

# Complex Predicates and Nuclear Serial Verbs\*

Avery D Andrews  
The Australian National University  
email: Avery.Andrews@anu.edu.au

**Proceedings of the LFG97 Conference**  
University of California, San Diego

Miriam Butt and Tracy Holloway King (Editors)

1997  
CSLI Publications  
<http://www-csli.stanford.edu/publications>

---

\*I would like to especially to thank Aleksandra Aikhenvald for cheerfully answering an endless stream of queries about Tariana, Joan Bresnan, Christopher Manning, and the attendees of the Sociable Syntax Supper Group at Stanford University for useful discussion of the analysis, and Xerox PARC and the Stanford Linguistics Department for providing facilities and support for the preparation of this paper.

In this paper I will discuss some serial verb constructions in Tariana, a North Arawakan language of Brazil, currently being described by Alexandra Aikhenvald (1994, in prep.). The SVC's I will be looking at are combinations of V (rather than VP or S), and therefore fall into Foley and Olsen's 1985 category of 'nuclear' serializations. They express a wide variety of different kinds of semantic composition; here I will analyse ones that have a similar function to the complex predicates discussed in Romance by Alsina (1996, 1997) and Urdu by (Butt (1995), Butt (1997)), although I will briefly illustrate some of the others.

One result of the discussion is some additional evidence for the existence of a complex predicate formation mechanism that is syntactically different from complementation in spite of having a similar semantic effect. Another is evidence that it is wrong to propose a strong distinction between serial verbs and complex predicates (as suggested for example by (Butt 1995:223-226)). Rather I will propose that 'serial verb' is a formally rather ill-defined concept comprising a variety of different c-structure configurations, each themselves able to encode a substantial variety of more abstract grammatical and semantic structures. Some of these will be the same as those conveyed by currently recognized types of complex predicates, while others will be quite different. A final result will be a reformulation of complex predicate formation for Romance within a modified LFG architecture that avoids some problems with Alsina's version.

## 1 Tariana and its SVC's

Tariana<sup>1</sup> has variable word-order, with some tendency for the subject to be initial, although the order of verbs and other complements is variable. The grammar is nominative-accusative in its overall organization. Dynamic verbs usually show person-number agreement with their subjects. Stative verbs on the other hand do not agree, and neither do Portuguese loan-verbs (Aikhenvald 1994:202).

There are also case-marking phenomena, in the form of a subject-nonsubject distinction for animate pronouns, and a 'topic-marking' affix *-nuku* that can be attached only to non-subjects. Clauses also contain a Tense-Aspect-Mood (TAM) particle, whose position is quite variable.

Tariana SVCs are sequences of verbs which normally each carry a PN-marker, but can have only one TAM marker and also only one negation. The verbs can be separated by the TAM marker and also by certain Aktionsart particles, but not by NP arguments or adjuncts. Here are some examples:

- (1) a. i-na-mha            yeme    hi-nuku  
           2pl-want-PROG 2pl.sniff DEM:AN-TOP

---

<sup>1</sup>Spoken by about 100 people in the Vaupes region of the Brazilian Amazon.

You want to sniff this (snuff).

- b. tarada-peni wa wama-daka  
 alive-PL:AN 1pl.go 1pl.look for-YET  
 We shall look for the live ones (fish) yet.

In (a) we see the object after the verb, bearing the *-nuku* marker, with the two components of the (serial) verb each bearing PN-marking, but separated by a Progressive TAM marker. (b) is similar, except that the object appears in front of the serial verb.

These SVCs can therefore plausibly be described as V-over-multiple-V constructions, obeying the structural constraints described by (a, b) below, which will allow an SVC component to carry something that can plausibly be adjoined to V to produce a V, but won't allow them to be separated by an NP or PP.

- (2) a. V → V\*  
 b. V → V V

The (a) formulation allows for any number of verbs to appear in the SVC, while (b) is binary. Both unlimited branching and binary branching SVCs seem to exist, and recursion of the binary-branching structures is also found.

## 2 'Pseudo-Complement' Serialization

Tariana SVCs fall into a considerable variety of types, of which I will be primarily concerned with ones that I call 'Pseudo-Complements'. These SVCs have at least approximately the same semantics as complement or modal auxiliary constructions, and are in fact exemplified by (1) above, where in terms of the semantics the first verb in the SVC seems to be taking the second as an argument.

Modal SVCs are among what (Aikhenvald 1995) classes as 'asymmetric' SVCs, wherein a closed-class verb combines with an open-class one. The verbs appearing in modal SVCs are: *-na* 'want', *-uma* 'strive for, look for', *-a* 'be going to (FUT)', and *ira* 'need'.<sup>2</sup>

One possible analysis for these constructions would be as XCOMPs. The second verb would be annotated as an XCOMP of the whole SVC, and the nonsubject NPs would be annotated as XCOMP\* OBJ (as has sometimes been proposed for objects in German). But there are two arguments against this.

The first is provided by a very similar looking SVC construction that forms causatives. In this construction the Causative verb comes first, and

<sup>2</sup>The last differs from the others in not taking a PN-marker, as indicated by the absence of the initial dash, and is furthermore involved in an interesting agreement phenomenon, as will be discussed in (Andrews and Manning in prep).

is one of *a* ‘go, give, get to, say’ (for plain causatives), *wana* ‘call’, *matara* ‘give permission to, leave’ *ira* ‘order’:

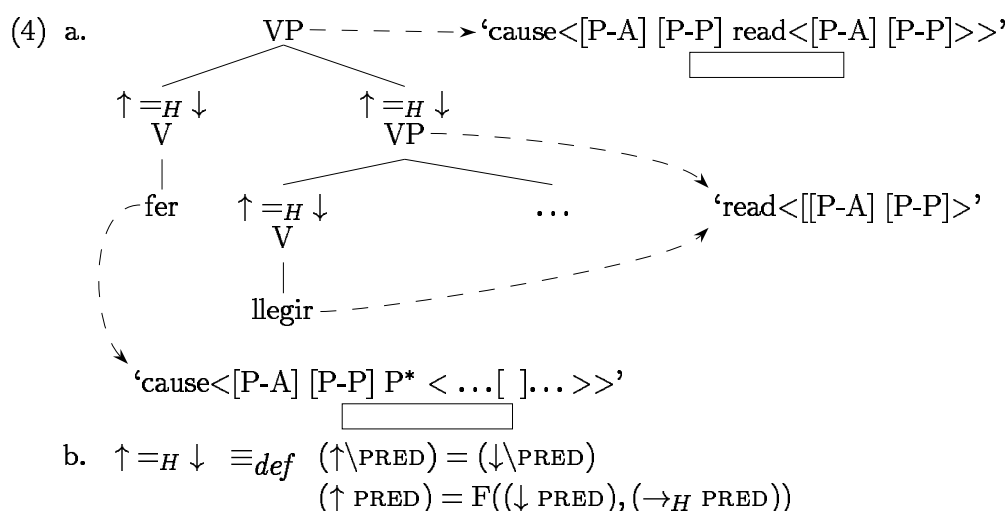
- (3) a. *ka:ru-ka nuha nu-a-mahka nu-hyã-niki*  
 fear-DECL I 1sg-give-NONVIS 1sg-eat-COMPL  
*piri-nuku di-a-pidana*  
 2sg.son-TOP 3sgnf-say-PART  
 Being afraid, I let (the fish) eat your son, he said
- b. ... *na-wana na-whe-ta-pidana*  
 3pl-call 3pl-stay.CAUSE-CAUSE-PART  
 ..., they called him to stop his canoe

In (a) we start out with a circumstantial adjunct *ka:ruka* ‘being afraid’, then we have a nominative subject pronoun (the nonsubject form would be *nu-na*), and finally our causative SVC. The first component V is unremarkable (the gloss NONVIS indicates ‘Non-Visual Past’ Tense, meaning that the event was not seen happening), but the second is rather surprising, because it shows subject agreement with the putative matrix subject, which is not even one of its own participants. To deliver an appropriate meaning, the XCOMP analysis would require that the object of the Causative verb be set as the SUBJ of the Effect verb in the XCOMP by functional control (see (Matsumoto 1996) for recent applications of this idea to causative and desiderative constructions). But this produces the expectation that the Effect verb will agree with the Causee (its own Agent), if it agrees with anything. The same surprising agreement pattern is seen in the second example, where in addition the semantic complement verb is itself a causative, in fact morphologically a double causative. Syntactically however this V is only singly causativized, the morphological doubling serving to intensify the effect on the Theme (Aikhenvald 1996). So he is supposed to stop (make-stay) the canoe, but the make-stay verb shows PN agreement with the subject of ‘call’.

‘Classic’ LFG (the theory of (Bresnan 1982b)) doesn’t provide any plausible analysis for this kind of construction. But more recent work on complex predicates by Alsina, Butt and others suggests a possible answer. The idea of a complex predicate is that two distinct argument-taking lexical items combine in such a way as to take their arguments as a single array of grammatical relations. On such an analysis, both the Cause and Effect verbs will have the same subjects, objects, etc., so the agreement of the Effect verb with the Cause verbs subject is not problematic.

### 3 Formulating Predicate Composition

Alsina’s conception of complex predicate formation is illustrated in the following diagram for a complex predicate in Catalan (Alsina 1997:236-237):

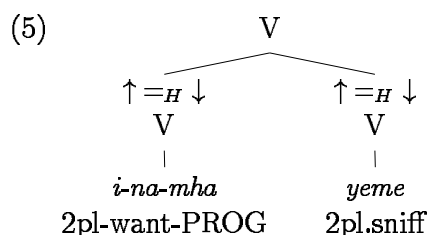


Here the causative verb *fer* has an ‘incomplete’ PRED-feature that provides semantic roles for Causer and Causee arguments, and also contains a position for an Effect predicate, one of whose arguments is to be identified with the Cause. This PRED-value is to be combined with that of the Effect verb *llegir* to produce the PRED-value the matrix VP. The linking theory then associates the argument-positions in the resulting complex-predicate with a set of grammatical relations that are shared across all the VP-levels (the ‘[P-A]’ and ‘[P-P]’ notations in the argument-positions represent ‘Proto-Agent’ and ‘Proto-Patient’ specifications that are related to but different from those of (Dowty 1991)).

The sharing is accomplished with the ‘ $\uparrow =_H \downarrow$ ’ annotation, which uses the concept of ‘restriction’ of (Kaplan and Wedekind 1993) to equate all the attributes of the mother and daughter nodes except for PRED.<sup>3</sup> The PRED attributes of the conceptually ‘main’ and ‘complement’ verbs are then combined by means of the somewhat complex function F, which is supposed to be defined in such a way as to combine the PRED-values of the two daughter nodes so as to produce an appropriate one for the mother. I won’t go into the details here.

Adapting these general ideas to the V-under-V c-structure that seems motivated for Tariana, we get the following annotated c-structure for a modal serialization, and similarly for causatives:

<sup>3</sup>If *f* is an f-structure, *f*\GF is defined as *f*, less any attribute-value pair with first member GF, if one occurs in *f*. In the apparently simplified presentation of (Alsina 1996:201-202), restriction is not used, rendering the account inconsistent, since identical f-structures can’t have different PRED-values.



I find two problems with this analysis. First, the phrase-structure is symmetric, both verbs being annotated with ‘ $\uparrow=H \downarrow$ ’. This provides no way of expressing the constraint that the modal or causative verb should come first.<sup>4</sup> A deeper problem is that the function  $F$  is an inherently undesirable addition to LFG. LFG has traditionally used an extremely restricted range of devices for producing the  $f$ -structures of mother nodes from those of their daughters: the value of a daughter-path can either be equated to or set as a member of the value of a mother-path.

This is a significant difference between LFG and HPSG, where the framework allows the use of arbitrary functions to assist in building the feature-structures of constituents (of course, any particular version of HPSG will only allow a specific assortment of such functions, but the general framework imposes no limits on what they might be). While it may turn out that LFG has been too restrictive in this respect, I think it’s still worthwhile to try to see what can be done using the original spartan inventory of techniques.<sup>5</sup>

To formulate predicate-composition in an explicit way, it is helpful to dissect the classic PRED-feature into at least two distinct attributes, including one which I’ll call LCS (lexical-conceptual structure), giving the meaning of the lexical item, and another I’ll call TERMS, which is an ‘argument-structure’ in the sense of (Grimshaw 1990), (Rosen 1989) and much subsequent work. The TERMS attribute is similar to the  $a$ -structures of (Manning 1996a), or the SUBCAT (or, more recently, ARG-ST) lists of HPSG, except that it includes only the ‘direct’ or ‘core’ arguments, and not obliques, which I assume to be handled by some different mechanism that I will not investigate here.<sup>6</sup> It is organized as a Prolog- or LISP-style list-structure, which is a right-branching binary tree whose the left (HEAD) member is the first list element and whose possibly empty/nonexistent left (REST) member is the remainder of the list. This kind of organization, which has been proposed by many authors, will be important for defining

<sup>4</sup>This is not a problem for Alsina’s analysis of Romance languages, because the annotation appears on two different categories,  $V$  and  $VP$ , which are therefore orderable by the  $c$ -structure rules.

<sup>5</sup>I’d also like to conjecture that there is a processing motivation for LFG’s restrictiveness in the regard: an LFG processor won’t have wait for the output of arbitrary functions, which might not be able to return a value until all of the daughters of a node have been processed. Instead, the  $f$ -structure of each daughter can be combined with that of the mother incrementally.

<sup>6</sup>See (Jackendoff 1990) and (Wechsler 1995) for some relevant proposals.

argument-structure composition. These two attributes are not intended to be exhaustive; I believe that there has to be at least one more, which we might call ‘LEXFORM’, to provide the interface to morphological spellout (allowing for the various distinct senses of the word GO for example to all have the same irregular past form *went*). But since this attribute is irrelevant to the subject of this paper, it will be omitted from the diagrams.

Here is a presentations of the feature-structure that we might have for *kill*, utilizing an old but obsolete proposal for its lexical decomposition, although we will elsewhere simply use the lexical item in italics as its own semantic representation:

- (6) a. PRED ‘Kill< X, Y >’  
 b. 
$$\left[ \begin{array}{ll} \text{LCS} & [\text{Cause}(X, \text{Become}(\text{Not}(\text{Alive}(Y))))] \\ \text{TERMS} & < X, Y > \end{array} \right]$$

The use of capital italics to connect the positions in the LCS to the positions in the TERMS list is intended to leave room for some latitude as to how this linking is to be accomplished. One possibility is conventional unification, but I would like to leave open other possibilities, such as the use of ‘glue-language’ semantics. On this kind of account, the LCS attribute would serve as a location in which to place semantic information produced by meaning-constructors such as:

- (7)  $(\uparrow \text{TERMS FIRST LCS}) \rightsquigarrow X \otimes (\uparrow \text{TERMS REST FIRST LCS}) \rightsquigarrow Y \multimap$   
 $(\uparrow \text{LCS}) \rightsquigarrow \textit{Kill}(X, Y)$

On this account, the italic variables represent not unification in the syntax, but rather co-instantiation of variables between premises and conclusion in a meaning-constructor.

There will also have to be a linking theory, to associate positions on the argument-list with grammatical functions. For English, the effect we want from the linking theory is that the top position on the TERMS-list gets equated to the value of the SUBJ grammatical function, while the remaining positions are equated with OBJ-values. There are various ways in which this could be achieved, and a number of empirical issues to be resolved, such as whether to remain with the standard LFG distinction between ‘restricted’ and ‘unrestricted’ object functions developed in the Lexical Mapping theory of (Bresnan and Kanerva (1989), Alsina and Mchombo (1993)) and others, or whether to adopt the proposal of (Alsina 1996) of a multiple-valued OBJ function. For present purposes I will assume without argument or discussion that the top TERM is linked to SUBJ, the second to OBJ, and the last to OBJ2, although nothing hinges on this particular treatment of objects.

With this background in place we can formulate a proposal. The PS rule for these constructions needs to introduce the second V as a some

kind of argument of the first V. None of the conventional GF's displays the kind of behavior we want, so I will use a new one, 'ARG' (for 'argument'), yielding:

$$(8) \quad V \quad \rightarrow \quad V \quad \begin{array}{c} V \\ \text{ARG} \\ =\text{SUBJ} \end{array}$$

Lexical entries for the causative and modal serial verbs can now be formulated to combine the LCS and TERMS-values so as to produce semantically appropriate results, in a manner generally reminiscent of HPSG analyses such as (Manning et al. to appear).

Below appear feature-structures for the 'light' verb 'want' and its 'semantic complement' 'sniff':

$$(9) \quad \text{a.} \quad \left[ \begin{array}{ll} \text{LCS} & [\textit{Want}(X, Y)] \\ \text{TERMS} & \langle X \mid \dots \rangle \\ \text{ARG} & \left[ \begin{array}{ll} \text{LCS} & Y \\ \text{TERMS} & \langle X \mid \dots \rangle \end{array} \right] \end{array} \right]$$

(ellipsis dots designating equated substructures)

$$\text{b.} \quad \left[ \begin{array}{ll} \text{LCS} & [\textit{Sniff}(W, Z)] \\ \text{TERMS} & \langle W, Z \rangle \end{array} \right]$$

The 'want' entry looks a bit complex, but note that it can be reduced to small number of equations, similar in complexity to a conventional LFG lexical entry:

$$(10) \quad \begin{array}{l} (\uparrow \text{LCS}) = \textit{Want}(X, Y) \\ (\uparrow \text{TERMS HEAD}) = X \\ (\uparrow \text{ARG LCS}) = Y \\ (\uparrow \text{TERMS}) = (\uparrow \text{ARG TERMS}) \end{array}$$

The equation for handling the TERMS lists is particularly straightforward.

Now when the 'sniff' structure is unified with the ARG-value of the 'want' structure, the result will be:<sup>7</sup>

---

<sup>7</sup>If one uses glue-language semantics, the following meaning-constructor for 'want' will produce the right results:

$$(i) \quad \begin{array}{l} (\uparrow \text{TERMS FIRST LCS}) \rightsquigarrow X \otimes (\uparrow \text{ARG LCS}) \rightsquigarrow Y \multimap \\ (\uparrow \text{LCS}) \rightsquigarrow \textit{Want}(X, Y) \otimes (\uparrow \text{TERMS FIRST LCS}) \rightsquigarrow X \end{array}$$



$$(11) \left[ \begin{array}{l} \text{LCS} \quad [ \text{Want}(X, \underbrace{\text{Sniff}(X, Z)}_Y) ] \\ \text{TERMS} \quad \langle X, Z \rangle \\ \text{ARG} \quad \left[ \begin{array}{l} \text{LCS} \quad Y \\ \text{TERMS} \quad \langle X, Z \rangle \end{array} \right] \end{array} \right]$$

Finally, the linking theory will apply, adding SUBJ and OBJ GF's, and, due to the '=SUBJ' annotation on the second V, the SUBJ-value will be shared between the upper level and the ARG-VALUE:

$$(12) \left[ \begin{array}{l} \text{LCS} \quad [ \text{Want}(X, \underbrace{\text{Sniff}(X, Z)}_Y) ] \\ \text{TERMS} \quad \langle X, Z \rangle \\ \text{SUBJ} \quad [ ] \\ \text{OBJ} \quad [ ] \\ \text{ARG} \quad \left[ \begin{array}{l} \text{SUBJ} \quad [ ] \\ \text{LCS} \quad Y \\ \text{TERMS} \quad \langle X, Z \rangle \end{array} \right] \end{array} \right]$$

The SUBJ-sharing implements the PN-marking of the second verb.

The causative verbs will be similar, but require a slightly more complex entry, so that the causative can introduce an additional argument of its own, not shared with the semantic complement. A reasonable straightforward proposal is (a) below, where (b) is an equation concisely expressing the desired mode of combination of the TERM-lists:

$$(13) \text{ a. } \left[ \begin{array}{l} \text{LCS} \quad [ \text{Cause}(X, Y, Z) ] \\ \text{TERMS} \quad \langle X, Y \mid \dots \rangle \\ \text{ARG} \quad \left[ \begin{array}{l} \text{LCS} \quad Z \\ \text{TERMS} \quad \langle Y \mid \dots \rangle \end{array} \right] \end{array} \right]$$

$$\text{b. } (\uparrow \text{TERMS REST}) = (\uparrow \text{ARG TERMS})$$

This entry says that Tariana causatives have 'full merger' of argument-structures, in the sense of (Rosen 1989), while the modal serializations simply identify the modal and semantic complement argument-structures. But this general method of combination is capable of producing other effects as well, such as Rosen's 'partial merger'. Further work on Tariana will be required in order to determine what mode of argument-structure composition is actually required.

The '=SUBJ' annotation discussed above here produces the spectacular effect of causing the Effect verb to agree with the Cause verb's subject. It is the use of predicate-composition rather than complementation to describe this construction that lets us describe this effect without contradiction, and perhaps even in some measure explain it: the SUBJ-sharing

annotation can be seen as part of the grammar of this construction that was presumably originally motivated by the overt form of the diachronic source of these construction, and produced this peculiar agreement effect when the construction extended to comprise causatives.

This analysis entail that when a two-place verb is serially causativized, a 3-place predicate results. It is therefore expected that the language should have basic three-place predicates whose arguments should be expressed in the same form as those of causatives, and this appears to be true: verbs like ‘give’ take their Recipient and Theme as bare accusative NPs, and this is also what happens with the Causee and the Effect Theme of a causativized transitive (Aikhenvald pc.)

- (14) a. kana nu-a-mahka          na-na  
           corn 1sg-give-NONVIS they-ACC  
           I gave them corn
- b. [hekuda-nuku    pi-na-ka          pi-a    pi-hỹa nun-na]  
           [fruit-TOP.ACC 2sg-want-SEQ 2sg-let 2sg-eat 1sg-OBJ]  
           pi-sepata pini-tha-sika                                  phia  
           2sg-suffer 2sg-do-FRUST-INFERRED you  
           Wishing to let me eat the fruit, you suffered

The second example, from a traditional narrative, also illustrates the embedding of a causative under a modal serialization. The marker *ka*, glossed SEQ, indicates that its clause (enclosed in square brackets) is subordinate. Our analysis predicts this straightforwardly, with the second daughter V in (8) clearly able to itself expand in accordance with (8). Modal inside causative serialization on the other hand does not seem to be possible, perhaps for semantic reasons (Aikhenvald p.c.).

## 4 Some Other Types

Modal and causative serialization are only two of the many types of Tariana SVC discussed by (Aikhenvald 1995). Here I will briefly review some of the others, without much formal analysis, in order to give an impression of the range of these constructions. Modal and causatives are amongst Aikhenvald’s ‘asymmetric’ serializations; in these an open-class item combines with a closed-class one. Some other asymmetric types include aspectual and directional serialization:

- (15) a. Aspectual (with causative):  
           na-na    dura          du-pita    du-yã-nhi  
           3pl-OBJ 3sgf.order 3sgf-bathe 3sgf-stay-ANT  
           She used to order them to bathe. or  
           She ordered them to usually bathe.

## b. Directional:

...nha ke:ča na-doReta nhe na-pidana  
 ...they black ants 3pl-carry 3pl.enter 3pl.go-PART

na-ya-yawa-se

3pl-POSS-CL:HOLE-LOC

[after the young man gave them some manioc flour], the black ants carried it away into their hole.

In both of these types the aspectual or directional verb comes second; in the aspectual example (a), the aspectual verb can be construed with either the causative verb or its pseudo-complement, as would be predicted by the structure given by (8). The directional example (b) illustrates a characteristic pattern whereby the first motion verb describes the motion relevant to the intrinsic features of the setting and the actions being described, while the second describes it relative to the current narrative point of view.

Some kind of complementation-like semantics might be appropriate for these serializations, but not at all for the next, which Aikhenvald calls ‘symmetric’. Symmetric serializations consist of two or more open-class verbs which must agree in transitivity. They express actions which either occur in sequence (for discrete events), or randomly interleaved (for activities):

(16) a. ma wa-wa wa-dana wa-yarupe-nuku  
 let’s 1pl-read/play 1pl-write 1pl-thing-TOP  
 Let’s read and write up our language!

b. phia-nihka phita pi-thake-ta pi-eme  
 you-DOUBT 2sg.take 2sg-cross.CAUS-CAUS 2sg-stand.CAUS  
 ha-ne-na hyapa-na-nuku?  
 DEM-DIST-CL:VERT hill-CL:VERT-TOP  
 Was it you who brought that mountain across?  
 (lit. take - cross - put upright)

c. na-ira-sita-pidana na-inu-kaka-pidana  
 3pl-drink-AFTER:SS-PART 3pl-‘kill’-REC-PART  
 na-ita neyuu  
 3pl-shoot arrow 3pl.go upstream  
 After they had drunk, they ‘killed’ each other with arrows and went upstream.

A possible formal analysis for these would be to combine the LCS-values of the daughter verbs into a set-valued LCS, but equate all the TERMS-attributes. This would provide a formally sharp analysis of the requirement for transitivity agreement:

$$\begin{array}{lcl}
 (17) \text{ V} & \rightarrow & \text{V}^* \\
 & & (\uparrow \text{TERMS}) = (\downarrow \text{TERMS}) \\
 & & (\downarrow \text{LCS}) \in (\uparrow \text{LCS}) \\
 & & (\uparrow \text{SUBJ}) = (\downarrow \text{SUBJ})
 \end{array}$$

One would expect that it would also provide a reasonable basis for the semantic interpretation, although this remains to be worked out.

The final main type recognized by Aikhenvald is ‘ambient’ serialization, wherein an open-class item seems to have some kind of adverbial or other modificatory relationship to the other verb:

- (18) a. *diha diha hipatu hiku-pana ma-ka-de-pidana*  
 he he snuff appear-ALL NEG+give-DECL-NEG-PART  
 He (the traditional God) did not give (her) the stuff in a totally visible way (lit. appear-give)
- b. *thuya ha-ehkwapeni-nuku nawiki-nuku di-kaRite*  
 all DEM-world+PL:AN-TOP people-TOP 3sgnf-tell  
*di-peya-ka-pidana*  
 3sgnf-be first-DECL-PART  
 He was the first to tell (lit. tell-be first) all the people in the world (about the discovery of fire).

In the (a) example the usual PN-marking does not appear, being missing from the first V because it is stative, and from the second because it is negative: the negative prefix *ma-* pre-empts the PN-marker. In (b) on the other hand the verbs are non-stative and positive, so the PN-markers appear as usual.

All of these constructions share with the pseudo-complement SVC’s the property of consisting of a series of V’s that are separable by TAM markers and various Aktionsart particles, but not by NPs or other full phrases. They thus illustrate the general idea, highly compatible with LFG architecture, of a single phrase-structure configuration being associated with a substantial variety of more abstract syntactic and semantic structures.

## 5 Complex Predicates

Now that we’ve adapted Alsina’s analysis to Tariana, it’s time to see if we can apply the modified version back to Romance. There are at least two motivations for trying this, the first, shared with Tariana, that the function ‘F’ in (4) makes it not quite LFG, and, second, the problem pointed out by (Frank 1996) that there’s more than just PRED that we don’t want to share between the various VP-levels of these constructions. In particular there are various ‘verbal form’ features such as those controlling the infinitive/participle distinction, and the use of verbal markers such as *de*

and  $a$ , which should remain in the region local to their V, rather than spreading throughout the whole VP-stack. And this problem gets worse if PRED is replaced by LCS, TERMS and ARG.

One possibility for dealing with the verbal form features would be to use the  $\mu$ -projection (Butt et al. 1996). The  $\mu$ -projection is a projection coming off of c-structure that is similar to f-structure, but contains only verbal form features and an attribute DEP for holding the form features of a VP that's dependent to a verb.

By using a  $\mu$ -projection we can have the upper and lower V in a complex predicate have the same f-structure, but different form-features, since for two c-structure nodes to have the same f-structure does not imply that they have the same  $\mu$ -structure. But using both projections and restriction together in this way in the same theory is not desirable, because the effects being produced by these devices are so similar: they are two different ways of allowing distinct nodes in the c-structure to share many but not all of their properties.

What I propose instead is to use a variant of the approach to LFG architecture proposed in (Andrews and Manning 1993), currently being elaborated in (Andrews and Manning in prep). We take the position that rather than putting different kinds of information in different places (which creates difficult problems of integration), we put all featural information (which is information in a format intended for easy checking of consistency) in a single feature-structure. This feature-structure is formally similar to an f-structure: the c-structure as a whole has a feature-structure correspondent, and each daughter-node in the c-structure corresponds to a substructure of the feature-structure. But this feature-structure differs substantively from an f-structure in containing all of the featural information relevant to the node, include X-bar category information, and not just the standard f-structure information.

The effects of having different projections in standard LFG are produced not by putting attributes in different places, but by dividing them into classes that are made to spread differently. This can be done by an adaptation of the Kaplan and Wedekind concept of restriction which we call a 'restriction projection'. Kaplan and Wedekind's concept of restriction is negative: one specifies the attribute which is to be excluded from the restriction. A restriction projection on the other hand is a set of attributes that is to be taken positively, in fact as a standard set-theoretical restriction (as used by Kaplan and Wedekind to define their concept): the restriction of an f-structure  $f$  by set of attributes  $\alpha$ , which we denote as  $f_\alpha$ , is formed by taking  $f$ , and throwing away those attribute-value pairs in  $f$  whose attribute is not a member of  $\alpha$ .

The main thing we want to do with restrictions is equate the restrictions of two f-structure by the same set of attributes, for which we therefore define some convenient notations:

(19) where  $\alpha, \kappa, \dots$  are sets of attributes:

$$= \{\alpha, \kappa, \dots\} \equiv_{def} \uparrow =_{\{\alpha, \kappa, \dots\}} \downarrow \equiv_{def} \uparrow_{\alpha \cup \kappa \cup \dots} = \downarrow_{\alpha \cup \kappa \cup \dots}$$

Some attribute-classes (restriction projections) which presently seem to be useful are:

- (20) a.  $\beta$ : for the BAR attribute(s) or equivalent.  
 b.  $\kappa$ : X-bar categories like N, V, etc.  
 c.  $\mu$ : Morphosyntactic features (PPART, VFORM, ...; status of nominal features unclear). Could  $\mu$  actually be the same as  $\kappa$  (Bresnan p.c.)?  
 d.  $\rho$ : Grammatical Relations (SUBJ, OBJ, ADJUNCT, ...).  
 e.  $\alpha$ : Argument-structure related attributes such as LCS, TERMS, ARG.

With these classes we can formulate a rule for Romance complex predicate VPs as follows:

$$(21) \text{ VP} \quad \rightarrow \quad \begin{array}{l} \text{V} \\ =\{\alpha, \kappa, \rho, \mu\} \end{array} \quad \begin{array}{l} \text{VP} \\ \text{ARG} \\ =\{\rho\} \end{array}$$

The annotation on the V causes all attributes except bar-features to be shared between the V and the VP, while the annotation on the VP introduces it as the ARG-value of the VP, and shares only the grammatical relations (this sharing expresses the idea that the two VP's have the same f-structure, and provides for clitic-climbing and adjunct-interpretation as discussed in (Andrews and Manning 1993)).

The multiple sharings on the V look rather stipulative, but it is important to note that with proper use of the X-bar theory, they should appear only once: (21) should be seen as one instance of a very general scheme for introducing  $X^0$ -level heads and their complements, which will also be interacting with general principles, as discussed in Bresnan (1982a, in prep).

The lexical entries for 'light verbs' in Romance languages can then be identical or similar to their counterparts in Tariana, with the different c-structure rule producing the major apparent differences between the constructions. These are of two kinds: (a) in Romance the pseudo-complement is a VP rather than just a V (as discussed at length by Alsina (1996) and Manning (1996b)) (b) in Romance the light verb governs some form-features of the pseudo-complement (via the ARG-attribute, fulfilling the role of DEP in the  $\mu$ -projection proposal), while in Tariana the SUBJ-attribute is shared. Note that in fact in Tariana all of  $\rho$  could be shared, as in Romance, but there is only direct evidence for the sharing of SUBJ. These rather different-looking constructions can thus be given analyses which straightforwardly express the differences and similarities between them.

## Bibliography

- Aikhenvald, A. Y. in prep. A grammar of Tariana.
- Aikhenvald, A. Y. 1994. Grammatical relations in Tariana. *Nordic Journal of Linguistics* 201–217.
- Aikhenvald, A. Y. 1995. Serial verb constructions and verb-compounding in Tariana. paper presented at SSILA Annual Meeting, Albuquerque NM; Australian National University ms.
- Aikhenvald, A. Y. 1996. Causative constructions and transitivity increase in Tariana. talk delivered at the 1996 meeting of the Australian Linguistic Society; Australian National University ms.
- Alsina, A. 1996. *The Role of Argument Structure in Grammar*. Stanford CA: CSLI Publications.
- Alsina, A. 1997. A theory of complex predicates: Evidence from causatives in bantu and romance. In A. Alsina, J. Bresnan, and P. Sells (Eds.), 203–246.
- Alsina, A., J. Bresnan, and P. Sells (Eds.). 1997. *Complex Predicates*. Stanford CA: CSLI Publications.
- Alsina, A., and S. Mchombo. 1993. Object asymmetries and the chichewa applicative construction. In S. Mchombo (Ed.), *Theoretical Aspects of Bantu Syntax*. Stanford CA: CSLI Publications.
- Andrews, A. D., and C. D. Manning. 1993. Information-spreading and levels of representation in lfg. Technical Report CSLI-93-176, Stanford University, Stanford CA. Available at: <http://www.sultry.arts.edu.au/cmanning/papers>.
- Andrews, A. D., and C. D. Manning. in prep. *Projections and Information Flow in LFG*. Stanford CA: CLSI Publications. (provisional title).
- Bresnan, J. W. 1982a. Control and complementation. *Linguistic Inquiry* 343–434. Also in Bresnan (1982), pp. 282-390.
- Bresnan, J. W. (Ed.). 1982b. *The Mental Representation of Grammatical Relations*. Cambridge MA: MIT Press.
- Bresnan, J. W. in prep. *Lexical-Functional Syntax*. Stanford University ms.
- Bresnan, J. W., and J. M. Kanerva. 1989. Locative inversion in chichewa: a case study of factorization in grammar. *Linguistic Inquiry* 1–50.
- Butt, M. 1995. *The Structure of Complex Predicates in Urdu*. Stanford CA: CSLI Publications.

- Butt, M. 1997. Complex predicates in urdu. In , 107–149.
- Butt, M., M. E. Niño, and F. Segond. 1996. Multilingual processing of auxiliaries within LFG. In D. Gibbon (Ed.), *Natural Language Processing and Speech Technology*. Berlin: Mouton de Gruyter.
- Dowty, D. 1991. Thematic proto-roles and argument-selection. *Language* 547–619.
- Foley, W., and M. Olson. 1985. Clausehood and verb serialization. In J. Nichols and A. C. Woodbury (Eds.), *Grammar Inside and Outside the Clause*, 17–60. Cambridge University Press.
- Frank, A. 1996. A note on complex predicate formation: Evidence from auxiliary selection, reflexivization and past participle agreement in french and italian. available at:  
<http://www-csli.stanford.edu/publications/LFG/toc-lfg1.html>.
- Grimshaw, J. 1990. *Argument Structure*. MIT Press.
- Jackendoff, R. S. 1990. *Semantic Structures*. MIT Press.
- Kaplan, R. M., and J. Wedekind. 1993. Restriction and correspondence-based translation. In *Proceedings of the Sixth European Convergence of the Association for Computational Linguistics*, Utrecht. Available at: <ftp://parcftp.xerox.com/pub/nl/restricton.ps>.
- Manning, C. D. 1996a. *Ergativity: Argument Structure and Grammatical Relations*. Stanford CA: CSLI Publications.
- Manning, C. D. 1996b. Romance complex predicates: In defence of the right-branching structure. paper presented at the Workshop on Surfaced-Base Syntax and Romance Languages, 1996 European Summer School on Logic, Language and Information, Prague. Draft available at: <http://www.sultry.arts.edu.au/cmanning/papers>.
- Manning, C. D., I. A. Sag, and M. Iida. to appear. The lexical integrity of japanese causatives. In G. Green and R. Levine (Eds.), *Readings in HPSG*. CUP. Currently available at:  
<http://www.lcl.cmu.edu/~manning>, and soon at:  
<http://www.arts.su.edu.au/Arts/departs/ling/manning>.
- Matsumoto, Y. 1996. *Complex Predicates in Japanese*. Stanford CA: CSLI Publications.
- Rosen, S. 1989. *Argument-Structure and Complex Predicates*. PhD thesis, Brandeis University, Waltham MA.
- Wechsler, S. 1995. *The Semantic Basis of Argument Structure*. CSLI Publications/CUP. originally Stanford PhD dissertation, 1991.