

More Appendices for Sets, Heads and Spreading in LFG

Avery D Andrews

The Australian National University

September, 2018

This is some additional material for ‘Sets, Heads and Spreading in LFG’ (Andrews 2018; SHS, link here: <http://jlm.ipipan.waw.pl/index.php/JLM/article/view/175>), which did not seem entirely suitable for appearance in the published article, due to making less of a contribution to the essential point, and generating a certain amount of distraction. Nevertheless, I think that some people might find it interesting, so here it is. Much of this involves revisiting materials from Andrews and Manning (1993, 1999), which are therefore referred to as AM93 and AM99, respectively. Since the article has two appendices, the numbering here starts with ‘C’.

The major topics omitted here are the treatment of parallel modification of adjectives, and symmetric serial verbs in Tariana, for the reason that these are closely related to the analysis of coordinate structures. Some proposals about this were made in earlier drafts of SHS, available on lingbuzz (<http://ling.auf.net/lingbuzz/002522>), but I hope to give these topics substantially more work before too long.

C Membership as an Attribute

Since the introduction of ‘inside out functional uncertainty’ expressions in Dalrymple (1993), set-membership has led a shadowy semi-existence as a sort-of but not-quite attribute. In the classic formal LFG interpretation of an f-structure as a set, it really can’t be any kind of attribute at all, since the attributes are functions, while set-membership is a basic concept of the mathematical framework. With the introduction of hybrid objects, this also leads to a slightly puzzling distinction between sets that have nondistributive adjuncts (coordination), and those that appear not to (presumably, the traditional ADJUNCTS values). What I suggest here is that the notion of ‘set’ be removed from the inventory of basic formal ideas of LFG, and be replaced by the notion of ‘multivalued attribute’. Set-valued ADJUNCTS would then be replaced by multi-valued ADJUNCT, whose values would in fact constitute a set.

But with a subtle difference. In the standard LFG formalism, an f-structure is a function that assigns values to attributes, and so there is a

clear difference between a function that does not have ADJUNCTS in its range, and one that does, but assigns it the null set as its value. With the multivalued attribute proposal, this is not possible: the attribute has the empty set its set-value iff it has no f-structure as a value. This is a clear if minor advantage of the proposed approach.

But what then of the hybrid objects? These are typically constituents of some kind, and so are either the top-level f-structure, or bear some GF to a higher f-structure. The proposal is that these involve some multivalued attribute, and the obvious candidate is something more or less equivalent to ‘head’, which I will label ‘H’ (more or less following AM93). In ‘ordinary’ structures, H will in fact have a single value, whereas in coordinate structures it will have multiple values.

Therefore, in place of structures with set-membership, we have structures with an H-value, and ADJUNCT is replaced by ADJUNCT:

- (1) a.
$$\left[\begin{array}{l} \text{ADJUNCT} \left[\text{PRED} \text{ 'former'} \right] \\ \text{H} \left[\begin{array}{l} \text{ADJUNCT} \left[\text{PRED} \text{ 'unscrupulous'} \right] \\ \text{H} \left[\text{PRED} \text{ 'developer'} \right] \end{array} \right] \end{array} \right]$$
- b.
$$\left[\begin{array}{l} \text{H} \left[\begin{array}{l} \text{PRED} \text{ 'José'} \\ \text{PERSON} \text{ 3} \\ \text{NUM} \text{ SG} \end{array} \right] \\ \text{H} \left[\begin{array}{l} \text{PRED} \text{ 'pro'} \\ \text{PERSON} \text{ 1} \\ \text{NUM} \text{ SG} \end{array} \right] \\ \text{PERSON} \text{ 1} \\ \text{NUM} \text{ PL} \end{array} \right]$$

Distribution would then be a property of the H attribute, yielding an approach that is conceptually perhaps most similar to Andrews and Manning (1993), but with a far more restrained use of attribute-spreading.

I suspect that this restraint has become considerably easier to attain by increasing understanding of glue semantics, which provides the possibility for a looser connection between semantic composition and syntactic structure, with a correspondingly lesser tendency to try to give semantic significance to

the f-structure attribute labels. For example in AM93:22, there is a special attribute O for modal adjectives, but with the availability of a glue analysis along the general lines of Dalrymple (2001), the motivation for this disappears. This looser connection is illustrated strongly by the entirely glue-based treatment of Misumalpan in the next appendix.

D Misumalpan

Earlier drafts of Andrews (2018) on lingbuzz provide a reanalysis of the treatment of the Misumalpan languages Miskitu and Susmu from Andrews and Manning (1999) which is entirely glue based and makes no use of innovative attribute-spreading. I present this analysis here.

Basic examples of the Miskitu causative are:¹

- (2) a. [Yang yul ba ra yab-ri] wîna pi-n
 I dog the ACC give-OBV:ACT.1 meat eat-PAST.3
 ‘I made the dog eat meat.’ (MCD:29)²
- b. [Yang yul ba ra yab-ri] wîna pi-ras
 I dog the ACC give-OBV:ACT.1 meat eat-NEG
 ‘I didn’t make the dog eat meat.’ (MCD:29)

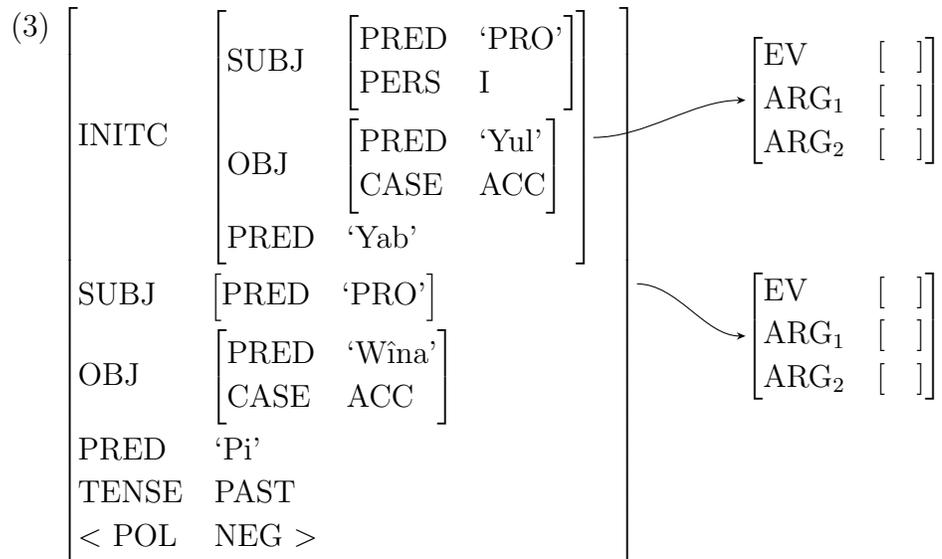
In these examples, the superficial form is that of a ‘consecutive’ structure, where the bracketted material is a morphosyntactically subordinate clause indicating what happens first, and the remaining material looks like a main clause, saying what happens next.

But the semantics of these constructions are different, and essentially causative. In particular, the negative applies to the entire construction rather than just the second apparent clause: (b) does not mean that I caused the dog to not eat meat, but rather to the entire proposition that I made the dog eat meat. We can continue to use the three-place causative predicate in the semantics, but the connection to the syntactic structure will have to be unusual (at least relative to current familiar analyses).

For the morphology to work, we would appear to want the first clause to be the value of some grammatical function, which we’ll call INITC (Initial Clause), with the second being the main clause:

¹AM99, based on work by Ken Hale and Danilo Salamanca.

²This is an undated handout by Ken Hale of Misumalpan causatives data.



From a traditional perspective, this seems like a rather poor prospect for semantic interpretation, but glue can manage it far more deftly than AM99 realized.

In the first place, there does not appear to be obligatory control between the arguments. In particular, we don't seem to need anything to enforce identity or coreference between the object of the Cause verb and the subject of the Effect verb (AM99, pp. 99-100). For example Bittner (1998:pp. 64-65) shows that there can be a variety of relationships between the Causee Object or its possessor, and some argument in the caused clause:

- (4) a. Upla kumi sin mai mun-an yul mai
 person one 'also' you(OBJ) cause-OBJ:ACT.3 dog you(OBJ)
 sam-ras kan
 bite-NEG PAST.3
 Noone will cause you to be bitten by the dog
 (Causee object = caused object)
- b. Witin upla kumi sin yula (ra) pruk-an law-ras
 he person one his dog (ACC) hit-OBV:3 get.angry-NEG
 kan
 PAST.3
 He didn't get anyone angry by hitting his_i dog (Possessor of causee
 object = caused subject)

In these examples, the fact that the subject of the first clause is within the scope of the negative marker indicates that this is a causative rather than a consecutive structure. Furthermore, AM99:100 note an example from Sumu where there is no coreference at all:³

- (5) Kâarak ârasyang dai, yang alas âranayang
 he.laugh.OBV.3 laugh.NEG.1 PAST I self laugh.PAST.1
 He didn't laugh me into laughing; I laughed by myself

Note that the obviatively marked verb is intransitive in the Sumu original, in spite of being rendered transitively in the free English translation. So AM99 conclude that there is no formal control or coreference requirement, but only a tendency, deriving from a requirement for a causal relationship.

Therefore the arguments can all be realized independently as NPs, with the further consequence that the Causee Agent does not have to be an ARG₃, but can be an ARG₂, and that is what it in fact appears to be.⁴ A rather simple meaning-constructor suffices:

- (6) $\lambda P y z . Cause(P)(y)(x) : (INITC \uparrow)_{\sigma} \rightarrow (\uparrow ARG_2) \rightarrow (\uparrow ARG_1) \rightarrow (INITC \uparrow)_{\sigma}$

Its mode of application is similar to that of a sentence-adverbial such as *apparently*, although it manages arguments that appear in its own clause. For both of the clauses, virtually any contemporary linking theory will suffice.

There is one interesting issue which the present literature does not entirely settle, as far as I am aware, which concerns the scope of negation. If we assume that the causative verb takes an argument of type *p*, which the caused verb provides, we predict an ambiguity in examples such as (5) and (2b), which apparently does not occur. This is a *prima facie* problem for the glue analysis, for which a possible solution would be to use more types. For example if the causatives complements were of an ‘Event’ type, and the negation only applied to ‘Propositions’, the scopes would be restricted as stated in the literature. On the other hand, there are some potential examples that could use further investigation. Causation of a negative might be semantically suspicious and therefore rejected, but permissive verbs also occur in the causative construction:

³From an undated ‘miscellaneous causative data’ handout compiled by various people including especially Danilo Salamanca, and distributed by Ken Hale.

⁴Note that because the Miskitu accusative case-marker *ra* is highly multifunctional, also marking various kinds of obliques, it is the possibility of object agreement with the verb that indicates ARG₂ status of the Causee Agent.

- (7) Witin yang ra ai swi-n skol ra wa-ri
 he me ACC me let-OBV:ACT.3 school ACC go-PAST.1
 He let me go to school

Nobody appears to have investigated whether negation of the caused verb here could produce the meaning ‘he let me not go to school’. So there is a margin of doubt about the nature of the facts here, and whether a distinction between *p* and *ev* types is truly required (although, this is very useful elsewhere, and, I would guess, probably correct here). My suspicion is that there is some work to be done in working out exactly what the semantic types involved in causative constructions are.

E Tariana

Tariana is an Arawak language described by Aikhenvald (2003), henceforth Aikh2003, whose serial verb constructions (SVCs) were a major topic of AM99 (on the basis of a number of papers and personal communications; the grammar does not change the relevant parts of the landscape in any significant way). Aikhenvald distinguishes on mostly semantic grounds a rather large number of SVC types, of which AM99 was and this paper will be concerned with only a few: ‘symmetric’, and certain ‘asymmetric’ SVCs including ‘modal’ and ‘causative’.

E.1 Types of Serialization

Symmetrical serialization is defined by Aikh2003:424 as consisting of two or more open class verbs, where none of the components is uniquely responsible for determining the semantic or syntactic properties of the construction, but rather all are on an equal footing. A typical example is:

- (8) ma [wa-wa wa-dana] wa-yarupe=nuku
 let’s 1PL-read/play 1PL-write 1PL-thing=TOPIC
 ‘Let’s read and write up our language!’ (Aikhenvald p.c.)

AM99 treat these as being essentially a kind of coordinate structure, which still seems to me to be a viable treatment. However a revision of the treatment of coordination provided by Asudeh and Crouch (2002) is not yet ready to go, so I do not consider these further here.

More unusual are the ‘asymmetric’ SVCs, many with the semantics of (VP) complement constructions, whose striking feature is that all of the verbs show person and number agreement with the syntactic subject of the clause, regardless of the semantic role that that argument bears to them:⁵

- (9) ka:ru-ka nuha [nu-a=mahka nu-hyã=niki]
 fear-DECL I 1SG-give=RECPAST:NONVIS 1SG-eat=COMPLT
 piri=nuku di-a=pidana
 2SG.son=TOPIC 3SGNF-say-REMPAST:INFR
 ‘Being afraid, I let (the fish) eat your son, he said.’
 (asymmetrical SVC, causative semantics, Aikh2003:425)

Another example, revealing at least some capacity for recursive combination, is:

- (10) nu-na=tha nu-ra nu-sata dineiru
 1SG-want=FRUSTR 1SG-order 1SG-ask money
 ‘I want to order (him) to ask for money.’
 (Modal on Causative; elicited, Aikhenvald p.c.)

The symmetric and this asymmetric type need significantly different treatments, but building the f-structures with set-membership will cause the grammatical relations to share/distribute fully, accounting for the agreement pattern.

E.2 C- and F-structure

For the asymmetric constructions, following the grain of AM99, the obvious thing to do would be to characterize the semantically subordinate member of an asymmetric SVC as a member, and the other as a top-level head. On the one hand, there is no real evidence for this treatment of the semantic head. In particular, there doesn’t appear to be any general constraint on the relative positions of the semantic head and complement (putative set member), which instead seems to be determined by the individual verbs, mostly in accordance with semantic classes. So for some verbs, such as the ones we’ve seen so far, the semantically head comes first, the other second; for others, the order is reversed:

⁵This phenomenon is called ‘concordant dependant inflection’ (Durie 1997).

- (11) tuiRi-kere na-hwa nema
 bird-island 3PL-stay 3PL.stand
 ‘They stayed at Bird Island for a long time.’
 (Aspectual; Aikhenvald 1999, p. 480)

Aikhenvald argues very plausibly that the ordering restrictions are based on iconicity of the historical word orders of the constructions, but that they are synchronically just facts. Introducing only one of them as a member doesn’t seem to add anything to this explanation. But neither does it cause any problems, and, furthermore, it allows the meaning-constructors to be slightly simpler than if both are members. Therefore we will adopt it, but also allow both to be members for the analysis of symmetric SVCs.

Therefore, the proposed annotated c-structure rule for SVCs is:

$$(12) V \quad \rightarrow \quad (V)^* \\
\langle \downarrow \in \uparrow \rangle$$

In the first place, we interpret this expansion of V as an alternative to lexical insertion, which produces simple verbs. A constraint against unfilled nodes will then block an empty expansion of the $(V)^*$, and glue semantics together with the Offline Parseability Constraint will then require that at least two V ’s are produced.⁶

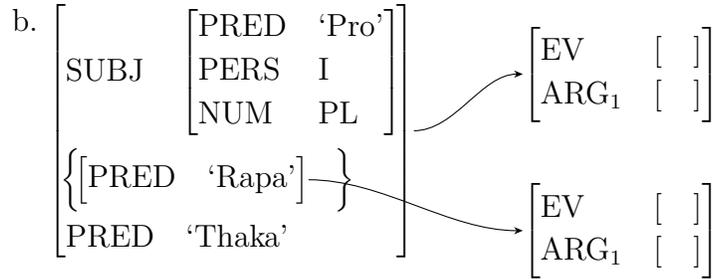
These can then appear with various combinations of the $\uparrow = \downarrow$ annotation, which we take to be present by default if nothing else is specified, and $\downarrow \in \uparrow$. This analysis will extend to symmetric serializations, which we can assume to have all their members introduced with the set-member annotation. But the rule (12) overgenerates substantially, and so needs to be constrained by glue.

E.3 Glue Analysis

Interestingly, in spite of the surface differences, the meaning-constructors for the Tariana ‘light verbs’ can be exactly as they are in Romance, since the ‘semantic complement’ is specified as a member, and the NP arguments are associated with s-structure ARG-attributes. We illustrate with a simple example with an intransitive light verb, effects of the Kibort-Findlay Lexical Mapping theory of linking not yet represented, although the semantic projection/argument structures are in the format for that:

⁶If one is produced with out the annotation, it will be ruled out by Offline Parseability, if one is produced with the annotation, the glue semantics won’t be able to effect assembly.

- (13) a. [wa-Rapa wa-thaka] wha
 1PL-dance 1PL-stop we
 We stopped dancing for a while
 (Aspectual; Aikhenvald 2003, p 433)



The relevant constructors will be as below:

- (14) $\lambda yxe. Rapa(e, x, y) : (\uparrow ARG_1) \rightarrow (\uparrow EV) \rightarrow \uparrow_\sigma$
 $\lambda Pxe. (\exists e_2) Thaka(e, x, P(e_2, x)) : [(\uparrow \in ARG_1) \rightarrow (\uparrow \in EV) \rightarrow (\uparrow \in)_\sigma] \rightarrow$
 $(\downarrow ARG_1) \rightarrow (\uparrow EV) \rightarrow \uparrow_\sigma$

These will assemble the same way, and linking can work in the same way, as it does in Romance.

For causatives, however, although the glue is the same, there is an apparent, but not, I claim, an actual difference in the linking. The apparent difference is that causatives take what might be seen as two objects, both in the ‘accusative’ case (Aikh2003:275):

- (15) na-na wa-yaRuphe-nuku [ma-sape-kade-ka
 3pl-OBJ our-language-TOPIC.NON.A/S [NEG-speak-NEG-RECP.VIS
 na-pala]
 3pl-put]
 She did not teach them our language

-*na* here marks pronominal non-subjects, while -*nuku* marks non-subject topics, both applying to both the theme and recipient of a ditransitive verb. The location of the negative in the SVC is also interesting, but I won’t pursue that here. The problem is that it looks like we might have two objects, both the Caused Patient/Theme and the Causee Agent.

But this is not clearly motivated, because Tariana has apparent double object constructions where the Theme appears to acquire the object properties, even though the Recipient looks the same (Aikh2003:236-238, 143-148). An example is:

- (16) na-na kuphe-nuku di-walita
 3pl-OBJ fish-TOPIC.NON.A/S 3sg-offer
 He offered them fish

But only the Theme can be passivized (Aikh2003:236, 259), indicating that given AGT, it is ARG₂ linked to OBJ. But then the Recipient will be ARG₃ realized as OBJ_θ, which lacks overt properties clearly distinguishing it from OBJ.

So now, if the application of KFMT in section 4.4 of SHS is basically correct, we derive the predication that if passivization can apply at all to causatives, only the Caused Object, not the Causee Agent, will be able to be passivized. This is because the linking theory naturally assigns ARG₂ to the Caused Object, allowing it to be realized as subject, leaving only ARG₃ for the Causee Agent, which should not be passivisable given Aikhenvald’s account of ditransitives. Aikhenvald’s grammar does not indicate whether causatives can be passivized, so this is a genuine prediction rather than a retrodiction of previously known facts from theoretical ideas devised later. But it is in accord with how things work in Romance languages, where passive is either inhibited with causatives, or applies only to the Causee Objects.

A further, somewhat unsettling, consequence of the framework that this analysis reveals that ‘concordant dependent inflection’ is actually what is expected if the caused verb in a clause union construction is of the same morphological type, e.g. ‘finite’, as a main verb. That is, the reason that the subordinate verbs in Romance Complex Predicates don’t show person-number agreement with the grammatical subject of the whole construction is perhaps just that they are ‘non-finite’ rather than ‘finite’. Since concordant dependent inflection does not seem to be especially common, this is something that needs closer investigation. An alternate way of blocking concordant dependent inflection is with an undersharing specification for SUBJ, which can be historically motivated by positing a non-clause-union origin for the constructions (clearly correct, in the case of Romance).

E.4 Linear Order

But, beyond the differences in c-structure and verb marking, there is a further difference between Romance and Tariana, which is the need for some way to express the verb-order restrictions in Tariana, which plausibly involve iconicity as their diachronic explanation, but are synchronically just facts

that have to be stated. Our asymmetric c- and f-structures make it slightly easier to state the restrictions. These are stipulative, but this can be regarded as not a serious problem, because each appearance of one of these verbs indicates what the order should be for it, so any theory relying in a sensible way on real-time prediction or MDL data-scores should be able to explain their acquisition.

I propose that a reasonable way to formulate the constraint is with a constraining equation with a universal quantifier, using inverse projections to refer to linear order. For a verb like *thaka* that follows its semantic complement, what we want is that if a c-structure node's f-structure correspondent is a member of the set component of *thaka*'s f-structure, then it precedes the c-structure node that *thaka* is introduced under. ' $\hat{*}$ ' is a standard notation for the c-structure node that a lexical item is introduced under, and we can use $\phi(c)$, $\phi(\hat{*})$ to represent the f-structure correspondent of an arbitrary c-structure node c and $\hat{*}$, respectively, leading to:

$$(17) \phi(c) \in \phi(\hat{*}) \Rightarrow c \ll \hat{*}$$

This will be part of the lexical entry of verbs that come after their semantic complements in an SVC, while those that come before will be the same but with the ordering statement of the consequent of the conditional reversed.

References

- Aikhenvald, A. Y. 1999. Serial verb constructions and verb-compounding in Tariana. *Studies in Language* 23:469–497.
- Aikhenvald, A. Y. 2003. *A Grammar of Tariana*. Cambridge University Press.
- Andrews, A. D. 2018. Sets, heads and spreading in LFG. *Journal of Language Modelling* 6:131–174. url=<http://jlm.ipipan.waw.pl/index.php/JLM/article/view/175>.
- 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. <http://nlp.stanford.edu/~manning/papers/proj.ps> (checked 5 Oct 2016).

- Andrews, A. D., and C. D. Manning. 1999. *Complex Predicates and Information Spreading in LFG*. Stanford, CA: CSLI Publications.
- Asudeh, A., and R. Crouch. 2002. Coordination and parallelism in glue semantics: Integrating discourse cohesion and the element constraint. In *Proceedings of the LFG02 Conference*, 19–39, Stanford, CA. CSLI Publications. <http://csli-publications.stanford.edu>.
- Bittner, M. 1998. Concealed causatives. *Natural Language Semantics* 7.
- Dalrymple, M. 2001. *Lexical Functional Grammar*. Academic Press.
- Dalrymple, M. E. 1993. *The Syntax of Anaphoric Binding*. Stanford CA: The Center for the Study of Language and Information.
- Durie, M. 1997. Grammatical structures in verb serialization. In *Complex Predicates*. Stanford CA: CSLI Publications.
- Hale, K., et al. undated. Misumalpan Causatives Data handout, collected from various sources and distributed by Ken Hale.