Clause Structure and Argument Realization in Tongan
A Dissertation Proposal

Douglas L. Ball

September 5, 2007

1 Introduction

1.1 Background

In the most recent comprehensive study in comparative Polynesian syntax, Chung (1978) examined various classical syntactic phenomena (raising, passive, control, and others), looking to better understand the distribution and asymmetries among the kinds of verbal arguments. In theorizing about the these argument-categories, she considered several possible places where the correct generalizations could be found. First, she considered the possibility that these various syntactic phenomena were sensitive to (surface) linear order. If this were to turn out to be the correct kind of generalization, then one would need to find evidence for phenomena that picked out NP₁ over NP₂ from (1):

(1) Verb NP₁ NP₂

Second, Chung considered the possibility that certain phenomena are sensitive to case-marking, the view found in traditional Polynesianist grammars. Thus, in ergative languages (as many Polynesian languages are), one would expect rules to be sensitive whether a given NP was marked by the absolutive preposition, the ergative preposition, or any of a host of other, more adjunct-like prepositions.

Finally, she considered the possibility that certain phenomena are sensitive to (accusatively-based) grammatical relations. In some cases, especially in the accusative languages of Eastern Polynesian, this view is difficult to differentiate from the case-marking viewpoint. However, especially in the ergative Samoic and Tongic, this proposal is easier to see independent of case-marking, because, here, there is a salient gap between the two. The case-marking hypothesis will predict that S (intransitive subjects) and O (second arguments of transitive verbs) arguments will pattern together while the (accusatively-oriented) grammatical relations hypothesis will predict that A (transitive subjects) and S (intransitive subjects) arguments will pattern together. And, in fact, Chung concluded that the latter was the case, citing the accusative basis for raising and control as well as from pronominalization behavior. That this evidence converged from both ergative (e.g. Tongan
and Samoan) and accusative languages (e.g. Māori) led her to further conclude that the abstractness level of grammatical relations was justified for all the Polynesian languages she examined.

Turning to one of these languages and its phenomena, let us look at Tongan. Tongan, as alluded in the above, is one of the ergative Polynesian languages. Among the phenomena that Chung proposes (in ch. 4) are, in fact, sensitive to the notion of subject and object, includes the Tongan case-marking pattern. It is a subject-sensitive rule, Chung claims, in that it marks the subject in clauses with two unmarked direct case NPs. However, given that such a statement makes key reference to a clause having two unmarked direct case NPs raises the question: how grounded in grammatical relations are the case determinants (rules or constraints)? Chung’s formulation would seem to suggest that case-marking is ‘parasitic’ on grammatical relations, but of course there are two other logical possibilities: (1) grammatical relations are ‘parasitic’ on case-marking (essentially the traditional Polynesianist view) or (2) grammatical relations and case-marking are separate (a whole family of views that would conceivably include the denial of grammatical relations). The open question remains which of these views offers a better account of the grammar of arguments in Tongan (and other Polynesian languages).

Furthermore, to account for the additional case-marking pattern of the so-called ‘middle verbs’, which have their semantically most prominent argument marked by the absolutive preposition and their second argument marked by either a locative or dative preposition, Chung proposes that middle verbs are marked as [+middle] in the lexicon and there is a special case-marking rule that takes [+middle] verbs as input and gives out the appropriate ‘middle frame’ (with either the locative or dative preposition). These case proposals raise two important questions about how Tongan case should be dealt with: (1) how idiosyncratic are the middle verbs? and (2) what interaction, if any, does there need to be between the determinants of case in Tongan? With regards to (1), Chung’s solution seems to treat the presence of this ‘middle frame’ as being a pattern without a generalization and, taken to a logical extreme, a randomly occurring ‘frame’. So, is this actually true, and if it’s not, then what sorts of factors is the notion of [+middle] standing in for? The question in (2) deals with more formal concerns, but is a question related to how a theoretical architecture for this phenomenon is developed: what kind of interactions, if any, should the case determinants have?

Chung also notes that certain phenomena – such as relativization and clefting/topicialization – are not sensitive to ‘deep’ grammatical relations but rather to the case-marking. In these particular constructions, the relevant patterning involves how the argument is realized at the foot of the potentially unbounded dependency between the filler and the lower argument position. The generalization appears to be that absolutive arguments lack overt realization while obliques (including, usually ergatives) require resumptive pronouns. This patterning raises yet another question related to the above: what kind of theory will allow for the correct interactions between these case-marking and these so-called more superficial phenomena?

Furthermore, this separation of phenomena into those sensitive to ‘deep’ relations and those sensitive to case-marking again returns us to the question of how ‘deep’ is case-marking. While Chung doesn’t take a strong stand about the “level” of case-markings in her 1978 work (though she offers some ideas), many others have. Certain current (and older) theories of syntax (such as mainstream generative grammar in numerous different incarnations) has treated case as a
relatively superficial property; in contrast, others have taken case to be tied closely to deep structure or deep structure’s analog, the semantics (such as in Charles Fillmore’s Case Grammar (Fillmore 1968)). Thus, the open question is which of these views can better account for the empirical phenomena?

In trying to motivate the necessity of ‘deep’ grammatical relations, Chung focuses more on those phenomena that support a subject grammatical relation at the expense of object-related phenomena. In fact, among the Tongan phenomena, there are two that she claims are as uniquely sensitive to objects (and not just core arguments): noun incorporation and quantifier float. Given this small amount of ‘object-referring rules’, as Chung calls them, one might wonder just how relevant the notion of ‘object’ may be in Tongan grammar. Certainly, these two phenomena require further investigation to see if the notion of ‘object’ is the right one for them. One important question is whether some notion of ‘object’ can adequately capture the generalizations surrounding these phenomena. (The other important question is that, if it can’t, what can?)

Also, if more phenomena could be added to the list of at least potentially object-related, this could provide further testing grounds for understanding the relevance of the notion ‘object’. One such relevant phenomenon (discussed first for the related language Niuean by Seiter 1980) involves the word ‘aki. ‘Aki marks what are semantically instruments, but differs from the kinds of prepositions mentioned above in that its complement looks like an absolutive nominal expression, complete with the possibility for an absolutive preposition (Contrasting with ‘aki, the other prepositions govern what look like DPs; that is, determiners plus nominal expressions). Furthermore, ‘aki can, alternatively, appear within the verbal complex. When it does, its notional complement is discontinuous from it, and this complement is coded like and behaves like any other absolutive-marked nominal. Thus, these arguments, seeming to share some properties with other, more seemingly prototypical objects, appear to be yet another source for understanding how objects (or perhaps, more properly “objects”) work in Tongan.

The universe of analytical possibilities for dealing with the distributional and asymmetrical properties of arguments, as outlined above, is much larger than it was in 1978. Beyond just surface/linear position, tree-theoretic approaches now give the option of quite a few positions within the tree that argument properties could conceivably be sensitive to. And among the possibilities is to re-interpret grammatical relations as being just positions on a tree, thus collapsing the linear order and grammatical relations approaches (a move foreshadowed by the seeming indeterminacy between them in chapter 4 of Chung 1978).

However, it is not just the tree-theoretic approaches that have been enriched. Alternative conceptions of the argument-predicate relations, stemming ultimately from more dependency-based approaches to grammar and extending earlier ideas about subcategorization frames, have been developed in many constraint-based theories, including Relational Grammar (RG), Lexical-Functional Grammar (LFG), Role and Reference Grammar (RRG), and Head-driven Phrase Structure Grammar (HPSG). In this last framework, theorizing in the last ten years has led to the positing of a thematic level of representation called argument structure.\(^1\) While this level does not posit

---

\(^1\)Although this corresponds more closely to what others have termed valency, actancy, and even transitivity.
grammatical relations in the same way as RG or LFG (i.e. nothing is labeled “subject” or “object”), it still imposes prominence relations between the arguments, as well as serving locus of interface between the semantics on the one hand and the phrasal constructions that make up the syntax on the other hand.

The question that arises from these different analytical possibilities is what kind of theory will yield the right kinds of representations and interactions between them. In this work, I look to explore this question, taking as my working hypothesis that a more dependency-based approach (as embodied in a framework with a thematic level representation and co-present semantic representations) can offer a more descriptively adequate and more principled account of the argument patterns of interest, discussed above, from Tongan.

The Main Claims of this Dissertation: The argument realization and, in fact, clause structure itself are driven by the lexically-mediated combinatorical potentials. Furthermore, the Tongan clause is far more minimal (i.e. flatter) than previous analyses (even Dukes 2000) would lead one to believe and that flatness does not require concomitant recognition of lots of grammatical functions (positionally or primitively) – instead, the work can be more economically parceled out to case, ordered lists of verbal arguments, the semantics proper, and, at most, one privileged syntactic argument.

1.2 Tools for an HPSG analysis

In framing my dissertation in HPSG (see Pollard and Sag 1994; Ginzburg and Sag 2000; Sag, Wasow, and Bender 2003), that affords me as the analyst with certain tools. In general, I will be assuming the sign-based architecture of HPSG, where levels of phonology, morphology, syntax, semantics, and pragmatics exist and potentially constrain each other in parallel. I will also assume a type hierarchy, or signature, whereby the constraints of the grammar are factored into a hierarchy, disallowing impossible feature combinations, and allowing for greater and smaller generalizations to be captured in a uniform way.

The particular HPSG-based tools of dealing with the argument realization phenomenon are the following:

ARG-ST: The ARGument STructure list (see Manning and Sag 1998). Though a SYN(tax) attribute, it interfaces with the semantics. It is the locus of binding relations and of unrealized arguments, such as “incorporated” pronouns and gaps. In some ways, it is like the initial stratum of Relational Grammar.

Case: This is feature of nominal expressions (NPs), though actually won’t be called CASE in the description language, due to the fact that Tongan case is not morphological. This is among the many categories that governing elements can “see” when they select for certain dependents.

XARG: A newcomer to the HPSG scene, the eXternal ARG(ument) (see Sag 2007). Though motivated by locality concerns, this feature is a way to translate the notion of ‘privileged
syntactic argument’ from Role and Reference Grammar into HPSG. Items taken to be the value of the XARG feature have generally been highly topical, prototypically left-periphery, items.

**VAL:** The VALence list. This actually won’t play too much of a role in the core part of the analysis, but will be present in diagrams as a consequence of the ARG-ST. This list deals with the surface syntax; items like “incorporated” pronouns and gaps do *not* appear on this list. It is very much like the final stratum of Relational Grammar.

**Semantics:** The whole of the semantics is not involved; however, individual predicates will be. They will either be represented à la predicate calculus (like \(eat'(e, x, y)\)) or using the notation of minimal recursion semantics, like in (2).

\[
\begin{array}{c}
\text{eat-rel} \\
\text{SIT } s_1 \\
\text{ARG1 } x \\
\text{ARG2 } y \\
\end{array}
\]

1.3 **The Plan**

- The first three chapters deal with the foundation for the later analysis and offer initial reasons for the need to recognize ARG-ST with regard to Tongan grammar. The first two deal with the clause structure I’m proposing, and the last with case.
  - Chapter 2 is on the lower part of the clause, dealing largely with a putative VP
  - Chapter 3 is on the upper part of the clause, dealing largely with the TAM word and its pronominal affixes (also brings up the first of several argument realization challenges)
  - Chapter 4 looks at the case patterns in canonical clauses and provides an analysis for them that will form the basis for the three case studies below

- The rest of the dissertation is three chapters that detail three case studies that (1) support or are consistent with the structure I’m proposing and (2) develop further the idea that various interactions among arguments must be dealt with on argument structure or in the semantics entirely, not syntactically. The case studies are:

  1. Noun Incorporation
  2. The Instrumental Particle, ‘Aki
  3. Quantifier Float with Kotoa, ‘all’
2 On the VP

**Issue:** Does Tongan have a traditional VP, grouping verb and internal arguments together?

**Supplemental Issue:** Can Tongan word order be derived from head movement of the verb or VP-remnant movement?

2.1 In Search of Positive Evidence for the VP

The first question: Is there any positive evidence from the usual tests for VP-hood?

Looking at the usual suspects

- Displacement of a Putative VP
  - *ko*-fronting with a putative VP looks initially to be possible
  - However, closer inspection reveals the structure is nominal, not in alternation with verbal structure, and is not isomorphic to a traditional VP

- Coordination
  - Coordination of traditional VPs is impossible

- Ellipsis
  - Problems arise for any ellipsis account that generates a VP, then deletes it

2.2 Problems Solved by a Flat Structure

- Scrambling – can be handled by linear precedence constraints. The locality of scrambling falls out of the fact that ID-rule still enforces locality. The A vs. A’-movement debate surrounding movement-based approaches to scrambling is rendered moot because this is ordinary head-valent combination and there is no movement

- Adverb Placement – can also be handled by linear precedence constraints. The degree of freedom in their syntax is unified with the freedom exhibited by NPs in a simple way

- Typological Considerations – this eliminates the need to treat verb-initial languages as special. Under mainstream X-bar-based, binary branching analyses, verb-initial languages have to be derived by movement. On this analysis, verb-initial languages are just as undervived as SVO languages or SOV languages, and operate much like them, combining heads with arguments to form larger units.
3 On the Upper Part of the Clause

Issue: What is the syntactic relationship between the clause-initial words and the PredP motivated in section 2?

Supplemental Issue: How are the “clitics” related to their governing predicate?

3.1 Flat Structure All The Way?

• No – TAM + clitics are a higher constituent that takes a PredP as its complement. Why?
  – Coordination data: PredP’s coordinate just fine
  – Problems for a Verbal Complex analysis:
    * This would create strange low conjunction coordination structures
    * Preverbal adverbs would necessitate very complex, multiword verbal complexes
      and might complicate statements about adverb distribution

3.2 “Clitics”

• Assuming a more configurational structure means that “clitics” are not locally realized with respect to their governing predicate. How then do we handle this relationship?

3.2.1 Affixes or Prosodically Weak Words?

• One solution: Clitic Movement

• However this predicts that “clitics” are syntactically independent. Is this true?

• No, there is a lot of evidence for the morphologization of “clitics”
  – Stem Selection with na’elna’a – clitics select the latter, while the latter is used for non-clitic instances
  – Allomorphy with 1SG – the form of the morph realization various depending on the TAM word
  – Stress Placement – stress treats TAM + single mora clitics as one word
  – Doubling
  – Coordination
    (though coordination data appears not to be so persuasive with 2-morae clitics, so it’s possible that maybe all clitics aren’t uniformly affixes)
3.2.2 2 Possible analyses of Clitics

Proposal 1: Copy Raising

- “Clitics” are related to downstairs predicate by copy raising
- Key component of the analysis: the XARG feature
- Synopsis of Analysis: Inflected initial selects a PredP with a XARG value that coin-
dexed with the clitic (coindexation is bolded) as in (3):

(3) 
\[ \begin{array}{cccccccc}
\text{inflected-initial} & \left[ \begin{array}{c}
\text{ARG-ST} \\
\text{proj} \\
\text{CAT} \\
\text{XARG} \\
\text{XP} \\
\text{SEM} \\
\text{IND} \\
\text{j} \\
\end{array} \right] & \end{array} \]

- Gets “Doubling”: Through underspecified XARG | CAT value (the XP), the lower
  argument could be unrealized pronoun (≈ PRO) or it could be a realized pronoun
  (as in the doubling case)
- Gets pronoun restriction on “doubling” – through restriction on semantics to items
  of type pro-obj
- Gets A + S orientation of pronominal clitics through XARG value, which only
  allows A’s and S’s (this needs to be made more precise how this works)

Proposal 2: Raising

- “Clitics” and downstairs predicate related by run-of-the-mill raising
- Inflected initials have the following lexical entry

(4) 
\[ \begin{array}{cccccccc}
\text{inflected-tam} & \left[ \begin{array}{c}
\text{ARG-ST} \\
\text{proj} \\
\text{PredP} \\
\text{VAL} \\
\end{array} \right] & \end{array} \]

- To Get “Doubling”: A special VAL list (5) must be created

(5) 
\[ \begin{array}{cccccccc}
\text{VAL} & \left[ \begin{array}{c}
\text{proj} \\
\text{NP[erg]} \\
\end{array} \right] & \end{array} \]

- Gets pronoun restriction on “doubling” – through binding on special VAL list
- Gets A + S orientation of pronominal clitics because raising is always subject-
  oriented (Is this really true?: It seems with an undifferentiated VAL list, the anal-
  ysis predicts that raising of objects could be possible. Thus, this analysis would
  need some auxiliary notion to work out – either a VAL list divided into SUBJ and
COMPS or a requirement that clitics are actually CASE:NOM and must unify with CASE:NOM argument “from” downstairs)

3.2.3 A Testing Ground: Clitic Climbing with \textit{ka}

- The two proposals for clitics seem quite similar and are difficult to distinguish
- However, there is one domain where I believe they can be distinguished: Clitic Climbing with \textit{ka}
- The Phenomenon:
  - In certain conditional clauses, the clause begins with what is thought to be a conjunction, \textit{ka}. This word is then followed by another \textit{ka} (The second one is made out, by Churchward, to be of the same category as the TAMs), and a(n) (incorporated) pronoun intervenes between the two \textit{ka’s}
- The predictions:
  - The copy raising analysis predicts that \textit{ka} clitic \textit{ka} clitic ... should be possible – multiple copies are possible
  - The raising analysis predicts only \textit{ka} clitic \textit{ka} ... should be possible (raising only realizes the argument once)
- Preliminary investigations support the copy raising analysis: \textit{ka} clitic \textit{ka} clitic is possible

3.2.4 Why Should “Clitics” raise when other arguments don’t?

Good question – perhaps historically-motivated (or functionally-motivated)

4 Case in Tongan

Issue: As case is configurationally regulated in mainstream generative theories, if Tongan has a flat structure, how is case for nominal expressions (NPs/DPs) ever licensed?

Answer: On the ARG-ST list of words

4.1 Background – How is Case Marked in Tongan?

- Case in Tongan is not marked in the canonical Indo-European style, with suffixes, but with prepositions
- 2 Choices for an Analysis – Does either make a difference?
1. Treat some/all as “weak heads”, altering the **MARKING** value
2. Treat them all as prepositions, contributing their own semantics, and all verbs just select for PPs, with particular **LEXICAL-ID/PFORM** values

### 4.2 Arguments for treating case uniformly and in the lexicon

#### 4.2.1 Morphologically-contingent Case-Marking: Verbs with -(C)i suffix

- The first argument comes the interaction of case-marking and morphology
- Many Tongan verbs can appear unaffixed (bare) or with the suffix -(C)i.
- All verbs, when suffixed with -(C)i, appear with the ERG–ABS frame, even if the base verb took a different frame
- -(C)i is clearly morphological
  - The suffix always attaches to verbs stems, and is immovable
  - The C is lexically determined. There is a default consonant, the glottal stop, but many stems select their own thematic consonant (often, it’s the consonant that was historically the stem-final consonant in Proto-Austronesian, but not always)
- If the morphology affects case-marking, and morphology is pre-syntactic, then it follows that case-marking must also be determined pre-syntactically.
- The morphology notwithstanding, an analysis that tries to do the case-marking with -(C)i verbs syntactically runs into one of two problems for non-ERG–ABS base verbs, either:
  1. It must license all “quirky cases” syntactically as well (next section will offer arguments against doing that)
  2. It must employ some sort of look-ahead mechanism to ensure that right case is selected downstairs – renders moot checking theory and would seem to compromise a constrained sense of locality

#### 4.2.2 “Quirky” Case

- Tongan has a fair amount of ABS–OBL verbs (but hardly any OBL–ABS verbs, in contrast to many Eurasian languages)
- On a GB analysis (or earlier), these would be considered lexical exceptions and it would be stipulated (where this occurred was never made precise) that these verbs govern “quirky” cases
- This sort of account is undesirable for 2 reasons
1. We would like a uniform theory of case assignment—it’s unparsimonious to have case sometimes being regulated by the syntax and sometimes by the lexical entry.

2. Their exceptional quality is overstated—the “exceptional” classes in Tongan follow patterns of lexical classes in other languages.
   - Interaction verbs: ABS–DAT in Tongan. NOM–DAT in German (and other caseful IE langs), Hebrew, others?
   - Perception verbs: ABS–DAT in Tongan. Unattained perception is NOM–DAT in Turkish, others?
   - Pursuit verbs: ABS–DAT in Tongan. Also NOM/ABS–DAT in other languages? (maybe Chukchi)
   - Feeling Predicates: ABS–LOC in Tongan. Also NOM/ABS–LOC/OBL in other languages?

- Conclusion: These are not random “exceptions”, but regular patterns that are part of the diversity of argument realization in natural languages.

- I argue that these should be lexically accounted for, not because of their “exceptional” nature, but because they rely on lexical classes that need to be independently recognized in the lexicons of natural languages. More to the point, they rely on characteristics of what in lexical semantics are called “roots,” that is on rather specific and lexeme-particular aspects of the meaning.
  - Not correlated with telicity
  - Not precisely correlated with event template

- Also, they should be accounted for lexically because of the following problem, which shows that there is no syntactic evidence for treating these exceptions as syntactically special.

**The Infeasibility of Treating “Quirky Case” as a Generalized OBJ function**

- A possible out: Tongan is like Icelandic, but instead of having all sorts of differently marked cases mapping to a general subject grammatical relation (Andrews 1982), Tongan has lots of cases that map to a general object grammatical relation.

- Implausible, because absolutive objects and oblique objects do not pattern the same way.

There are few available object tests, but:
  - Absolutive objects can float the quantifier *kotoa*, ‘all’; oblique objects cannot
  - In “extraction” contexts: oblique objects pattern with other obliques and require a resumptive pronoun; absolutive objects require a gap
Furthermore, first arguments of verbs with oblique objects don’t behave as if first argument has been shifted to a different category (“subject”) – they behave as if they are still absolutive arguments (in extraction and in floating *kotoa*).

### 4.2.3 An Analysis

- Case is a part of linking – it is done on the portal with the lexical semantics: the ARG-ST list.

- Case is licensed via linking constraints, taking properties of semantic predicates and mapping them into morphosyntactic categories (Doing linking Koenig and Davis-style [see Koenig and Davis 2003; Koenig and Davis 2006] means hierarchically partitioned verbal relations will be constrained to be associated with certain kinds of ARG-ST lists).

- Though partial constraints can be placed on high supertypes within the verbal relations type hierarchy, the cases will not be fully specified until the word’s ARG-ST list (To do: make sure this is formally enforced).

- This is done because various morphological additions can effect the semantics, which in turn can effect the linking (such as in the instances of the *-Ci* suffix, noted above).

- The treatment of case on the ARG-ST list is also important because of LDDs:
  - Particular cases correlated with particular kinds of LDD-foot realizations: absolutive → gaps, just about everything else → resumptive pronouns.
  - However, assuming a traceless account of LDDs (following others in the HPSG literature, such as Bouma, Malouf, and Sag 2001), case could not be done on a VAL list because higher arguments of dyadic predicates are still in the ergative even when there is a gap for the absolutive. If case were done on the VAL list, a transitive verb with an absolutive gap would look intransitive and so the higher argument would appear in the absolutive. If case were done on the ARG-ST list, there would still be two arguments on the ARG-ST list, ensuring the correct kind of linking.

- The question, then: How to implement these correlations within HPSG? A lexical rule (a so-called word-to-word rule)? Special gap/resumptive pronoun linking rules? Freely allowing gaps to appear on ARG-ST lists, supplemented with other, hopefully independently-motivated, constraints on their distribution?

### 4.2.4 A Comparative-Historical Sidebar

- There is also a historical argument for doing case on a list instead of by tree position (configurationally).

- In relatively recent times, the case-marking system of some Polynesian languages changed (the direction of change is controversial – in other work, I have supported an accusative to ergative change, but the argument here holds regardless of the direction).
Among the Polynesian family, certain languages (mostly of the Eastern Polynesian subgroup) are accusative; other languages (entirely in the Tongic and Samoic-Outlier subgroups) are ergative.

If case were done configurationally, this would mean that these languages would have had to radically reconfigure their syntax to get the right positions for case checking/assignment.

There is little evidence of a wholesale change in syntax (intraclausal syntax is pretty much the same across the family, modulo the different marking patterns).

If this was a change on an argument structure list, this would be a relatively minor change—and fit with the overall variation we see in Tongan (and other languages); argument realization is not necessarily uniform across all lexical classes and all kinds of lexical constructions.

5 Noun Incorporation

Phenomenon: Transitive (and some extended intransitive) verbs alternate with an intransitive instantiation, which has a noun (sometimes with modifiers) right next to the verb.

(6) a. Canonical Clause
   Na'e inu 'ā e kavā 'e Sione.
   PST drank ABS DET kava ERG (name)
   ‘Sione drank the kava.’ (Churchward 1953, 76)

b. Clause with Incorporation
   Na'e inu kava 'a Sione.
   PST drink kava ABS (name)
   ‘Sione drank kava.’ (Churchward 1953, 76), my emphasis added

Issue: Does Noun Incorporation necessitate a VP?

Answer: No, it can be handled with ordinary compounding and compounding-like semantics—no VP involved; no ARG-ST item involved.

5.1 The Morphosyntax of NI

5.1.1 Data

NI in Tongan is unusual, given the received typological wisdom on NI

- The construction is apparently detransitivizing (external argument is in the absolutive) → would appear to be a case of Compounding NI (in the terms of (Rosen 1989))
- but adjuncts can appear and modify the incorporated noun, such as the phrase mo e laise in (7)
(7) ‘Oku ne kai kale mo e laise.
PRS 3SG eat curry COM DET rice
‘He eats curry and rice.’ (Churchward 1953, 115), my emphasis added

- So, it might be a Stripping Construction (a construction of a verb + bare NP), not compounding (as suggested by Massam 2001)

- But the evidence for a Stripping Construction is not so good:
  - Incorporated noun + adjuncts can’t scramble
  - Incorporated noun + adjuncts must appear inside of postverbal particles
  - The kicker: Incorporated noun + verb can nominalize together (Incorporated noun + verb + adjuncts cannot)

∴ Better to treat the NI as a compound and deal with the adjuncts a different way

5.1.2 Problems for VP-remnant Movement and Head-Movement accounts

- VP-remnant Movement
  - Doesn’t capture the lexicality of Tongan NI – the adjacency of the verb and noun is happenstance (this analysis needs, for instance, auxiliary assumptions to rule out stranding of classifier-like adjectives)
  - If case is checked in-situ, not clear how the external argument (in spec, vP) will avoid getting ergative case (i.e. how is the intransitivization effect of the construction analytically actualized)

- Head-Movement
  - The behavior of Tongan modifiers poses a problem – head movement predicts there should be Mohawk-style stranding, when there isn’t
  - The ordering of the modifiers inside the postverbal particles compounds this problem – to capture the facts, there has be both VP-remnant movement and head-movement – an unparsimonious account
  - Suffers from many issues to be discussed in 5.2

5.1.3 My analysis

Part I: Verb + Noun is a lexical construction (I’m following the compounding NI analysis of Runner and Aranovich 2003)

- Tongan NI is a compounding structure: a Verb + Noun compound
The non-scrambling of incorporated noun, the location of the postverbal particles, and the nominalization data all falls out straightforwardly from this structure.

Semantically, the incorporated noun “restricts” the verb (whether by Restrict (Chung and Ladusaw 2004), Unification (Farkas and de Swart 2003), or something else [e.g. Carlson 2005]) and restricts the incorporated noun from even appearing on the ARG-ST list of the newly created incorporating verb.

With only 1 core argument on the incorporating verb’s ARG-ST, the verb will have to govern ABS case (on any linking theory for Tongan).

Thus, the case-marking falls out, too.

This predicts that NI verbs should behave as intransitives
- They do with respect to “extraction”
- They should with quantifier float (needs to be checked)

**Part II:** But what about the adjuncts? How do they fit in?

- The leading idea here: these are not too different from “classic” bracketing paradoxes like [transformational [grammar]ian] and [nuclear [physical]ist]
- Adjuncts are to be treated as syntactically independent phrases that combine with incorporating verbs and end up modifying the incorporated noun.
- Like nominal modifiers outside of this construction, they combine via the SEL(ect) feature.
- These adjuncts are specified as in (8)

\[
\begin{align*}
\text{SEL} & \leftarrow \text{CAT} \quad \text{universal} \\
\text{SEM} & \rightarrow \text{IND} \quad \text{REL} \\
\text{REL} & \leftarrow \text{nom-rel} \quad \text{IND}
\end{align*}
\]

- The CAT specification of *universal* means that these modifiers are underspecified in their category – they can combine with nouns (usually) or verbs (in this special case). (I use Māori grammarian Bruce Biggs’ term for this supertype of nouns and verbs)
- These modifiers are further specified to require a semantic nominal relation. With the proper identity with indices, the modifier comes out modifying the noun.

An outstanding question: how to ensure this doesn’t create the possibility for nominal adjuncts to attach higher.
5.2 The Lexical Semantics of NI

**Issue here:** What kind of relationship to the verb does a noun have to have to be licensed in noun incorporation?

- Assuming a syntactic approach, either VP-remnant movement or head movement, would suggest that NI is isomorphic to some particular syntactic category.
- With the advent of L-syntax (Hale and Keyser 1993), the notion of syntactic category on some analyses is much broader than it once was: it can include a number of lexical semantic categories.
- But a broad survey of the ability of over a 100 predicates to take incorporated nouns reveals that no syntactic category (even in this broader sense) is isomorphic to incorporanda.

5.2.1 Problematic Syntactic Categories

**Case:** Case turns out to be a rather bad predictor, as there are mismatches in both directions.

- A significant amount of second arguments taking oblique case do incorporate, e.g. dative arguments of perception verbs, dative arguments of searching/waiting verbs, the locative arguments of various verbs including *heka*, ‘ride’; *fonu*, ‘be full of’, *manako*, ‘want’
- But there are also dative and locative arguments that don’t incorporate (datives of interaction verbs, contact verbs, locatives of some psych verbs, datives and locatives of motion verbs)
- Similarly with absolutes: a vast majority of absolutes allow incorporation. But a few don’t: e.g. those absolutes of *motu(hi)*, ‘snap’; *u’u*, ‘bite’, *amo(hi)*, ‘pet’

**Generalized Object Grammatical Function:** Supposing that we combined some of the cases into a generalized notion of “logical object” or “undergoer” (as suggested by Dukes 1998), would the results be any better than just looking at the cases?

No, precisely for the same reasons that case was problematic: the splits in oblique objects and splits in absolutes suggest that generalizing does not help with the problem at all.

**Structural vs. Root Argument:** Another possible cut of the NI pie is structural arguments vs. root arguments. In lexical decomposition theories of lexical semantics, the structural arguments are those that occupy a position required by the event structure (i.e. the y argument in the following: [x ACT [CAUSE [BECOME [y CUT]]]). Root arguments, as their name suggests, are those permitted by the root, but not integrated into an event structure position. This distinction could easily be translated to a syntactic framework, whereby structural arguments are in specifiers, while root arguments are the complement of roots (i.e. items under radicals). However, this is still problematic in two ways.
1. Comparing across structural and root arguments, they can pattern together. E.g. the root argument of *holo*, ‘wipe’ and the structural argument of *tofi*, ‘cut’ can both be incorporated.

2. Within the two kinds of arguments, the behavior is not uniform: e.g. the root argument of *holo*, ‘wipe’ incorporates while the root argument of *tokoni*, ‘help’ does not.

**Atelic predicates:** Perhaps given the work on ‘measuring out’ and AspP, telicity is the key factor for NI.

- Problem #1: Aktionsart is a phrasal property, so it’s a bit unclear how to use it to determine whether particular verbs allow for noun incorporation.
- Even assuming a prototypical aktionsart classification, there are still problems:
  - As with the structural/root argument distinction, not all instances of a particular aktionsart category allow for incorporation, cf. *holo* and *tokoni* – both activities, but not both incorporaters.
  - And there is, once again, cross-category similarities in behavior: ‘*ilo*, ‘know’ – a state – allows for incorporation, just as many activities do.
  - So event type doesn’t seem to be the right way to go either.

Conclusion: A syntactic category approach is not right for Tongan NI. A new kind of approach to the lexical semantics of NI is needed.

### 5.2.2 My Analysis

- Ability to incorporate is licensed within a specific kind of part-whole relationship, similar to the relationship in other kinds of compounding.
- A first pass at an analysis: Those participants that lie outside the event (i.e. are autonomous from the event) are not licensed in incorporation.
- Many of the ungrammatical incorporanda are literal or metaphorical directed motion verbs – as the “goal” of these motion events is autonomous from the event (that is, these “goals” are not an intrinsic part of the event), these participants are not licensed as incorporanda.
- Extends to goals and benefactives, which are not incorporable.
- Manner incorporation is also quite common in Tongan. Licensed under this theory because manner is intrinsically always part of the event.
- Effects of Noun: A contrast exists between *ui kuli*, ‘call dogs’ and *ui manupuna*, ‘call birds’ – suggests that semantics of the noun itself plays a role in determining autonomy.
- Issue: How to make this more precise?
- Note that this is a purely semantic notion and very difficult to implement syntactically.
6 ‘Aki

Phenomenon: Prepositional ‘Aki alternates with Applicative ‘Aki (and the doubling construction) as in (9)

(9) a. Prepositional ‘Aki
   Na’e tofī ‘a e mā [‘aki ‘a e hele].
   PST cut ERG (name) ABS DET bread INST ABS DET knife
   ‘Sione cut the bread with the knife.’

   b. Applicative ‘Aki
   Na’e tofī ‘aki ‘e Sione [e hele] e mā.
   PST cut INST ERG (name) DET knife DET bread
   ‘Sione cut the bread with the knife.’

   c. Na’e tofī ‘aki ‘e Sione e mā [‘aki e hele].
   PST cut INST ERG (name) DET bread INST DET knife
   ‘Sione cut the bread with the knife.’

Raison d’être: As the only double absolutive construction in Tongan (aside from maybe a few verbs), the instance of ‘aki poses an interesting challenge to analyses, as one needs to license the second absolutive and get ‘aki over to the verbal complex. However, closer examination of the data reveals that ‘aki offers an independent argument for the lexicality of case marking in Tongan

6.1 Category of ‘Aki

• First, of all, what category is ‘aki?

• I argue (deficient) verb, because:

  – ‘Aki assigns absolutive case, like verbs do
  – ‘Aki takes the verbal pattern as a complement of the conjunction ka
  – “Compacted” ‘Aki-phrases appear clause-finally (consistent with either preposition or verb)
  – ‘Aki can appear within the verbal complex as a complex predicate (consistent with either verb or preposition, cross-linguistically; consistent only with verb Tongan-internally)

• Has to be deficient (or otherwise exceptional kind of) verb, because:

  – ‘Aki can’t appear as a main predicate
  – ‘Aki can’t take any of the verbal derivational marking
6.2 Analyzing the Alternations

6.2.1 “Prepositional” ‘Aki

- Unique properties
  - ‘Aki assigns absolutive case
  - ‘Aki and complement are contiguous unit
  - The unit of ‘aki + complement tends to appear clause-finally

- Two possible analyses
  1. ‘AkiP as Complement: Main Verb’s ARG-ST list is extended by an ‘akiP; combines to form structures like:

   (10)
   
   \[
   \begin{array}{c}
   \text{PredP} \\
   \text{Pred} \quad \text{Basic Arguments} \quad \text{‘akiP}
   \end{array}
   \]

  2. ‘AkiP as Pure Adjunct: ‘akiP combines directly with PredPs
     ‘Aki specified as SEL <PredP>; combines to form structures like:

   (11)
   
   \[
   \begin{array}{c}
   \text{PredP} \\
   \text{PredP} \quad \text{‘akiP}
   \end{array}
   \]

   Pred Arguments

- Unclear at present which is better (but there are some things to follow that might distinguish the two)

6.2.2 Applicative ‘Aki

Unique Properties:

- ‘aki in Verbal complex

- But not an affix → can scramble with other verbal complex particles, preferred order with kotoa is after kotoa

- Applied object appears later in the clause
• Applied object coded like and behaves like an object:
  – Marked by absolutive case
  – Can scramble like an absolutive
  – Can “raise” like an absolutive

Analysis:
• Complex Predicate: ‘Aki as a (deficient) verb can form a complex predicate with the main verb
• With that, ‘aki and its argument are promoted to the main verb (argument attraction)
  verb < x y > + ‘aki < z > → verb’ < x y z ‘aki >
  (Underspecified analysis, where the attraction largely just involves semantic arguments is maybe not necessary for transitives, but is necessary for scaled-up view of this process)
• Case linking constraints then need to come in then and link (via lexical semantics of complex predicate):
  x → NP[erg]
  y & z → NP[abs]
  Yielding
  \[
  \text{ARG-ST} \left( \text{NP[erg]}_x, \text{NP[abs]}_y, \text{NP[abs]}_z, \text{Pred}^0[\text{‘aki}] \right)
  \]
• This newly formed ARG-ST/VAL then enters the syntax in one of two ways:
  1. Entirely by the head-valents-cxt – all arguments enter the syntax at the same time. ‘Aki is the positionally restricted by LP constraints
  2. By two constructions: head-word-cxt for ‘aki itself and head-valents-cxt for all other arguments, no specific LP constraints required

6.2.3 The Doubling Construction
• A third construction involving ‘aki: The doubling construction
• It features both a verbal complex ‘aki and a “prepositional” ‘aki. Schematically:
  V ‘aki NP NP ‘akiP

Two ways to analyze the doubling construction:
1. Parasitic on applicative ‘aki

   One possible analysis: Using underspecification to our advantage, the doubling construction is viewed as a variant of applicative ‘aki.

   A verb with the arguments \(< x y z \) ‘aki> is created, like with applicative ‘aki.

   But there is different linking (assuming z was the original complement of ‘aki):

   \[ x \rightarrow \text{NP}[\text{erg}] \]
   \[ y \rightarrow \text{NP}[\text{abs}] \]
   \[ z \rightarrow \text{PredP}[\text{‘aki}] \]

   Yielding:

   \[ \text{ARG-ST} \langle \text{NP}[\text{erg}]_x, \text{NP}[\text{abs}]_y, \text{PredP}[\text{‘aki}]_z, \text{Pred}^0[\text{‘aki}] \rangle \]

   Question, then: how is this exceptional linking licensed?

2. Parasitic on the prepositional variant

   - On this approach, the doubling construction is just the the ‘akiP (however it is to be analyzed) plus an additional “free” ‘aki, perhaps like the following ARG-ST:

     \[ \text{ARG-ST} \langle \text{Pred}^0[\text{PFORM ‘aki}], \text{PredP}[\text{PFORM ‘aki}] \rangle \]

   - Remaining Q: What happens with the argument of ‘aki in the first instantiation of it in (12)?

The test: what is the behavior of base intransitives with the doubling construction?

   - If they transitiivize, that would be support for “parasitic on argument attraction” approach

   - If they stay intransitive, then that would be support for the “parasitic on the prepositional” approach

   - If they don’t allow doubling, then we’re back to square one

Base intransitives appear to transitive, supporting the “parasitic on applicative ‘aki” approach

6.3 The Behavior of Applicative ‘Aki with Intransitives

   - With ‘aki, intransitives are involved with a transitivity alternation

     \[ \text{V NP[abs]}_y \text{PredP}[\text{‘aki}]_z \leftrightarrow \text{V ‘aki NP[erg]} \text{NP[abs]} \]
• This phenomenon supports a lexical approach to case assignment with ‘aki.

Why?

– Because if we assume that the applicative is formed in the syntax (whether it’s argument attraction in the syntax or head-movement in the syntax), the base dependencies of will be unchanged. Therefore, we’d predict two absolutes when applicative ‘aki appears with an intransitive

– We, in fact, need the lexical semantics-syntax interface to see a verb with the arguments \( <x \ y \ ‘aki \ > \), with the linking

\[
\begin{align*}
x & \rightarrow \ NP[\text{erg}] \\
y & \rightarrow \ NP[\text{abs}]
\end{align*}
\]

The only way I see to get that is if the argument attraction occurs in the lexicon and crucially is not syntactic

– The rest of the syntax should follow from the existing analysis

6.4 Long-Distance Dependencies and ‘Aki

• Unique Properties

  – ‘Aki requires a gap realization of the foot of a LDD

  – Verbal Complex ‘aki is needed (just in case the location of ‘aki wasn’t convincing enough, intransitives also transitivize in LDDs)

  – Solutions

  **Gap realization problem:** This should fall out if we have done the relationship between case and LDD-foot realization right in Chapter 4

  **VC ‘aki problem:** I see this as contextual neutralization problem (the alternation between “prepositional” ‘aki and applicative ‘aki is neutralized in favor of applicative ‘aki in the context of an LDD). The key questions are how and why.

  2 options, depending on the analysis of “prepositional” ‘aki:

  1. ‘akiP is purely an adjunct (not a complement) → can’t “extract” from a true adjunct (in HPSG terms, there’s no way to, as ARG-ST list has no access to non-complement adjuncts)

  2. ‘akiP is a complement, but is an island (islandhood status should be derived from an independent property, perhaps its property of being a deficient/quasi-participial verb)

6.5 Interaction of NI and ‘Aki

• It has been suggested by Massam (2001) that the ability of instruments to be incorporated has something to do with the ability of instruments to appear in the absolute
• If this is so, the relation probably is not direct in the syntax. To derive instrumental-like incorporanda from the syntax would involve a deletion of ‘aki – a huge problem for theories that assume that all transformations are structure-preserving.

• The manner/means incorporation facts seem to fall out quite nicely from semantic account given in 5, so it seems unlikely that we need to relate ‘aki with these kinds of NI, except (possibly) by giving them somewhat similar semantics.

7 Quantifier Float with Kotoa

Phenomenon: The quantifier kotoa, ‘all’ can both appear as an NP modifier and as a member of the verbal complex.

Issues: Is quantifier float an object property as Chung claims? Does kotoa floating offer independent evidence for ARG-ST?

7.1 Empirical Facts

• The alternation schematically: V ... [NP(abs) kotoa] ↔ V kotoa_i ... NP[abs], as in (13)

(13) a. Modificational Kotoa
   Na’e kai ‘e he kau tangata ‘a e mata’i ika kotoa.
   PST eat ERG DET PL man ABS DET pieces fish all.
   ‘The guys ate all the fish.’

   b. Floated Kotoa
   Na’e kai kotoa ‘e he kau tangata ‘a e mata’i ika.
   PST eat all ERG DET PL man ABS DET pieces fish.
   ‘The guys ate all the fish.’

• The “launching site” is restricted to absolutives. However, this includes absolutives of either the O or S variety.

∴ Not an object property, but a case property.

7.2 Problem for Configurational Approach to Quantifier Float

• “Traditional” approach: adverbial use of quantifier is derived from adjectival use of quantifier through stranding. The restrictor of the quantifier moves to another site in the clause, leaving the quantifier behind.

• This faces numerous problems if applied to Tongan kotoa.
– The ideal spot for the starting point of the quantifier in a stranding analysis is spec, vP: for a language like English, this appears to neatly account for both the position of the floated quantifier and the subject-orientation of it. However, the NP containing kotoa is not necessarily based-generated there because it’s the absolutive that launches quantifiers.

– Supposing, though, that the NP with the kotoa moved to a spec, vP, one could possibly get the right order for kotoa, BUT the absolutive NP would have to move again to get its correct positioning. There’s no reason to (it’s already got case, it doesn’t need to check [+WH] feature or anything like that)

– A different route, where kotoa does the head movement, faces the issues of how an adjunct can head-move and how the ordering with other verbal complex elements might be achieved

7.3 Towards A Solution

• Kotoa, like ‘aki is a deficient verb. As such it has two realization options:
  1. Modificational
  2. “Adverbial”

• A bit more about each realization:

Modificational

• Here kotoa just is a regular modifier

• It’s specified SEL < NP[case] > – it’s specified to be outside the regular nominal expression because there’s evidence that it’s that far out (e.g. it appears after the definitive accent) Q: Why should that be?

• It combines in the usual manner for nominal modifiers; there’s no restrictions on cases; thus, kotoa appears with nouns of any case

“Adverbial”

• Analysis: The intuition is that kotoa is another instance of a complex predicate in Tongan, in particular the kind called in the serial verb literature an “ambient serialization” (Crowley 1987) or event-argument serial verb constructions (Aikhenvald 2006). Here, there doesn’t seem to be a clear morphosyntactic sharing of arguments between the two predicates involved in the serial verb construction and at most there’s just a semantic link between a main predicate argument and the “adverbial” verb.
The analytical question: How to model this?

Some ideas:

1. Complex Predicate that Just Raises
   - A possible way of dealing with ambient serialization is to treat it as an instance of somewhat ordinary raising. The argument of the deficient verb, kotoa is identified with the absolutive of the main verb. Thus, kotoa would be extended as in (14)
   
   $$(14) \begin{align*}
   \text{ARG-ST} & \left[ \left[ \text{NP[abs]} \right] \oplus \left[ \text{L-ID all-rel} \right] \right] \end{align*}$$
   
   - The absolutive case restriction would be a consequence of the raising. Kotoa as a single place predicate would have to be assigned absolutive, and because of the identity relation in (14), the realized argument would have to be
   - Q’s: Why should some complex predicates just do normal raising, while others do argument attraction “raising”? Why analyze this as raising when there’s no athematic arguments here?

2. Depictive Secondary Predicate with Complex Predicate Syntax
   - Somewhat like ‘aki, kotoa is extended to verbs’ ARG-ST lists like this:
   
   $$(15) \begin{align*}
   \text{ARG-ST} & \left[ \left[ \text{NP[abs]} \right] \oplus \left[ \text{L-ID all-rel} \right] \right] \end{align*}$$
   
   This extension requires co-indexing with an absolutive NP, following the above generalization that this is a case-specific problem
   - The syntax, then, is quite like ‘aki → it could combine using a head-word-cxt or combine via the head-valents-cxt and be subject to particular LP constraints
   - Some outstanding questions: Why is kotoa a depictive? How does this related to other floated quantifiers in the Polynesian family (some of which are cognate, some of which are not)?
   - The depictive part of the analysis might be key to answering (some of) these questions.
   * Depictives, cross-linguistically, are more often S and/or O-oriented, not A-oriented; thus, opening the possibility for an explanation for their absolutive orientation in Tongan
   * Furthermore, depictives are known to be resistant to modifying obliques. On the idea that e-marked NPs (the ergative in ergative Polynesian languages) and i-marked NPs (the accusative in accusative Polynesian languages) are “too oblique” for kotoa’s purposes (what does this actually mean?), this could explain the variation among the quantifier floating possibilities at least between Tongan kotoa and Mäori’s cognate kātoa
However, it is an open question how this analysis might be related “ambient serialization” to other instances of complex predicate formation.

8 Conclusions

8.1 Synopsis of Major Findings of Specific Chapters

- In the lower levels of the clause, the structure is much flatter than previous analyses have assumed. There’s no syntactic evidence for a VP; thus, Tongan patterns with many other verb-initial and free word order languages in lacking one
- In the higher levels of the clause, some hierarchy is needed between the TAM and the rest of the clause
- The Argument Realization Case Studies

Clitics: Clitic Movement proves to be untenable. A better analysis is through copy raising via the XARG feature

Canonical Case Marking: Contingent on both (the semantics of) verb classes and verbal morphology. Case assignment is best captured on the ARG-ST list of words (a late, ‘deep’ property, if that isn’t a contradiction in terms). LDD-foot realization is dependent on this structure because this is an even more surface-y property

Morphosyntax of Noun Incorporation: Both VP-remnant movement and Head-movement are untenable. Tongan NI is better analyzed as a compounding structure with somewhat exceptional modification possibilities. The incorporated noun has no syntactic life, both within the phrasal structure and the argument structure. The distribution of incorporanda with main predicates is mediated totally by the semantics.

The Lexical Semantics of Noun Incorporation: Existing syntactic/lexical semantic categories are insufficient in accounting for predicates that license incorporanda. A semantic solution based on event autonomy offers more descriptively adequate account.

The Instrumental Particle: Here there is significant interaction on the ARG-ST of the main verb, especially in the case of applicative ‘aki which involves complex predicate formation and concomitant argument attraction. Case must be licensed lexically and not syntactically within this construction.

Quantifier Float with Kotoa Structural account of quantifier float extended to Tongan kotoa run into a myriad of problems. This, too, is better involved as an ARG-ST interaction, but one that crucially relies on case, rather than grammatical relations (or other category)
8.2 Conclusions to be Drawn from this Larger Study

- Chung’s original argument that the argument realization patterns are not just dependent on case was on the right track

- But a closer examination of several argument realization phenomenon indicate that there is more complexity than was acknowledged in 1978

- Phenomenon are variously related to the semantics, the syntax-semantics interface, and more surface phenomenon

- However, there seems to be little reason to account for these interactions with a more articulated tree-structure; instead the enrichment in syntactic theory seems to be more profitably done with enriched dependency-related data structures producing minimal syntactic structures

References


