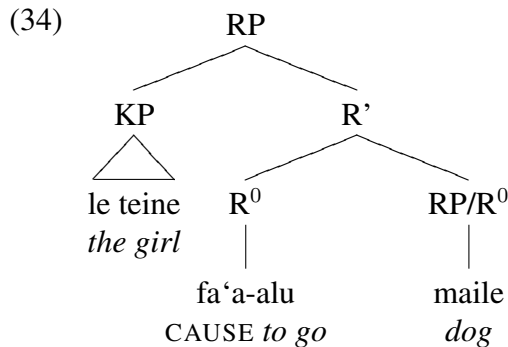


In contrast, *fa'a-* predicates with Unergative roots are unable to incorporate their Internal Arguments.

(33) \**e fa'a-alu maile le teine*  
 PRES CAUSE-go dog the girl  
 'The girl shoos dogs away.'

The model in (20) predicts that *fa'a-alu* is an Ergative verb and takes the Internal Argument as a complement. There is therefore no structural reason why (33) is impossible.





A hypothesis could be made that the failure of (33) is semantic. As a local rule of Samoan, incorporation is simply unable to target Agents. This generalisation appears to hold weight - no examples in my data demonstrate the incorporation of an Agent. However, this generalisation has no motivation other than to provide a rule for the ungrammaticality of (33).

Further, the generalisation is not supported by cross linguistic evidence. Hale & Keyser (2002 : 52) give the following example from Hopi, clearly demonstrating the incorporation of the Agent of an Unergative.

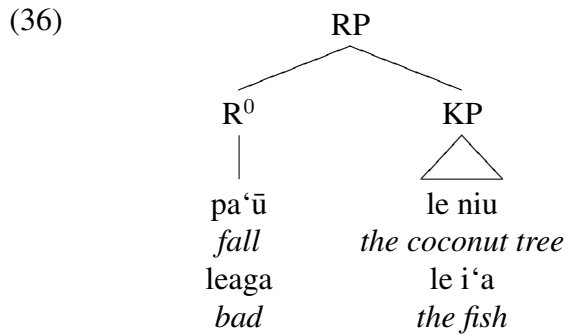
- (35) *itam tap- wari- k- na* (*tapwarikna*)  
 we rabbit run V CAUSE  
 ‘We made the cottontail rabbit run.’ or ‘We flushed the cottontail rabbit out of hiding.’

Instead, let us consider the Syntactic approach to morphology. By this approach, all morphemes (roots and affixes) enter into syntax unaffixed. By the laws of Conflation in the sense of Hale & Keyser (2002), summarised in Chapter 2, a bound morpheme selects its root as its complement in a tree structure. If *fa'a-* is a bound morpheme, it must select the root *alu* as its complement. The projections specified by the Argument Structure of the root are preserved.

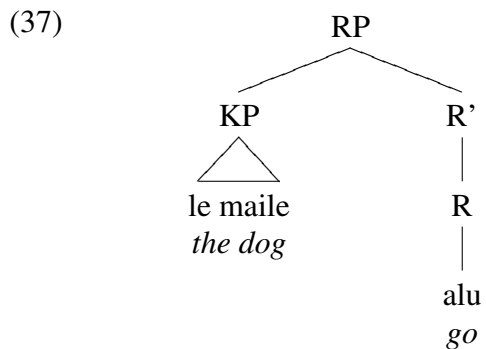
Earlier in this Chapter I discussed the Argument Structure of Unergative roots. I compared them in contrast to Unaccusative roots. By the Uniformity of Theta

Assignment Hypothesis (Baker, 1988), I proposed the structure in (36) for Unaccusative and Stative roots and the structure in (37) for Unergative roots.

For Unaccusatives, the Internal Argument is projected in a complement.

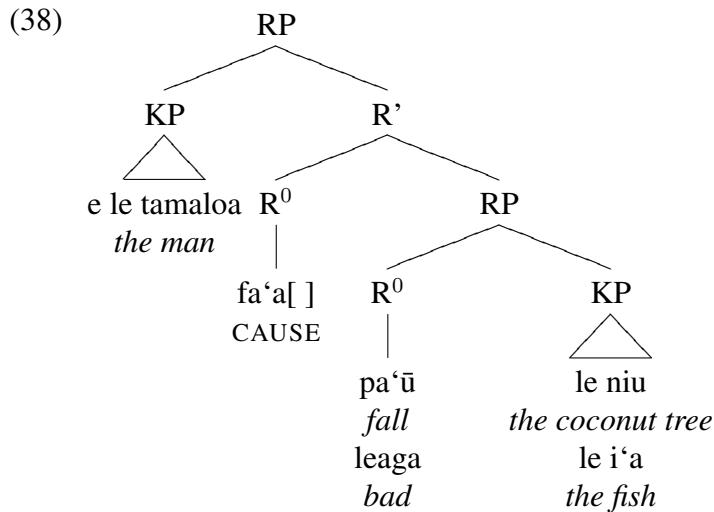


For Unergatives, the Agent is projected in a specifier.

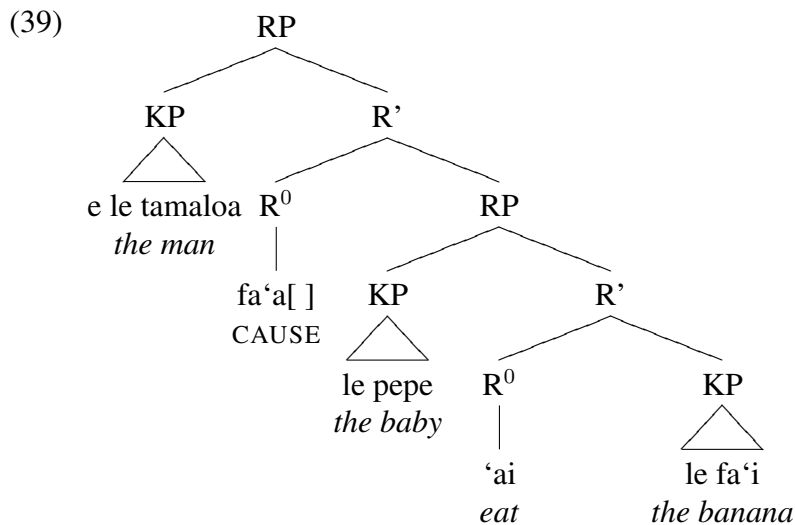


If the Syntactic approach to morphology is assumed, these structures are preserved when the root serves as the complement of *fa'a-*.

The tree in (38) demonstrates the *fa'a-* morpheme taking an Unaccusative root as a complement. The structural projections of the Unaccusatives are preserved. This structure accounts for the incorporation in (30) and (31). The Internal Argument is generated as a complement of the root and therefore can legally incorporate (if it is a bare root).

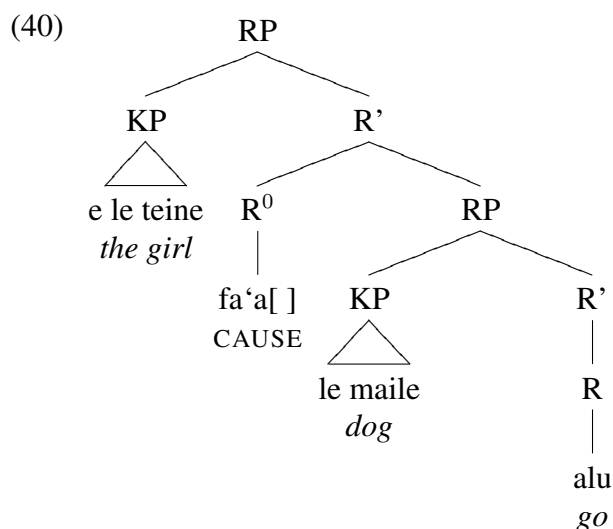


The structure in (39) demonstrates the *fa'a-* prefix taking an Ergative root as a complement. Again, the structural projections of the Ergative root are preserved. Incorporation is also licensed here, as the Internal Argument is a complement of the root. The Internal Argument may otherwise take an Oblique case.



Where *fa'a-* selects an Unergative root as a complement, the structure in (40) is generated. The Unergative root projects its Agent as a specifier. As such, incorporation is not licensed. The Agent is not projected in a complement position.

The adjunction of the agent to the predicate is impossible. The Agent would not c-command its trace.



The structures in (38-40) must Merge as the complement of the verbalising  $V^0$  (in TPs) or the nominalising  $N^0$  (in DPs) and thus derive predicate-initial word order, as discussed in Chapter 3.

This section demonstrates the advantage of a Syntactic model of morphology in describing *fa'a*- predicates. In this model, the structural projections of the root are preserved. The causative morpheme *fa'a*- is generated separately as a distinct head and does not alter the root's structural projections. This offers an explanation as to why the Agents of *fa'a*- affixed Unergative roots are unable to incorporate, despite being the object of an Ergative predicate.

The following section discusses another process where Ergative predicates are derived with additional morphology, the so-called *-Cia* suffix, using the terminology of (Cook, 1996, Milner, 1974). This section will argue that, in contrast, the Lexicalist hypothesis offers a simpler explanation as to how these predicates are derived.

### 6.3 The -Cia Suffix

For derivations with *fa'a*, I assumed a model of morphology where all morphemes (free and bound) are generated separately in the tree. They form morphologically complex words by process of conflation. As they are generated in the syntax, the argument structure of the root is preserved. However, not all affixes in Samoan fit this model effectively.

One use of the functionally very diffuse suffix *-Cia* is to create Ergative verbs from (Labile or True) Statives. A pair of examples follows below.

(41) *sā va'ai le tama*  
IMP see the boy  
'The boy was aware.'

(42) *sā va'ai-a e le tama le pepe*  
IMP see-CIA  $K_{Erg}$  the boy the baby  
'The boy was looking after the baby.'

If an analysis using the syntactic derivation is assumed, the Internal Argument of the root ("the boy") is re-assigned an Agent theta role when the root is affixed with *-Cia*. This is theoretically highly problematic. With a derivational approach it involves movement into a theta assigning position.

I will demonstrate in this section that forming these morphologically complex words with *-Cia* is simpler and accounts for certain phenomena if the Lexicalist hypothesis of morphology is assumed.

The symbol *C* in the *-Cia* suffix represents a variable consonant that is lexically selected by the root. Roots appear to select the consonant in an unpredictable manner. Further, the *-Cia* suffix may be spelled out as *-ina* or *-ia*.

(43) *alofa*, "be kind" → *alofagia*, "treat well"

*lava*, “be able” → *lavatia*, “manage”  
*ala*, “be the cause” → *alafia*, “be fair to”  
*sola*, “have a friend” → *soania*, “pair with”  
*sāuā*, “cruel” → *sāuāina*, “persecute”  
*ganagana*, “talk” → *ganaia*, “talk about”

In this respect, the suffix differs from *fa’a-*. The phonological processes involved when affixing *fa’a-* to a root are minimal and predictable. Consider the *fa’a-* predicate *fa’a-alu*. By a model of syntactic derivation, the two lexical entries, *fa’a-* and *alu* are simply fused together with a predictable phonological outcome.

Consider the predicate *alofa-gia*. In a syntactic derivation model, the morphemes *alofa* and *-Cia* exist separately in the lexicon. They are fused by conflation to form the morphologically complex *alofa-Cia*. The unspecified consonant must be valued. The value of the consonant is lexically specified by the predicate *alofa*. This means that some information regarding affixation to *-Cia* must necessarily be stored in the lexical entry of *alofa*. If at least one feature of the *-Cia* suffix is “visible” to the lexical entry of *alofa*, what motivation is there for (any) other features being obscured.

In addition to being highly variable in phonological form, the *-Cia* suffix is highly variable in grammatical function. I define a list of possible functions of *-Cia* in Samoan below found in my data. They are listed with examples.

- (44) Deriving an Unaccusative from an Ergative  
*tanu*, “x BURIES y” → *tanumia*, “y BECOMES BURIED”
- (45) Changing an atelic Ergative into a telic Ergative  
*sausau*, “x SPRINKLES WATER ON y” → *sausauina*, “x SATURATES y”
- (46) Changing an Unaccusative into an Ergative  
*va’ai*, “x IS AWARE” → *va’aia*, “x OBSERVES/SCRUTINISES y”

(47) Changing a Nominal into an Unaccusative

*afā*, STORM → *afātia*, “x IS STRUCK BY A STORM”

Note that this may well be a variant of the *-a* suffix which I discussed in Chapter 2. That suffix gives a meaning of “being provided with NOMINAL”.

(48) Optional suffix giving no semantic change or syntactic change to Ergatives  
*tipi*, “x CUTS y” → *tipiina*, “x CUTS y”

For simplicity I will restrict my discussion of morphology to the version of *-Cia* in (46), which creates Ergative verbs from Unaccusatives. I will demonstrate why this suffix cannot be represented with a syntactic approach to morphology.

Many Unaccusatives in Samoan (particularly Stative Unaccusatives) may take an additional argument with an Oblique case marking. Chung (1978) and Cook (1996) call these constructions Middles while Mosel & Hovdhaugen (1992) calls them Semi-Transitives. Mosel & Hovdhaugen state they typically encode events of perception, feeling and thinking (1992 : 730). Some examples follow in (49) and (50).

(49) *e alofa le tama i le fafine*  
PRES love the boy  $K_{Obl}$  the woman  
‘The boy loves the woman.’

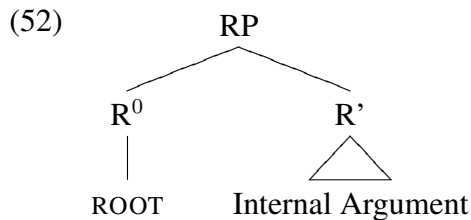
(50) *sā va'ai le tama i le fafine*  
IMP see the boy  $K^{Obl}$  the woman  
‘The boy saw the woman.’

An essential property of Middles is that they may optionally appear without the Oblique argument. The table in (51) demonstrates the variance in meaning when a Stative predicate takes or does not take an Oblique argument.

(51) **Derivation of Middles**

alofa	x IS KINDHEARTED	alofa i Tasi	x LOVES TASI
va'ai	x IS AWARE	va'ai i Tasi	x SEES TASI
fefe	x IS AFRAID	fefe i Tasi	x FEARS TASI
agaleaga	x IS CRUEL	agaleaga i Tasi	x TREATS TASI BADLY
musu	x IS SULLEN	musu i Tasi	x DOES NOT HELP TASI
muli	x IS AT THE BACK	muli i Tasi	x FOLLOWS TASI

The roots in (51) all appear to be regular Unaccusative roots, and thus can be represented with the structure in (52).



When the roots take an Oblique argument their meaning is restricted with respect to the Oblique argument. Without the Oblique argument, the root encodes an attribute of the Internal Argument which can be applied generally. The addition of the Oblique argument specifies the scope of the attribute.

I propose that the Oblique arguments in (51) are adjuncts. The first piece of evidence is their optionality. They do need to be syntactically or semantically present in the clause. As the state encoded by the root may be applied generally, there is no need to assume that the Oblique argument is necessarily implicit.

Further, they demonstrate essential syntactic properties of adjuncts. Literature on island effects (see Goodluck & Rochemont, 1992) states that a property of adjuncts is their inability to move. That is, a trace cannot occupy an adjunct position. This is demonstrated in Samoan by the fact that when the Oblique argument of middles is relativised or topicalised to a higher position, it binds a phonologically overt anaphor.



(53) *'o le fafine s̄a va'ai ai le tama*  
*K<sub>Top</sub> the woman IMP see ANAPH the boy*  
 'As for the woman, the boy saw her.'

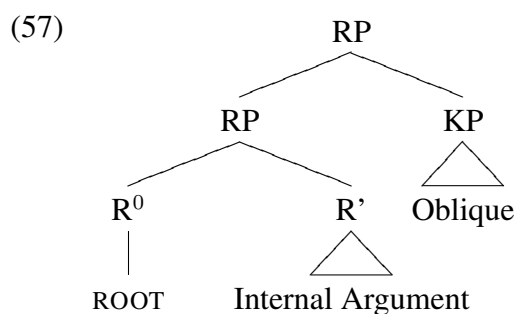
(54) *le fafine s̄a va'ai ai le tama*  
*the woman IMP see ANAPH the boy*  
 'The woman that the boy saw.'

This is evidence that the Oblique argument does not occupy the position of a core argument. Compare the examples in (53) and (54) with the examples below, where a core argument is relativised and topicalised. No anaphor is left in the matrix clause.

(55) *'o le vao s̄a tipi le tama*  
*K<sub>Top</sub> the grass IMP cut the boy*  
 'As for the grass, the boy was cutting it.'

(56) *le vao s̄a tipi le tama*  
*the grass IMP cut the boy*  
 'The grass that the boy was cutting.'

I therefore represent so-called Middles with the structure in (57)



If the roots in (51) may be suffixed with *-Cia*, they become regular Ergative predicates, assigning an Agent and Internal Argument theta role. These and some other examples are demonstrated in (58)

(58) **Derivation with -Cia**

alofa	x IS KINDHEARTED	alofagia	x TREATS y WELL
va'ai	x IS AWARE	va'aia	x SCRUTINISES/LOOKS AFTER y
agaleaga	x IS CRUEL	agaleagina	x TREATS y BADLY
ita	x IS ANGRY	itagia	x SHOWS DISLIKE TOWARDS y
sāuā	x IS CRUEL	sāuā	x PERSECUTES y

(59) *sā ita le fafine i le teine*  
IMP angry the woman  $K_{Obl}$  the girl  
'The woman was angry at the girl.'

(60) *sā ita-gia e le fafine le teine*  
IMP angry-CIA  $K_{Erg}$  the woman the girl  
'The woman shows dislike towards the girl.'

How are these examples modelled with a syntactic theory of morphology? The Internal Argument in the *-Cia* examples appears to relate semantically to the Oblique argument in the Middle constructions. The Agent argument in the *-Cia* constructions relates to the Internal Argument in the Middle Constructions.

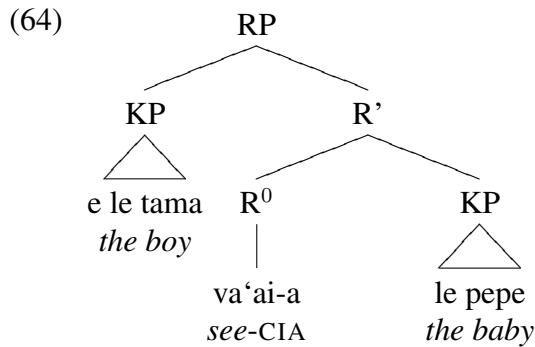
Say the *-Cia* suffix is generated at a structurally higher position. It does not introduce or delete any arguments in the Middle construction, but re-evaluates their theta roles. It must assign an Agentive theta role to one argument projected in the Middle construction. This is problematic for the model of theta roles established in this thesis. How is the assignation possible if arguments receive theta roles by being generated in certain structural positions? Boskovic (1994) proposes the model of movement into theta positions. Based on his model, the tree structure in (61) is generated. The morpheme *-Cia* is given the arbitrary label  $\delta$ . It projects an empty specifier which assigns the theta role of Agent to whichever argument occupies it. The Internal Argument of the root moves to fill it and take the Agent theta role.



The function of the *-Cia* suffix (in this instance) is to create an Ergative predicate. The Word Formation rule is represented by (63).

$$(63) \begin{array}{l} \text{ROOT} + \textit{Cia} \quad \rightarrow \text{ROOT-}\textit{Cia} \\ \text{UNACCUSATIVE} \quad \text{ERGATIVE} \end{array}$$

The newly formed Ergative root generates the established structure proposed throughout the structure for Ergative roots.



It is clear that the structure of the *-Cia* affixed Unaccusative cannot be derived from the middle construction. Rather, we must assume that the affixation of *-Cia* occurs outside of syntax. The process creates an Ergative verb, which then generates the “classic” structure proposed here for Ergative verbs. Theta role assignment, case assignment and incorporation are all correctly licensed in these positions.

## 6.4 Conclusion

This section models two theories of morphology to real phenomena occurring in natural language, the Syntactic model and the Lexicalist model. I assess the models based on their ability to account for the relevant syntactic behaviour of predicates. The Lexicalist hypothesis is untenable in accounting for Incorporation phenomena in Samoan. A derivational approach generates superior results,

accounting for examples where complex RPs are incorporated into the predicate.

In terms of the *fa'a-* and *-Cia* affixes, the crucial factor determining the applicability of the two models is theta roles assignment. The Internal Argument of the *fa'a-* causative appears to take its theta role from the root, demonstrating the syntactic behaviour associated with that theta role. For example, Unergative Agents are unable to incorporate. The Syntactic model of morphology allows roots to preserve their structures, therefore the theta role assignment of the root is unaffected by the affixation of *fa'a-*.

In contrast, the affixation of *-Cia* onto Stative roots strongly affects the theta role assignment of the Internal Argument of the Stative. It becomes a volitional Agent. The Syntactic model of morphology has great difficulty in formalising the assignation of Agentive properties onto the Internal Argument. The Lexicalist model has far greater success, generating the *-Cia* affixed stative as a regular Ergative predicate - demonstrating crucial properties of Ergative predicates such as incorporation and case assignment.

## 7 Conclusion

In this thesis I considered various theoretical perspectives on syntax and their applicability to the Samoan language. Specifically I looked at the syntactic behaviour of three predicate types, Ergatives, Unaccusatives and Unergatives. Aspects of their behaviour which proved particularly interesting in my analysis were their properties in relation to theta role and case assignment, event semantics, incorporation and their ability to take transitivity affixes such as *fa'a-* and *-Cia*.

I proposed a structure for the basic Samoan clause. In the derivation of the structure I discussed the fundamental Minimalist assumption of binary branching and proposed an instance where unary branching is more appropriate, namely the projection of an Agent by an Unergative. Otherwise, I found that binary branching more accurately captures certain facts about Samoan syntax problematic for a ternary branching structure.

I examined a number of other linguistic theories with the intention of assessing their effectiveness in modelling Samoan syntax. To model the Samoan Ergative-Absolutive case pattern, I gave an Optimality Theory analysis, formalising a number of constraints including one which stipulates that Absolutive case must be assigned to a non-pronominal argument projected by the root. I also deviated from Chomskyan assumptions in my use of the Lexicalist hypothesis to account for derivations with the *-Cia* suffix. The theta assigning properties of *-Cia* affixed Statives suggest that the argument structure of the root is not preserved syntactically.

Hopefully, this thesis will serve to raise certain questions about Samoan syntax. Further work must be undertaken to more fully understand certain syntactic phenomena. Particular interest must be paid to the *-Cia* suffix and its various uses. The analysis of its behaviour when affixed to predicates would benefit from a cross-linguistic survey of other Polynesian languages. Furthermore, other transi-

tivising and de-transitivising processes, not described in this thesis, are available in Samoan, such as the *ta-* causative and the *fe-i* circumfix. Investigations into these affixes should prove equally interesting.

I give a final note on objectives. The intention of a generative model of syntax (such as Minimalism and other theories described in this thesis) is to provide a machine which generates all and only the grammatical sentences of a particular language. An efficient generative model should avoid unnecessary complexities and ambiguities. This objective having been defined, any model of syntax proposed in this thesis does not claim to be a picture of the real-world cognitive ability of a Samoan speaker. The model only has the intention of replicating his or her output.

## References

- [1] Baker, Mark. 1988. *Incorporation*. University of Chicago Press. Chicago.
- [2] Ball, Douglas. 2005. "Tongan Noun Incorporation: Lexical Sharing or Argument Inheritance?". Proceedings of the HPSG05 Conference. CSLI Publications. Stanford.
- [3] Bittner, Maria. 1994. *Case, Scope and Binding*. Kluwer. Dordrecht.
- [4] Bittner, Maria and Ken Hale. 1996a. "The Structural Determination of Case and Agreement. *Linguistic Inquiry* Vol 27. No. 1.
- [5] Bittner, Maria and Ken Hale. (1996b). "Ergativity: Toward a Theory of a Heterogenous Class. *Linguistic Inquiry* Vol. 27. No. 4
- [6] Boskovic, Zeljko. 1994. "D-structure, theta criterion and movement into theta-positions". *Linguistic Analysis*. Vol. 24. No. 3. pp. 247-286.
- [7] Boskovic, Zeljko. 1998. "Scrambling and Last Resort". *Linguistic Inquiry*. Vol. 29. No. 3. pp. 347-366.
- [8] Carnie, Andrew and Eithne Guilfoyle. 2000. *The Syntax of Verb Initial Languages*. Oxford University Press. Oxford.
- [9] Chapin, Paul. 1970. "Samoan Pronominalization". *Language*. Vol. 46. No. 2. pp. 366-378.
- [10] Chomsky, Noam. 1981. *Lectures on Government and Binding*. Foris. Dordrecht.
- [11] Chomsky, Noam. 1995. *The Minimalist Program*. MIT Press. Cambridge.
- [12] Chomsky, Noam. 2000. "Minimalist inquiries: The framework". in *Step by Step: Essays on minimalist syntax in honor of Howard Lasnik* (Roger Martin, David Michaels and Juan Uriagereka, eds.) MIT Press. Cambridge.
- [13] Chomsky, Noam. 2001. "Derivation by Phase". in *Ken Hale: A Life in Language* (Michael Kenstovicz ed.) MIT Press. Cambridge. pp. 1-54.
- [14] Chomsky, Noam. 2004. "Beyond Explanatory Adequacy". in *Structures and Beyond: The Cartography of Syntactic Structures, Vol. 3* (Adriana Belletti, ed.) Oxford University Press. Oxford. pp. 104-131.



- [15] Chung, Sandra. 1978. *Case Marking and Grammatical Relations in Polynesian*. University of Texas Press. Austin.
- [16] Collins, Chris and Hoskuldur Thrainsson. 1996. "VP internal structure and object shift in Icelandic". *Linguistic Inquiry*, Vol. 27. pp. 391-444.
- [17] Collins, James. 2010. "Object Incorporation in Samoan". Essay submitted to the Department of Linguistics, University of Sydney.
- [18] Cook, Kenneth. 1996. "The Cia Suffix as a Passive Marker in Samoan". *Oceanic Linguistics*. Vol. 35. No. 1. pp. 57-76.
- [19] Di Scullio, Anna-Maria and Edwin Williams. 1987. *On the Definition of Word*. MIT Press. Cambridge.
- [20] Dowty, David. 1975. "The stative in the progressive and other essence/accident contrasts". *Linguistic Inquiry*. Vol. 6. pp. 579-587.
- [21] Dowty, David. 1979. *Word Meaning and Montague Grammar*. Kluwer. Dordrecht.
- [22] Dowty, David. 1991. "Thematic Proto-Roles and Argument Selection" *Language*. Vol. 67. pp. 547-619.
- [23] Fillmore, Charles J. 1968. "The Case for Case". in *Universals in Linguistic Theory*. New York (Bach and Harms ed.). Holt, Rinehart, and Winston. New York. pp. 1-88.
- [24] Fodor, Jerry. 1998. *Concepts: Where Cognitive Science Went Wrong, The 1996 John Locke Lectures* Oxford University Press. Oxford.
- [25] Goodluck, Helen and Michael Shaun Rochemont. 1992. *Island Constraints*. Kluwer. Dordrecht.
- [26] Grimshaw, Jane. 2008. "OT Syntax: an Interview with Jane Grimshaw". *Revista Virtual de Estudos da Linguagem*. Vol. 6. No. 10.
- [27] Hale, Ken, and Samuel Jay Keyser. 1993. "On argument structure and the lexical expression of syntactic relations". in *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger* (Ken Hale and Samuel Jay Keyser, eds.) MIT Press. Cambridge.

- [28] Hale, Ken and Samuel Jay Keyser. 1993. "On argument structure and the lexical expression of syntactic relations in *The View from Building 20: A Festschrift for Sylvain Bromberger* (Ken Hale and Samuel Jay Keyser, eds.) MIT Press. Cambridge.
- [29] Hale, Ken, and Samuel Jay Keyser. 2002. *Prolegomenon to a Theory of Argument Structure*. Linguistic Inquiry Monographs. MIT Press. Cambridge.
- [30] Halle, Morris and Alec Marantz. 1993. "Distributed Morphology and the pieces of inflection". in *The View from Building 20: A Festschrift for Sylvain Bromberger* (Ken Hale and Samuel Jay Keyser, eds.) MIT Press. Cambridge.
- [31] Jackendoff, Ray. 1990. *Semantic Structures*. MIT Press. Cambridge.
- [32] Kratzer, Angelika. 1996. "Severing the External Argument from its Verb in *Phrase Structure and the Lexicon* (J. Rooryck and L. Zaring eds.). Kluwer. Dordrecht.
- [33] Lakoff, George. 1966. "Stative adjectives and verbs in English". Report No. NSF-17: *Mathematical Linguistics and Automatic Translation* The Computation Laboratory. Cambridge.
- [34] Levin, Beth. 1983. "On the Nature of Ergativity". Doctoral Dissertation. MIT. Cambridge.
- [35] Levin, Beth and Malka Rappaport Hovav. 2005. *Argument Realization*. Research Surveys in Linguistics Series. Cambridge University Press. Cambridge.
- [36] Marantz, Alec. 1984. *On the Nature of Grammatical Relations*. MIT Press. Cambridge.
- [37] Massam, Diane. 2001. "Pseudo Noun Incorporation in Niuean". *Natural Language and Linguistic Theory*. Vol. 19. No. 1. pp. 153-197.
- [38] Milner, G.B. 1966. *Samoan Dictionary*. Oxford University Press. London.
- [39] Milner, G.B. 1973. "It is Aspect (Not Voice) Which Is Marked In Samoan". *Oceanic Linguistics*. Vol. 12. No. 1/2. pp. 21-639.
- [40] Mithun, Marianne. 1986. "On the nature of noun incorporation". *Language*. Vol. 62. pp. 32-38.

- [41] Mosel, Ulrike, and Evan Hovdhaugen. 1991. *Samoan Reference Grammar*. Scandinavian University Press. Oslo.
- [42] Pizzini, Quentin. 1971. "An Analysis of Predicate Complement Constructions in Samoan". *Oceanic Linguistics*. Vol. 10. No1. pp. 38-62.
- [43] Prince, Alan and Paul Smolensky. 2004. *Optimality Theory: Constraint interaction in generative grammar*. Blackwell. New Brunswick.
- [44] Pylkkänen, Liina. 2002. *Introducing Arguments*. Ph.D. Thesis. MIT. Cambridge.
- [45] Ramchand, Gillian. 2008. *Verb Meaning and the Lexicon*. Cambridge University Press. Cambridge.
- [46] Levin, Beth, and Malka Rappaport Hovav. 2005. *Argument Realization*. Research Surveys in Linguistics. Cambridge University Press. Cambridge.
- [47] Reinhart, Tanya. 2002. The Theta System An Overview. *Theoretical Linguistics* Vol. 28. pp. 229-290.
- [48] Scalise, Sergio and Emiliano Guevara. 2005. "The Lexicalist Approach to Word Formation and the Notion of the Lexicon". in *The Handbook of Word Formation* (Pavol Stekauer and Rochelle Lieber, eds.). Studies in Natural Language and Linguistic Theory. Vol. 64.
- [49] Stechow, Armin von. 1996. "The different readings of wieder: A structural account". *Journal of Semantics*. Vol. 13. pp. 87-138.
- [50] Talmy, Leonard. 1985. "Lexicalization patterns" in *Language Typology and Syntactic Description* (Timothy Shopen, ed.) Cambridge University Press. Cambridge.
- [51] Vendler, J. Van. 1957. "Verbs and Times". *Philosophical Review*. Vol. 56. pp. 143-60.
- [52] Woolford, Ellen. 2007. "Introduction to OT Syntax". *Phonological Studies*. Vol 10. pp. 119-134.