The Morphosyntax of 2P Pronouns in Tongan

Michael Dukes

University of Canterbury and Stanford University[1]

§ 1. Approaches to 2P Elements.

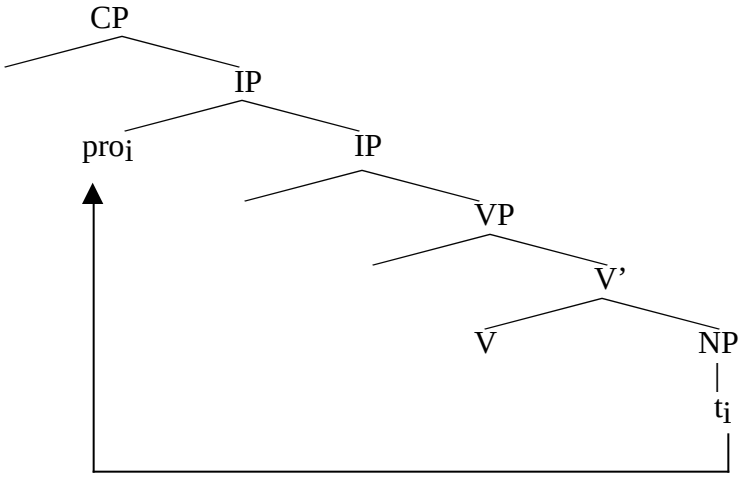
Many languages exhibit grammatical categories that are referred to as 'Second Position' (henceforth '2P') or 'Wackernagel' elements (Halpern & Zwicky 1996). These elements are typically functional morphs (often pronominal or auxiliary elements) which are phonologically weak and appear immediately after the first element of some constituent. Determination of what counts as the 'first element' may involve morphosyntactic factors, phonological factors or both.

Various proposals have been made in order to account for the language particular properties of pronominal and other 2P elements in a number of languages, often making appeal to special mechanisms of various sorts to derive their unexpected positional properties. Some of these mechanisms are reviewed below.

(i) **Movement-based accounts** (Kayne 1989, Fontana 1996):

Head movement along the spine of a clause or adjunction of an argument to a designated projection (aka 'clitic-climbing') is employed to describe the observed ordering facts. Such approaches are intended to explain the appearance of nominal elements in unexpected 'nonargument' positions while simultaneously characterizing their status as arguments, as illustrated in (1) below.

(1) **2P adjunction of a moved pronominal clitic.**



(ii) **Prosodic Inversion** (Halpern 1995):

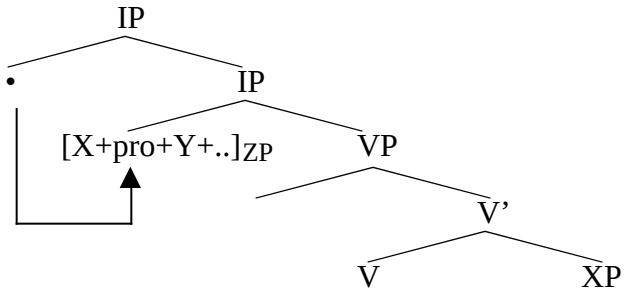
A post-syntactic process whereby a 2P element is embedded to the right of some 'first position' element, possibly even inside a constituent to the right of its first element.

[1] The research reported on in this paper was supported by grant UOC702 'The Representation of Tongan Syntax' from the Marsden Fund of the Royal Society to the University of Canterbury. I gratefully acknowledge helpful comments and questions from audiences at the Department of Linguistics, University of Hawaii at Mânoa, the HPSG2000 / Berkeley Formal Grammar conference at UC Berkeley and from participants in two courses at Stanford University in Spring Qtr, 2000 - Ivan Sag's 'HPSG2' course and Ling 225A 'Austronesian syntax'. I am indebted to Tolini Tonga for her help as a consultant on this project.

Proceedings of the 7th International HPSG Conference, UC Berkeley (22-23 July, 2000).
Dan Flickinger and Andreas Kathol (Editors).
2001. CSLI Publications. http://csli-publications.stanford.edu/

For syntactic purposes the 2P element is treated as external to the constituent into which it intrudes and is adjoined at the sentence level, as in (2) below.

(2) **Prosodic inversion of a sentence-level clitic into second position.**



(iii) **Extended inflection** (Miller 1991, 1993, Halpern 1995):

Extended inflection analyses have been employed to analyse elements often described as clitics, though to my knowledge this approach has not been applied specifically to the analysis of 2P elements. Under such approaches, morphosyntactic information carried by a phrase as a whole may filter down to some non-head element within it via constraints which determine that a feature must be realized as an 'edge feature' on some constituent. A 2P element could potentially be analysed as a left edge feature realized suffixally. Such an analysis would seem to be best restricted to elements that display a very low degree of autonomous wordhood, since the use of edge features implies the lack of an independent lexical category borne by the grammatical item in question. Furthermore, the analysis rests on finding reliable criteria for detemining the position of an edge element relative to some larger constituent on which it is parasitic.

The Polynesian language Tongan (like some of its close relatives, including Samoan, Tuvaluan, Tokelauan & East Futunan) displays pronouns that resemble 2P elements in many respects and which have been described as second position subject pronouns (Clark 1973, Chung 1978). It is argued here that, in contrast to what one might expect on the basis of the approaches sketched above, the grammatical properties of these pronouns can be shown to follow from standard theoretical machinery within HPSG, in particular, the standard set of phrase structure schemata and principles governing the satisfaction of valence requirements (Pollard & Sag 1994).

§ 2. An overview of the Tongan preverbal pronouns.

Tongan preverbal pronouns superficially display properties characteristic of 2P elements. They typically appear to the right of a clause-initial Tense/Aspect Marker (TAM), of which there are seven, as displayed in Table 1 below. As the table shows, the TAMs fall into two classes with respect to the distribution of preverbal pronouns. The non-past tense marker and one of the past tense markers display segmental allomorphy with respect to whether or not they are followed by a pronoun.[2] All other TAMs retain the same segmental shape whether or not a preverbal pronoun is present, though all bimoraic TAMs display stress shift when followed by a monomoraic pronoun. Illustrative examples can be found below.

When there is no TAM present in a sentence (these may be omitted under certain circumstances), a preverbal pronoun may occur following a conjunction word, of which

[2] The exact conditions on the choice between the two kinds of past tense marker are not entirely clear. Chung 1978 notes that the form *ne* appears to be associated with an 'as expected' presumption. This TAM is considerably less common in Tongan text than the alternating form.

the most common are listed in Table 2. In such cases, the pronoun again shows up in second position.

Table 1: Tongan Tense/Aspect Markers

Denotation		
	Nonpronominal form	**Prepronominal form**
Past Tense	*na'e*	*na'a*
Non–Past Tense	*'e*	*te*

	Neutral form
Past Tense	*ne*
Future Tense / Conditional Mood	*ka*
Progressive Aspect	*'oku*
Perfective Aspect	*kuo*
Nonindicative Mood	*ke*

Table 2: Tongan Conjunction Words

Denotation	**Conjunction word**	**May cooccur with TAM?**
'and (then), but'	*pea*	yes
'and, (in order) to'	*'o*	no
'lest'	*na'a*	yes (though rarely)
'if, when'	*ka*(*pau*)	yes (oblig. with *kapau*)
'but'	*ka*(e)	yes (*ka* with TAM)

When both a conjunction word and a TAM are present in the clause, the pronoun appears in third position.[3]

Basic word order in Tongan is typically described as VSO and generally obeys the simplified template given in (3).

(3) Conj TAM Pronoun Verb [ErgativeNP] [AbsolutiveNP]

Case marking in the postverbal domain follows an ergative/absolutive pattern. Ergative arguments are obligatorily marked with the preposition *'e*, while absolutive arguments are optionally marked with the preposition *'a*. However, the preverbal pronouns follow a nominative / accusative pattern and obligatorily crossreference the **ergative** argument of a 'transitive' predicate when that argument is pronominal. In all other cases, the preverbal pronoun crossreferences the **absolutive** argument of a predicate when it is pronominal. A postverbal pronoun crossreferenced by a preverbal pronoun is usually omitted, though it may be overt for purposes of emphasis. The following examples illustrate the basic pattern. The accents appearing on *na'a* and *pea* indicate stress shift in the presence of a following monomoraic pronoun. Note that Tongan orthography displays variation as to

[3] The particle *'o* displays a certain amount of ambiguity as to whether it is a conjunction word or a TAM. It seems likely that it must actually be assigned to both categories. Note that it is the only conjunction word listed in Table 2 that cannot cooccur with a TAM.

whether preverbal pronouns are written as separate words or as orthographic units with the TAM or conjunction.[4]

(4) (a) Na'a **mau**fu'u ongo'i 'aupito 'a e momokó ...
Past 1Pex Intns. feel Intns. Abs. Spec. cold
'We were really feeling the cold ...' (ergative argument) [MR]

(b) Na'á **ku** tatali ke hu'a 'a e tahi ....
Past 1Sg. wait Subj. be-in Abs. Spec. tide
'I waited for the tide to come in...' (absolutive argument) [MR]

(5) (a) .... pea **mau** ngâue 'aki 'a e Misiní ....
Conj. 1Pex work with Abs. Spec. engine
'... and we used the engine ....' (ergative) [MR]

(b) Peá **u** foki mai .....
Conj. 1Sg. return Dir1
'And then I went back ...' (absolutive) [MR]

(6) 'Oku **ou** talange ki ai, ...
Pres. 1Sg. say to him
'I say to him, ...' (absolutive) [CT]

(7) **Kau** 'alu mu'a ?
1Sg. go please
'May I go?' (absolutive) [Ch]

The forms of all the preverbal pronouns are listed in Table 3 below and the distribution of the first person singular allomorphs is outlined in Table 4. As can be seen in Table 3, only the first person exclusive singular displays any allomorphy.

Table 3: The Tongan Preverbal Pronouns [adapted from Churchward (1953:126)].

Person, Exclusive/Inclusive	**Number**	**Pronominal Forms**
1st Person Exclusive	Singular Dual Plural	*u, ku, ou, kau* *ma* *mau*
1st Person Inclusive	Singular Dual Plural	*te* *ta* *tau*
2nd Person	Singular Dual Plural	*ke* *mo* *mou*
3rd Person	Singular Dual Plural	*ne* *na* *nau*

[4] The following transcription conventions are used in the Tongan examples. Apostrophes indicate glottal stops, accents indicate stress shift and circumflexes indicate long vowels. Bracketted initials following each example indicate its textual source. Sources are listed at the end of this paper. The following abbreviations are used in examples: Abs.=Absolutive, Conj.=Conjunction, Dir1=first person directional ('towards speaker'), Dex.=Dual exclusive, Dl.=Dual, Erg.=Ergative, Fut.=Future, Gen.=Genitive, Intns.=Intensifier, Loc.=Locative, Neg.=Negation, NPst.=Nonpast, Perf.=Perfective, Pex.=Plural exclusive, Pl.=Plural, Pres.=Present, Sg.=Singular, Spec.=Specific, Subj.=Subjunctive.

There is a considerable body of evidence to indicate that the exact characteristics of the preverbal pronouns vary depending on which particular pronominal allomorph is present in the clause and which TAM it is adjacent to. In particular, unlike all the other pronouns, the first person singular pronoun exhibits four allomorphs depending on its context of occurrence.

Table 4: Distribution of the First Person Singular Preverbal Allomorphs.

Form	Contexts of occurrence	Examples in this paper
ku	Immediately follows past tense TAM *na'a* or conjunction *na'a* 'lest'.	(4)b, (9)
ou	Immediately follows the progressive TAM *'oku*.	(6), (15)b
kau	Occurs sentence-initially in concessive clauses.	(7)
u	Occurs in all other preverbal contexts. The default form.	(5)b, (8), (27)c

The free-standing first person singular variant *kau*, which occurs clause-initially in so-called concessive clauses ('Let me ...' , 'May I ...'), is almost certainly derived historically from a combination of the conditional/future TAM *ka* and the default pronominal form *u*, as the additional semantic contribution of this form suggests. All the other allomorphs are diachronically reduced forms of the Proto-Austronesian first person form **akú*.

There is also a fifth allomorph of the first person exclusive singular found in informal spoken Tongan which is not discussed in published grammatical descriptions of Tongan. This form is found in collocation with a reduced form of the past tense TAM *na'a*. Thus, in conversational contexts, *na'áku* is often reduced to *na'u*, as in (8).

(8) **Na'u** kai au 'i Christchurch peá**u** toki ha'u.
Past.1Sg. eat 1Sg. in Christchurch Conj.1Sg. next come
'I ate in Christchurch before I came.' [CT]

The morphosyntactic analysis of this form is discussed below since it is relevant to the issue of the constituency of TAM + pronoun units.

As noted in Table 4, the allomorph *ku* occurs following both the past tense TAM *na'a* and the homophonous conjunction word *na'a*, as seen in examples (4)b and (9).

(9) Na'e fekau 'e Kuku ia mo Kuku ke 'oua na'á **ku** hifo ki lalo.
Past order Erg. Kuku that with Kuku Subj. not lest 1Sg. descend to down
'Kuku and Kuku ordered me not to come down.' [TPF]

The issue of how to address this homophony is also addressed below.

Finally in this section I present a few examples illustrating preverbal pronouns occurring in initial position and in third position. Note that monomoraic forms are capable of standing word-initially despite the fact that are phonologically enclitic when following a TAM or conjunction. The 'third position' examples indicate that pronouns are expressed lower in the structure that one might have expected from the point of view of a strictly 2P analysis (though see section 6 for a counterexample).

(10) **Nau** ohu, ohu pea maha kae tekefili hake 'a e tuna ....
3Pl. scoop scoop Conj. empty Conj. exposed up Abs. Spec. eel
'They scooped and scooped until it was empty and the eel was exposed...' [TPF]

(11) **Na** a'u mai ki Vava'u ...
3Dl arrive Dir1 to Vava'u
'They got back to Vava'u...' [TPF]

(12) **Ta** ô mu'a ki Tonga ....
1Dex. go.Pl first to Tonga
'We will go first to Tonga...' [TPF]

(13) ...., he ne **ne** maumau'i 'ena fuakava ...
for Past 3Sg. break 3Dl.Gen. vow
'for he had broken their vow ..' [TPF]

(14) ..., pea kuo **nau** fai pongipongi, he kuo **nau** pehê pekia 'a e ta'ahine ....
Conj. Perf. 3Pl do morning for Perf. 3Pl. think die Abs. Spec. girl
'...and they did it in the morning for they thought that the girl had died...' [TPF]

In clauses where no TAM is overt, context determines the intended temporal role of a sentence.

§3. Evidence for and against constituency of the TAM + pronoun complex.

The conditions on the allomorphy of the pronouns and TAMs suggests that the 'bound' forms of these two categories are mutually selecting. An accurate formal characterization of this selection relies on an analysis of several finer grammatical distinctions discussed below. In essence, it is necessary to decide whether the pronouns should be analyzed as forming a syntactic constituent with the TAMs and conjunctions or whether they are simply phonological suffixes that form a phonological word with a preceding element. There is a substantial body of evidence pointing to the conclusion that the pronouns are generally autonomous syntactic elements. Conversely, there is also evidence pointing to a lexical treatment of the TAM/conjunction/pronoun complex in particular cases, namely those involving the first person singular exclusive allomorphs.

The first piece of evidence for the general syntactic autonomy of the preverbal pronouns is the fact that they can occur in at least three different contexts; either following a TAM or a conjunction (which must be treated as distinct syntactic types since they can often co-occur and generally have distinct functions in the clause, as noted in the tables above) or, less commonly, free-standing at the front of the clause. Since the free-standing forms will require an autonomous treatment anyway it seems pointless to start from the position that the pronouns require a host to their left in general. The opposite tack seems more justified.

The idea that the preverbal pronouns are, in the general case, syntactically autonomous is also supported by the facts of pronominal allomorphy. Only the first person singular pronouns show any sensitivity to the presence or absence of a preceding element. Certainly something special must be said about those cases but clearly they do not represent the typical situation.

Evidence from two kinds of allomorphy in the shape of the TAMs and conjunctions seems consistent with either position, though it perhaps offers more support to the constituency hypothesis.

The segmental allomorphy of the past and future tense TAMs is sensitive simply to the presence or absence of any following pronoun, independent of its phonological shape. This indicates that the shape of these two particular TAMs is indeed conditioned

morphosyntactically. On the other hand, the occurrence of stress shift on a TAM or conjunction in the general case depends purely on the phonological property of whether the following pronoun is monomoraic or not. Given the fact that stress falls consistently on the penultimate syllable of a lexical word in Tongan, it seems clear that the occurrence of stress shift on Tongan TAMs and conjunctions is simply a consequence of the fact that the phonologically weak pronouns form a prosodic word with the TAM. Thus the analysis of stress shift must be based on an account of how the pronouns form a prosodic word with the TAMs and conjunctions, so that if the pronouns are indeed analysed as morphosyntactically autonomous from the TAMs and conjunctions there must be a limited degree of mismatch allowed between morphosyntactic and phonological structure. The fact that the monomoraic pronouns display a low degree of selectivity with respect to their phonological host (i.e. attaching to a preceding TAM or a preceding conjunction) also supports a mismatch between phonology and morphosyntax.[5]

Given that the two first person singular 'bound' allomorphs *ku* and *ou* undergo phonologically unpredictable segmental changes, it seems that this allomorphy must be treated in one of two ways. Either *na'á ku* and *'oku ou* must be analysed as lexicalized inflectional complexes or the idiosyncratic forms *ku* and *ou* must be allowed in some way to select the TAM or conjunction that they cooccur with. A similar choice of options is available for the idiosyncratic form *na'u.* I will opt for the lexical approach, treating the subject pronoun as being incorporated into the TAM in these cases, as there appears to be no independent evidence for the lexicality of *ku* and *ou* given their rather limited distribution.

§ 4. Problems for movement-based approaches to Tongan clitics.

Movement-based analyses of pronominal clitics generally take the position that these elements raise from their normal argument positions (Kayne 1989) or are generated as heads which trigger the fronting of a (possibly null) pronoun from an argument position (Sportiche 1992). Such approaches run up against a number of difficulties with the Tongan data.

Firstly, a Tongan preverbal pronoun may be freely 'doubled' by an ordinary postverbal pronoun (though not by a nonpronominal NP), suggesting that nothing has moved out of the normal argument position (see (15) and also (8) above).

(15)(a) Te **ke** 'alu **koe** pea 'e 'alu mo Sitani ...
NPst. 2Sg. go 2Sg. Conj. NPst. go too Stan
'You will be going and so will Stan ..' [LMU]

(b) He 'ikai te **u** tamate'i **'e** **au** koe.
for Neg. NPst 1Sg. kill Erg. 1Sg. 2Sg.
'For I will not kill you.' [TPF]

Secondly, as can be observed in the above examples, the preverbal pronouns form a distinct paradigm from the postverbal pronominal series. The latter are either ergative or absolutive arguments of the predicate, while the preverbal series is not marked by any case particles. As noted above, the preverbal pronouns arguably instantiate a nominative category despite the ergative / absolutive case marking system found in the postverbal domain. This suggests, contrary to the expectations of a movement-based approach, that the preverbal pronoun is not associated with any particular case-marked position inside VP but rather with the thematically most prominent (i.e. 'agentive') argument, which I will refer to as the 'ARG-ST subject' (following Manning 1996).

[5] The treatment of phrasal phonological constituency is not further discussed here. See Asudeh 2000 for discussion of this issue from an HPSG perspective.

A third problem for a movement-based approach to the Tongan preverbal pronouns is that there is no obvous way to account for the cases in which the pronouns have merged morphologically with the TAMs save by adopting an analysis which appeals to rules of post-syntactic morphological fusion (such as proposed in Halle & Marantz 1993). Appeal to such processes severely weakens the generality of the analysis and renders any discussion of lexicality vs affixhood meaningless.

§ 5. Problems for prosodic inversion and extended inflection.

The motivation for invoking a process of prosodic inversion seems to be lacking in the case at hand. In particular, while one might argue that there is some prosodic motivation for the attachment properties of some of the preverbal pronouns (i.e. the phonological weakness of the monosyllabic enclitics such as first person *u* and second person *ke*) there is no motivation for placing the preverbal pronouns in a higher position syntactically than the TAMs and conjunctions and lowering them phonologically into the second (or third) position. Nor do the preverbal pronouns 'break up' any obvious constituent, as one might have expected under a prosodic inversion account. The appearance of the preverbal pronouns in third position after a TAM (i.e. in preference to the conjunction word) also counts against an inversion story. In short, none of the usual criteria for a prosodic inversion account are evident in the Tongan case.

The prospects of applying an extended inflection account to the data seem equally dim. Tongan 2P pronouns exhibit a high degree of wordhood and involve 'bundling' of several grammatical features, a situation that is not typical of the kind of phenomena that edge features were designed to deal with. There seems to be no motivation for treating preverbal pronouns as the spelling out of a set of inflections on the edge of some phrasal constituent.

§ 6. A relation-based analysis of the data.

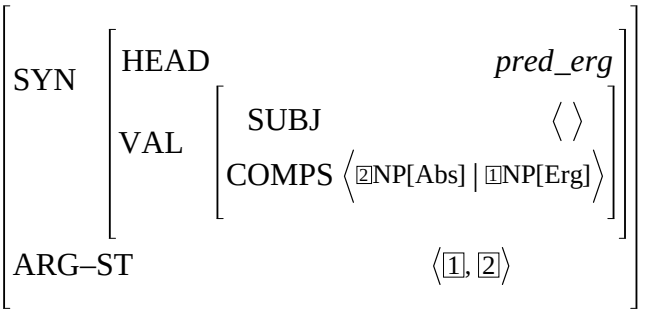
In the absence of much evidence for an account that seeks to treat Tongan preverbal pronoun behaviour via movement or subsequent postsyntactic rearrangement, it seems reasonable to attempt to develop a straightforwardly monostratal analysis which exploits an interesting split in the definition of subjecthood available in HPSG. Frequent 2P-hood in the Tongan case might just be a consequence of fairly ordinary grammatical machinery.

Following Borsley (1995) I assume that clauses in verb-initial languages fall into two classes depending on whether 'subjecthood', taken as a somewhat pretheoretical notion of grammatical prominence, is defined in terms of least obliqueness on the COMPS list of a predicate (as in older instantiations of HPSG (Pollard & Sag 1987, 1994)) or via specification of a nonempty value for the SUBJ feature. As noted above, a (presumably universal) notion of ARG-ST subject is also defined in least-obliqueness terms, providing us with three potentially competing notions of subject within HPSG.

I propose that Tongan (like Polynesian in general) be analyzed as a language in which full postverbal NP subjects are defined in terms of COMPS obliqueness (as in Borsley's analysis of Syrian Arabic). Thus, a full postverbal subject will be analysed as the first element on the COMPS list. Elements appearing in the postverbal domain will be uniquely associated with COMPS. This leaves the SUBJ value of the predicate empty in an ordinary clause with postverbal arguments, as in (16) below. An ergative verb (labelled *pred_erg*, a semantically 2-place verb which is allowed to coccur with an ergative NP) takes two full postverbal arguments and will have the partial category in (17) in which the SUBJ value remains empty.

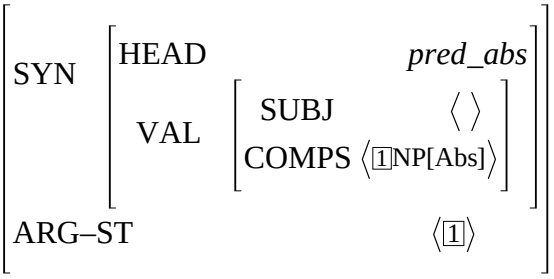
(16) Kuo lau 'e Siale 'a e tohí ni
Perf. read Erg. Abs. Spec. book this
'Siale has read this book'

(17) **Partial category for an ergative (semantically two-place) agentive predicate:**



A one-place absolutive predicate such as *'alu* 'go' (labelled *pred_abs*), which may not cooccur with an ergative NP, will have the partial category in (18).

(18) **Partial category for an absolutive (semantically one-place) predicate:**



As can be observed in the lexical entries above, I uniformly treat the **absolutive** argument as the least oblique element on the COMPS list of Tongan verbs. The motivation for this analysis is not directly relevant to the topic of this paper but revolves around the fact that absolutive NPs in the ergative languages of Polynesia display typical 'reference-related' subject properties (e.g. having a presupposed referent, controlling quantifier float, controlling (lexicalized) verb agreement, exhibiting unmarked wh-extractability, exhibiting (optionally in Tongan) unmarked case) which the ergative argument lacks despite its prototypical agency (Dukes 1998, Hooper 1999).[6] As is true of numerous other Malayo-Polynesian languages, Tongan exhibits a considerable degree of 'patient-prominence', implying that transitive agents are not the prototypically most prominent syntactic argument. As can be observed above, the absolutive argument is also the ARG-ST subject of a one-place predicate but not of a prototypically agentive predicate.

I adopt the idea that the preverbal pronouns instantiate the value of the SUBJ feature in those clauses in which they are present. The fact that the preverbal pronouns are neutral with respect to the ergative / absolutive case-marking distinction observed in the postverbal domain and refer to the most agentive argument of the main predicate using a series of pronouns distinct from the postverbal forms can be seen as a manifestation of nominative case in the value of SUBJ. For any predicate selecting full postverbal arguments (as in (8)) there will be a corresponding verb (of supertype *pred_nom*) that selects a nominative pronominal SUBJ which is coindexed with the (optional) pronominal postverbal ARG-ST subject (as in (9)). This relationship between

[6] See Schachter 1977 for discussion of the distinction between 'reference-related' and 'role-related' subject properties.

the nominative-taking predicates and the ordinary predicates taking full NP arguments can be formalized via the lexical rule given in (19).

(19) **Lexical rule relating *pred* to *pred_nom*:**

$$\left[\begin{array}{ll} \text{SYN} & \left[\begin{array}{ll} \text{HEAD} & \textit{pred} \\ \text{VAL} & \left[\text{SUBJ} \quad \langle\,\rangle\right] \end{array}\right] \\ \text{ARG–ST} & \langle \boxed{1}_i, .. \rangle \end{array}\right] \Rightarrow \left[\begin{array}{ll} \text{SYN} & \left[\begin{array}{ll} \text{HEAD} & \textit{pred\_nom} \\ \text{VAL} & \left[\text{SUBJ} \quad \langle \boxed{2}\textit{pro}[\text{Nom}]_i \rangle\right] \end{array}\right] \\ \text{ARG–ST} & \langle \boxed{1}\textit{pro}_i, .. \rangle \end{array}\right]$$

The postverbal NP corresponding to the ARG-ST subject may be ergative or absolutive depending on whether the verb is ergative or absolutive. Thus the relationship involved is one of argument addition (valence-increasing) that obligatorily requires the added SUBJ argument to share the interpretation of the ARG-ST subject. In (20) for example, the preverbal pronoun doubles the optional postverbal ergative.

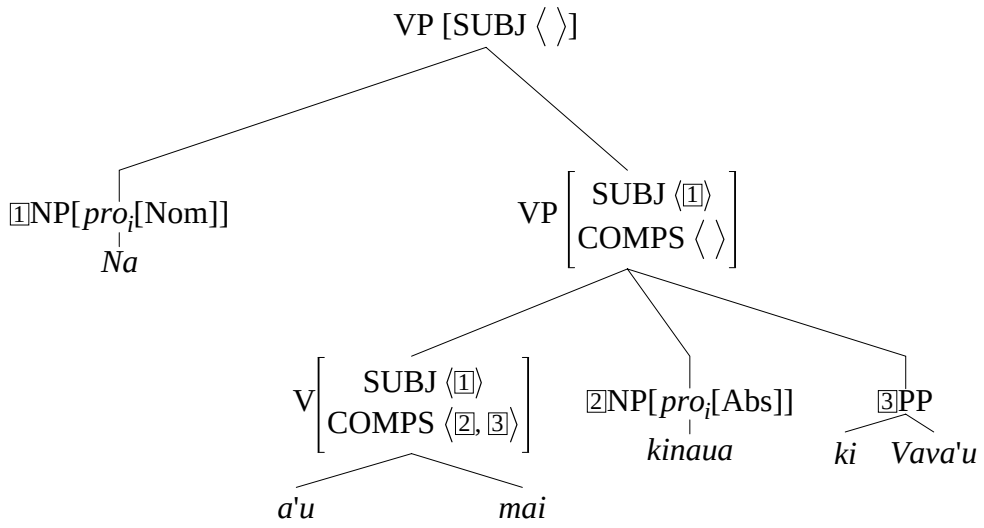
(20) Kuo **ne** lau ('e ia) 'a e tohí ni
Perf. 3Sg. read Erg. 3Sg. Abs. Spec. book this
'He has read this book'

There are several appealing consequences of this approach to subjecthood and preverbal pronouns in Tongan. Aside from the fact that the nominative case analysis of the SUBJ pronoun correlates nicely with the nominative target for the rule, the positioning of the preverbal pronouns is now straightforwardly assimilated to the standard Head-Subject Schema of Pollard & Sag 1994 (ch. 9), in which the value of SUBJ is treated as a kind of 'external argument'. Thus there is no need to invoke an otherwise unmotivated language-specific rule of preverbal pronoun placement (nor, by the way, is there a need to invoke the sometimes postulated Head-Subject-Complement Schema (Pollard & Sag 1994) to account for the verb-initial word order in clauses with full NP arguments)). The preverbal pronouns are 'base-generated' under the TAM without recourse to any corrective procedures. Thus a sentence, such as the modified version of (11) found in (21) will have the straightforward structure given in (22).[7]

(21) Na a'u mai (kinaua) ki Vava'u.
3Dl[Nom] arrive Dir1 3Dl to Vava'u
'They got back to Vava'u.'

[7] Note that the sequence [a'u + mai] is treated as a complex verb. The issue of verb constituency is not discussed further here. Note also that the locative modifier is treated as an (optional) complement.

(22)



Under this proposal, Tongan is treated as a language with 'defective' configurational subjecthood that is restricted to pronouns. This analysis accords nicely with its observable ergative properties.

I propose to treat the TAMs and conjunctions as lexical heads of category I(NFLECTION) and C(OMPLEMENTIZER) respectively. This analysis provides general syntactic categories for what are semantically diverse collections of competing particles. However, it also leads to a nonuniform description of Tongan sentence-types, which may be VP, IP or CP under this approach.

In the case of those TAM allomorphs which require the presence of a following pronoun (i.e. *na'a* (category I, 'past tense') and *te* (category I, 'non-past tense')), it is necessary only to stipulate that they select a complement VP with a head of type *pred-nom*, as in (23).

(23)

$$\begin{bmatrix} \text{HEAD} & I\_prepron \\ \text{VAL} & \begin{bmatrix} \text{SUBJ} & \langle\,\rangle \\ \text{COMPS} & \left\langle \text{VP} \begin{bmatrix} \text{HEAD} & pred\_nom \\ \text{SUBJ} & \langle\,\rangle \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

Thus only a VP with a nominative argument will occur as complement of the prepronominal TAMs. The nonpronominal TAMs, which by contrast disallow preverbal pronouns (i.e. *na'e* and *'e*) will be restricted to subcategorizing for a complement that is not of type *pred_nom*.

The appeal to selection via HEAD types above is crucial because there is no other way for an I or C to check whether its complement has an expressed SUBJ, assuming the correctness of the proposed structure. However, in the cases in which the first person

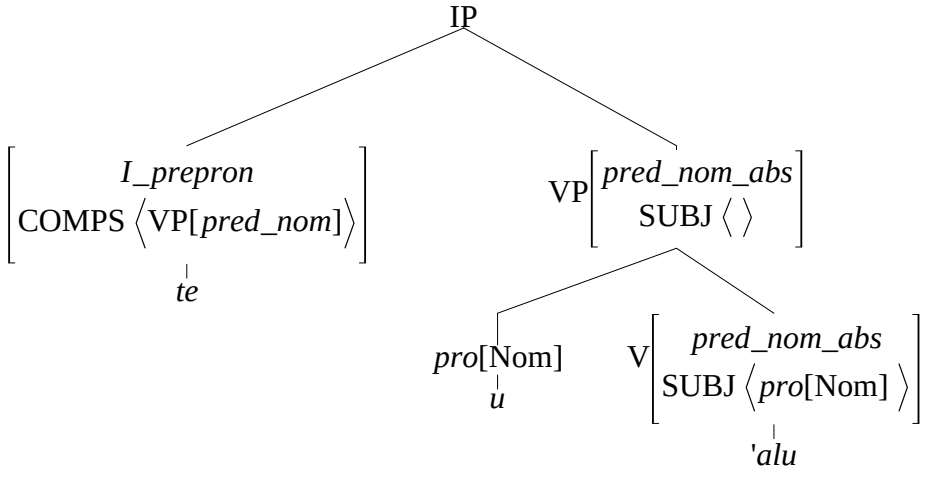
singular pronouns *ou* and *ku* are idiosyncratically selected by the TAMs *'oku* and *na'a* (or the conjunction *na'a*) it is necessary for the TAM or conjunction to be able to check whether or not the dominated VP SUBJ is first person or not. So the type-theoretic approach will not work in this case. Fortunately, this breakdown in the more general analysis correlates with the fact that these allomorphs (including the conversational register form *na'u*) are morphologically idiosyncratic and so suggests that they can be treated as incorporated material modifying I or C. More specifically, these forms will be elements of category I which have a composed category seeking a complement VP that has an undischarged first person singular SUBJ value. Thus, the entry for *na'aku* will specify that it is an I whose complement VP still seeks a first person singular SUBJ (as opposed to one that is fully discharged). The (partial and abbreviated) entry for *na'aku* is given in (24) and is identical in selectional respects to that for *na'u* or *'okuou*.

(24)

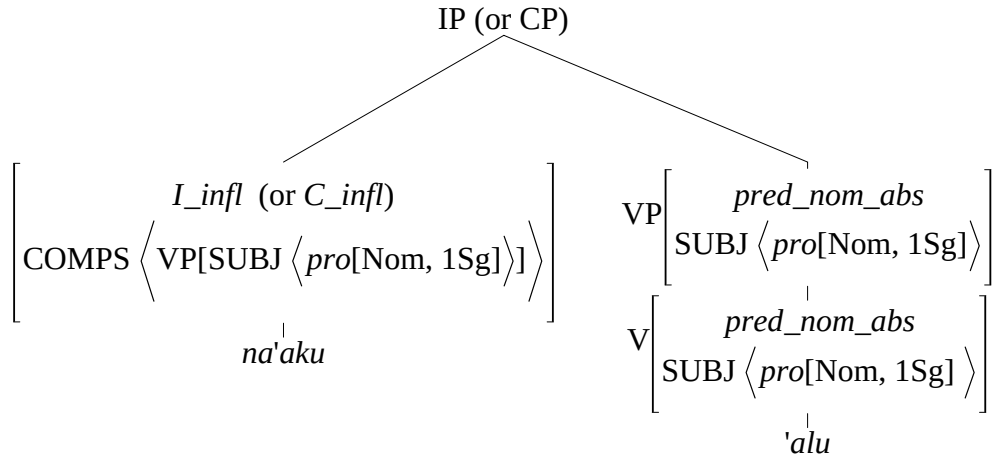
$$\begin{bmatrix} \text{PHON} & \langle \text{na'aku} \rangle \\ \text{HEAD} & \textit{I\_infl} \quad (\text{or } \textit{C\_infl}) \\ \text{VAL} & \begin{bmatrix} \text{SUBJ} & \langle\,\rangle \\ \text{COMPS} & \left\langle \text{VP} \left[\text{SUBJ} \ \langle \textit{pro}[1\text{Sg}] \rangle \right] \right\rangle \end{bmatrix} \end{bmatrix}$$

Note that the entry in (24) actually creates a subjectless structure in which the missing subject is ‘absorbed’ into the head of Infl. Thus, under this analysis, the structure of *Te u ‘alu* ‘I will go’ in (25) differs markedly from that of *Na’aku ‘alu* ‘I left’ in (26).

(25)



(26)



Tongan orthography appears to reflect the fact that speakers sometimes think of the TAM+pronoun sequences as one complex word and sometimes as two separate words, probably indicating incipient morphologization of the preverbal pronoun. In any case, a compositional approach of the sort proposed above seems to be ideal in general for treating diachronic reanalysis of originally autonomous adjacent words into single lexical entities.

An issue that remains to be addressed is how to block the occurrence of such sequences as *na'a u* and *'oku u* where the default pronominal form appears in a context where a particular idiosyncratic allomorph should appear. Since *na'a+ku* and *'oku+ou* are listed in the lexicon it might be possible to appeal to the idea that these more restricted listed forms block the appearance of the phrasal units bearing the same information, though on the face of it this appears to be a somewhat poorly motivated appeal to blocking since the TAM+ pronoun sequence does not form a phrasal constituent under the approach developed here. What appears to be required for a more adequate account of this problem is to allow for the possibilty of 'inside-out selection' so that the form *u* specifies that it cannot occur with the past tense TAM. Under the analysis presented here this could only be achieved via allowing for a kind of type-raising that lets the pronoun select the predicate of which it is an argument. A more promising alternative in my view is to rethink the idea that the TAMs and conjunction words are heads. Treating them as modifiers of the predicate (of type MARKER for example) would allow more straightforwardly for the possibility that the predicates could allow interactions in the form of their arguments and selected modifiers.[8]

§ 7. An interesting wrinkle.

The analysis presented above works rather straightforwardly with no recourse to (post)syntactic reordering. However, there is one intriguing problem that requires some additional machinery. Tongan exhibits a common conjunction word *ka* 'if, when', which is (perhaps) a shortened form of *kapau*. After *kapau*, the normal 'future' marker is *'e* or *te*. After *ka* however, the future marker is also realized as *ka*, in which case a preverbal pronoun, when present, must occur **between** the two homophonous particles, as seen in (27).[9]

[8] Yet another possibility is to take seriously the idea that the TAM+pronoun units form a constituent uniformly. Indeed, the data seems potentially compatible with any of these approaches under certain choices of axioms.

[9] The lengthening of the pronominal vowels in (27)c is the result of emphatic stress, not due to the existence of a morphologically specified alternative form.

(27)(a)... pea ka **nau** ka 'ilo, ...
Conj if 3Pl. Fut. know
'But if they find out, ...' [TPF]

(b) Ka **ke** ka ha'u 'oku 'atâ 'a to'ohema, ...
if 2Sg. Fut. leave Pres. free Abs. left
'If you leave, the left will be free ...' [Ef]

(c) Pea ka **u** ka nofo au pea û maumau'i 'e au pea û mate au.
Conj. if 1Sg. Fut. stay 1Sg. Conj. 1Sg. break Erg. 1Sg. Conj. 1Sg. die 1Sg.
'But if I stay and I break (the pledge) then I will die.' [TPF]

Thus, in this context, the preverbal pronoun unexpectedly 'skips' over the TAM and intervenes between the two homophonous particles. The skipping only occurs when both these particular particles are present. Thus the reordering seems to motivated by a restriction on sequencing two homophonous morphemes (c.f. the 'Repeated Morph Constraint' (Menn & MacWhinney 1984)) or (perhaps more likely) is motivated by historical changes relating to the replacement of *ka* by *te* as a future/nonpast tense marker. It appears in fact from examination of corpus material that the *ka...ka* sequence rarely occurs without an intervening 2P pronoun (though *ka 'o ka* is also possible and both are listed in Churchward 1959). A prosodic inversion analysis is not motivated since this again is not a true 'second position' phenomenon (note the appearance of the pronoun in third position in (27)a and (27)c).

From the point of view of the analysis presented so far, there does not seem to be any particularly appealing way of accounting for the data in (27). One could treat any sequence $ka_1+pro+ka_2$ as a lexical constituent, presumably a C with a composed category seeking a VP missing a subject with the person features specified by the pronoun. But the lexicon will then require 12 such forms, one for each pronoun. Since both ka_1 and ka_2 can independently take a following pronoun, it would be necessary to appeal to grammatical (or lexical) blocking to prevent the sequence *ka+ka+pro*, which should otherwise be grammatical. One could also appeal to a linearization rule operating on the relevant elements either in the syntax or the phonology, but this would require us to treat these elements as part of the same ordering domain just in this single case.

A tentative approach to dealing with this data along selectional lines runs as follows. It is possible to provide the 'future tense' *ka* with a raising category which selects a nominative pronominal subject and sets its complement VP as having an unsaturated SUBJ value structure-shared with the subject of *ka*. The conditional conjunction *ka* can be provided with a secondary category which specifies that it selects an IP headed by 'raising' *ka* just in case the sentence as a whole is associated with future time interpretation. This is information which can be listed in the set of contextual restrictions associated with conditional *ka*. The pronoun will thus be realized as the subject of future tense *ka* just in case it is preceded by the conditional conjunction *ka*.

The plausibility of this analysis rests on evidence for treating INFLECTION as a SUBJ selecting category. From a crosslinguistic perspective this is obviously not an unusual situation and can be seen as involving the partial development of an auxiliary-like category for Tongan *ka*. It seems peculiar in the case at hand because Tongan does not otherwise have any TAMs which raise their complement subject. However, as we will see in the next section, there is evidence for a more general raising category for one of the TAMs found in Samoan, a close relative of Tongan.

§ 7. Variations in the analysis of related languages.

The analysis of Tongan 2P pronouns presented here must allow variation in at least two respects in order to account for similar data in related languages. **Ordering variations** between TAM and pronoun are required and variation in the **grammatical target** for the rule that licenses preverbal pronouns is also necessary.

As hinted at in the previous section, the use of a raising category for one TAM in Samoan appears to be necessary. The data from Samoan is generally very compatible with the analysis proposed here with the same ordering relations holding between TAM and pronoun. However the 'general' (Gnr.) TAM *te* (which is a kind of default TAM cognate with Tongan *te*) obligatorily requires a raising analysis of the sort proposed for Tongan *ka*. Unlike all other Samoan TAMs, *te* always **follows** rather than precedes the preverbal pronoun, as illustrated in the examples in (28) (from Mosel & Hovdhaugen 1992:364-5).

(28)(a) Tou **te** le fia aai i uluvai...?
2Pl. Gnr. Neg. want eat Loc. shrimp
'You do not want to bite at the prawns?'

(b) ..., ma **te** le toe fia vaai i le auivi ...
1Dex. Gnr. Neg. still want see Loc. Det. skeleton
'..., we didn't want to see the skeleton again...'

As in Tongan, the form *te* generally alternates with the form *e*, which occurs in contexts lacking a preverbal pronoun. However in Samoan it is also possible for both the TAMs to cooccur in the same sentence when there is also a preverbal pronoun present. In such cases, the two TAMs bracket the pronoun in identical fashion to the *ka ... ka* sequence in Tongan.

(29)(a) **E** ta **te** feiloai foi i luma o le Atua ...
Gnr. 1Dinc. Gnr. meet again Loc. front Gen. Det. god
'We will meet again in front of God.'

(b) **E** matou **te** falemoe faatasi ...
Gnr. 1Pex. Gnr. sleephouse together
'We slept in the same house ...'

Thus the kind of strategy for treating these cases of bracketting adopted in Tongan can also be applied to Samoan where it will be necessary to have at least two categories for the TAM *e*, one in which it cannot cooccur with a pronoun and another in which it must occur with the pronoun when the pronoun itself is followed by *te*. The doubling prepronominal category for *e* may be derived via category shift from I to C, thus accounting for its clause initial position and the occurrence of a following TAM.

A more challenging set of difficulties for the analysis of 2P pronouns proposed here is brought up by the second kind of variation noted at the start of this section. The valence-based approach to Tongan 2P pronouns rests on the assumption that the preverbal pronoun is a (nominative) SUBJ. However, Polynesian languages with cognates of the Tongan pronouns exhibit considerable variation in the constraints on their expression.

The data from **Samoan** broadly supports the kind of approach outlined here. The lexical rule deriving the preverbal pronoun has a nominative target, just as in Tongan, and differs only in the following respects. Whereas the preverbal pronoun in Tongan increases the valence of the predicate by adding a coreferential SUBJ, the rule in Samoan must obligatorily delete the least oblique element from the COMPS list when the

pronominal SUBJ is present. Thus in Samoan there is no pronominal doubling of the sort observed in Tongan. Furthermore, unlike Tongan, the rule of pronominal fronting is optional so that pronominal subjects may occur postverbally. Finally, when pronominal fronting of an ergative argument occurs it is typically (though not obligatorily) accompanied by the appearance of the so-called 'transitive suffix' on the verb, arguably an overt indicator of the change in grammatical functions determined by the verb (Chung 1978).

On the face of it, the facts of **Tokelauan** are somewhat more problematic for the account presented here. In this language, pronominal fronting is restricted exclusively to ergative arguments and is obligatorily accompanied by the appearance of the transitive suffix on the verb (Hooper 1999). The fact that the rule is organized ergatively raises problems for the nominative approach to SUBJ assumed here. However, there appears to be good evidence for thinking that the pronominal fronting rule in Tokelauan is grammatically quite distinct from that observed in Tongan and Samoan. Apart from the difference in grammatical target, the Tokelauan process involves fronting of full pronominal forms, not just reduced 'clitic' pronouns of the sort observed above. The likeliest hypothesis for explaining this difference is that Tokelauan is in the process of losing its older reduced forms and the grammatical constraints that govern their appearance, subsequently innovating a new 'topicalization' process involving unreduced pronouns restricted to ergative arguments. Loss of reduced preverbal pronouns has also occurred in several other parts of Polynesia, especially Eastern Polynesian (e.g. Maori) and in Niuean (Clark 1973). The obligatory appearance of verbal morphology is probably a key factor in the retention of the fronting of ergatives.[10]

The data from East Futunan present yet another twist on the story. According to Biggs 1974 and Moyse-Faurie 1997, preverbal pronouns are not restricted solely to logical subjects. While logical subjects may be fronted as in Tongan and Samoan, absolutive pronouns corresponding to logical objects may also be fronted, as in (30) (examples from Moyse-Faurie 1997:16-17).

(30)(a) E **kau** tamate e koe.
Gnr 1Sg. hit Erg. 2Sg.
'You hit me.'

(b) E **ke** 'u'uti e le kulî mokâ ke 'aga o sa'u lana ne'akai.
Gnr 2Sg. bite Erg. Det. dog when 2Sg. face C remove his food
'The dog will bite you if you try to remove his food.'

Although this possibility is synchronically restricted to East Futunan, it seems possible that it reflects the situation that was prevalent in Proto-Polynesian since the possibility of having fronted patients is also observed in Fijian which still allows a patient pronoun to be realized clause initially in typical pronominal subject position (Dixon 1988). Thus East Futunan and Fijian arguably reflect a situation in which the feature SUBJ has a more independent grammatical role from the ARG-ST subject than is observed in Tongan and Samoan, while still being restricted to having a pronominal expression. Further investigation may shed light on the diachronic changes involved in the role of SUBJ in Polynesian and on the typology of valence features in the world's languages as analysed within HPSG.

[10] The facts from the southern dialects of Tuvaluan are broadly similar to those in Tokelauan, requiring obligatory morphology on the verb and being restricted to ergative pronouns. Northern Tuvaluan dialects however bear more resemblance to Samoan since verbal morphology is not obligatory (Besnier 2000).

Textual Sources

Fanua, Tupou Posesi. 1976. *Po Fananga: Folk Tales of Tonga.* Nuku'alofa: Friendly Islands Bookshop. [TPF]
Fonua, Pesi. 1983. *La'â mo 'Uha* (Sun and Rain). Nuku'alofa: Vava'u Press. [LMU]
Hafoka, Loseli (ed.). 1992. *Minerva Reef.* Tonga: Friendly Islands Bookshop. [MR]
Moala, Masiu. 1994. *Efinanga.* Nuku'alofa: LALI Publications. [Ef]
Rogers, Garth (ed.). 1986. *The Fire Has Jumped.* Suva: USP. [FHJ]
Conversational Texts, collected and transcribed by Tolini Tonga. [CT]

References

Asudeh, Ash. 2000. A formalization of precompiled phrasal phonology. Ms; Stanford.
Besnier, Niko. 2000. *Tuvaluan*. Routledge; London & New York.
Biggs, Bruce. 1974. Some problems of Polynesian grammar. *Journal of the Polynesian Society* 83:401-426.
Borsley, Robert. 1995. Some similarities and differences between Welsh & Syrian Arabic. *Linguistics* 33:99-122.
Chung, Sandra. 1978. *Case marking and grammatical relations in Polynesian*. University of Texas Press; Austin.
Churchward, C. Maxwell. 1953. *Tongan grammar*. Oxford Universoty Press; Oxford.
Churchward, C. Maxwell. 1959. *Tongan dictionary*. Govt of Tonga; Nuku'alofa.
Clark, Ross. 1973. Transitivity and case in Eastern Oceanic languages. *Oceanic Linguistics* 12: 559-605.
Dixon, R.M.W. 1988. *A grammar of Boumaa Fijian*. University of Chicago Press; Chicago.
Dukes, Michael. 1998. Evidence for Grammatical Functions in Tongan. In M. Butt and T. Holloway King (eds.). 1998. *Proceedings of the 1998 Conference on Lexical Functional Grammar*. CSLI; Stanford.
[Online at: http://www.sultry.arts.usyd.edu.au/LFG98/austro/dukes/dintro.htm].
Fontana, Josep. 1996. Phonology and syntax in the Tobler-Mussafia law. In Halpern & Zwicky (eds). 1996.
Halle, Morris & A. Marantz. 1993. Distributed morphology and the pieces of inflection. In Ken Hale & Jay Keyser (eds) *The View from Building 20*. MIT Press; Cambridge, MA.
Halpern, Aaron. 1995. *On the Placement and Morphology of Clitics.* CSLI; Stanford.
Halpern, Aaron & Arnold M. Zwicky (eds). 1996. *Approaching Second*. CSLI; Stanford.
Hooper, Robin. 1999. Revisiting the subject: properties of ergative and absolutive arguments in Tokelauan. Paper presented at the 4th International Conference on Oceanic Linguistics, Niue.
Kayne, Richard. 1989. Null subjects and clitic climbing. In Osvaldo Jaeggli &Ken Safir (eds), *The Null Subject Parameter*. Dordrecht: Reidel.
Manning, Christopher. 1996. *Ergativity: argument structure and grammatical relations*. Stanford: CSLI.
Menn, Lise & B. MacWhinney. 1984. The repeated morph constraint. *Language* 60.3.
Miller, Philip H. 1991. *Clitics and Constituents in Phrase Structure Grammar*. PhD dissertation, Rijksuniversiteit, Utrecht.
Miller, Philip H. 1993. Morphological marking misses the head. In J. Mead et al (eds), *Proceedings of the 11th West Coast Conference on Formal Linguistics*. CSLI; Stanford.
Mosel, Ulrike &Even Hovdhaugen. 1992. *Samoan reference grammar*. Scandinavian University Press; Oslo.
Moyse-Faurie, Claire. 1997 Syntactic and pragmatic functions of pronominal arguments in some western Polynesian languages. *Oceanic Linguistics* 36.1:6-28.

Pollard, Carl & Ivan A. Sag. 1987. *Information-based syntax and semantics, volume 1*. CSLI; Stanford.
Pollard, Carl & Ivan A. Sag. 1994. *Head-driven phrase structure grammar*. University of Chicago Press; Chicago.
Sag, Ivan & Thomas Wasow. 1999. *Syntactic Theory: a formal introduction*. Stanford; CSLI.
Schachter, Paul. 1977. Reference-related and Role-related Properties of Subjects. Peter Cole & Jerrold Sadock , eds, *Syntax and Semantics 8: Grammatical Relations.* Academic Press; New York.