

# Idioms in Lexical Functional Grammar and Glue Semantics

Avery D Andrews  
v3.1, ANU, Feb 2025

In this paper, I will apply the proposals about semantic interpretation in LFG+glue advanced in Andrews (2007, 2008, 2019) to some problems in the semantic interpretation of idioms. These problems are:

- (1) a. idioms such as *make headway*, where one of the components has restricted distribution, while the other seems to have its ordinary meaning
- b. idioms such as *Penny's paper got up the reviewer's nose/reviewers' noses*, where it can be argued that the components of the idiom have their literal meanings as well as a role in the idiomatic one
- c. the prepositional and double passives such as *Mary's warnings were made light of* (which will require a look at the much-discussed issue of the passivizability of some idioms and not others)

The theory will require a small modification, but seems able to meet the challenge.

It can be seen as an alternative to the approach of Findlay (2019), which requires replacing the c-structure component of LFG with Tree-Adjoining Grammars. Instead, we change the nature of the way in which glue semantics interacts with the syntax, having semantic contributions introduced on the basis of constellations of items in f-structure, rather than individual lexical items as in current mainstream glue, as presented for example in Asudeh (2023) or Dalrymple (2001, ch 9).

In the first section I will present LFG+glue, using Andrews' technique for introducing the glue meaning contributions. Then in the following sections I will take on various problems, and show how, with a slight modification, Andrews' technique can deal with them. In the final section I will observe that the approach has at least some degree of transportability to Minimalism, and briefly discuss some of its general features.

## 1 LFG+glue

The basic LFG architecture consists of two levels, c-structure, and f-structure. The former is ordinary phrase structure,<sup>1</sup> the latter a structure made up of un-

---

<sup>1</sup>Technically, not produced by phrase structure rules as in mathematical linguistics, but by recursive transition networks, since they allow node types to be repeated indefinitely by means of the  $()^*$  notation.

ordered sets of pairs of attributes and their values. Some attributes are grammatical functions/relations, such as SUBJ and OBJ, whose values are other f-structures, and others are grammatical features, such as TENSE and NUM, whose values are atomic symbols. The f-structures are produced from the c-structures by annotations on the c-structure rules, in a manner that we don't need to consider quite yet. For a comprehensive introduction, see Belayev (2023). Especially important amongst the features was PRED, whose values were taken to be lexemes, comprising a morphological stem and a meaning. A sample f-structure for *Fido chased Spot* would be:

$$(2) \left[ \begin{array}{ll} \text{SUBJ} & \left[ \text{PRED} \quad \text{'Fido'} \right] \\ \text{TENSE} & \text{PAST} \\ \text{PRED} & \text{'Chase'} \\ \text{OBJ} & \left[ \text{PRED} \quad \text{'Spot'} \right] \end{array} \right]$$

This is assuming a 'privative' account of features whereby the singular value of number (NUM) is represented by the absence of a value, and also drops the assumption that the PRED-feature encodes the valence of the predicate, a function that is increasingly thought to be taken over by the glue semantics.<sup>2</sup>

In very early LFG as presented for example in Bresnan (1982a), there was no explicit theory of semantics, although the PRED-features were assumed to encode meaning in some way. The treatment of idioms and prepositions and particles that weren't making a clear individual contribution to the meaning was that they lacked PRED-features, but introduced other attributes such as FORM (for parts of idioms) and PCASE (for prepositions). A bit later, a kind of semantics was provided by Halvorsen (1983), but this was not capable of managing scope ambiguities such as in *John doesn't understand many of the problems*, because it had no way to create a semantic structure that was less flat than the f-structure.

Later, in the 1990s, a more powerful kind of semantics, glue semantics, was developed, as surveyed in Dalrymple (1999), Dalrymple (2001), Asudeh (2023). Idioms did not play a prominent role in the development of mainstream glue, which Andrews (2007, 2008, 2019) did not find satisfactory, and made a proposal whereby PRED and other features, sometimes in constellations, would introduce semantic contributions via a 'semantic lexicon' distinct from the morphological lexicon. Findlay (2019) proposed a more radical change to LFG, replacing its c-structure rules with TAGs, with a major focus on idioms and other multi-word expressions (MWEs). Andrews' proposal can be regarded as

---

<sup>2</sup>See Andrews (2008) for discussion of the idea that PRED should be retained, but primarily for the morphology.

an attempt to capture many of the properties of Findlay’s, with an arguably less extensive revision of the LFG framework, especially to the older parts of it.

In mainstream glue (Dalrymple, 2001; Asudeh, 2023), lexical items including any inflections introduce ‘meaning constructors’, which consist of contributions to the meaning combined with specification of connections to the syntax that restrict the ways in which the meaning contributions combine, so that the syntax can give different interpretations to *Spot chased Fido* and *Fido chased Spot*. The basic idea of Andrews (2007, 2008, 2019) is that these constructors do not appear as parts of individual conventional lexical items appearing in overt structure (‘Codescription’ in the sense of Halvorsen & Kaplan (1988)), but as elements of a ‘Semantic Lexicon’, that introduces meaning-contributions on the basis of constellations of feature-values in f-structure.

An SLE consists of two parts:

- (3) a. A functional description indicating what kind of environment the SLE can apply to. An extremely important point is that the functional description is entirely constraining: it cannot deposit things into the f-structure.<sup>3</sup>
- b. A ‘semantic contribution’ presenting a component of the meaning, to be assembled with such other components as may be introduced by other SLEs

In the hope of easier comprehension, both components will be presented diagrammatically. However, f-structure (constraining) equations could be used for the functional description (and could be thought of as the ‘official’ representation, and a box notation inspired by Construction Grammar will be used for the semantic contributions. These boxes will express the meanings as formulas in the usual formal semantics notation of simply-typed lambda-calculus. The box notation is a reformulation of proof nets,<sup>4</sup> which is an alternative formulation for a certain range of linear logic proofs, more limited. The components of the meaning contribution will be connected to (some of) the components of the functional description by association links, which will allow the syntax to constrain the semantic assembly, without necessarily dictating it.

A very simple sample SLE is:

$$(4) \left[ \text{PRED} \quad \text{'Fido'} \right] \overset{\dots\dots}{\text{---}} \boxed{\text{Fido}}$$

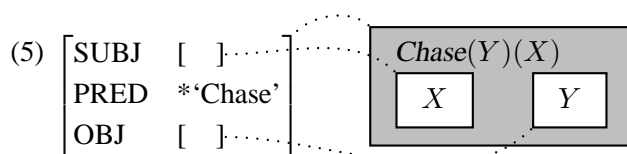
An SLE such as this one can be regarded as ‘licensed’ by the feature values it mentions, and every ‘interpretable’ feature must be involved in the licensing of

<sup>3</sup>If an SLE could add feature values to f-structure, then these would want interpretation, allowing application of more SLEs, allowing the f-structure to grow indefinitely.

<sup>4</sup>See Moot (2002) for an overview of these.

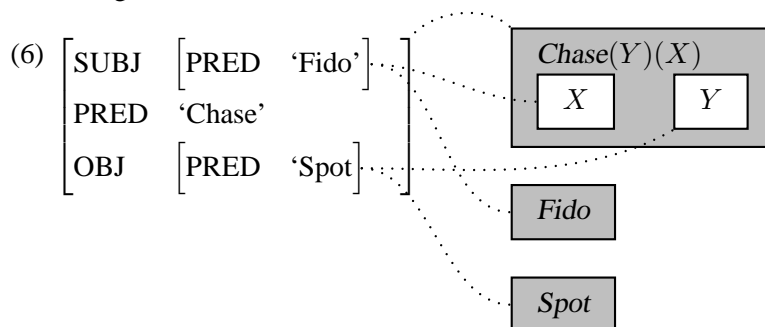
at least one SLE. Some features, the ones that aren't interpretable, such as syntactic case, can be licensed in other ways; I will not look into that here. Semantic assembly is controlled by association lines from the boxes to f-structures in the functional description; this is a function from the boxes, but is neither one-to-one nor onto the f-structures. A change from the earlier versions is that a single feature is able to license more than one SLE; in the discussion of the *nose* idiom we'll see why.

The SLEs for verbs are more complicated. A verb with arguments has one unshaded box for each argument, contained within the shaded box, and all the boxes have association links to substructures in the f-description. The upper case variables in the unshaded boxes are usually repeated in the argument positions of the meaning formula of the shaded box:



The upper case variables can be eliminated, but without them, we would have to stick to some conventions for ordering the unshaded boxes to indicate argument positions, which is sometimes inconvenient for a graphical presentation.<sup>5</sup>

So now, supposing that *Spot* gets the same kind of treatment as *Fido*, and ignoring tense, we have enough to explain how meaning-assembly for (2) would work without the TENSE PAST feature. The SLEs introduce meaning contributions as follows, including with the dotted lines which now connect boxes in the meaning contribution to substructures of the full f-structure:



We perform semantic assembly by drawing links from shaded to unshaded boxes, obeying four rules, three of which can be seen as intrinsic to semantics, the third concerning the way in which the syntax constrains the semantics. These rules involve two kinds of links, the explicit ones we add, from shaded

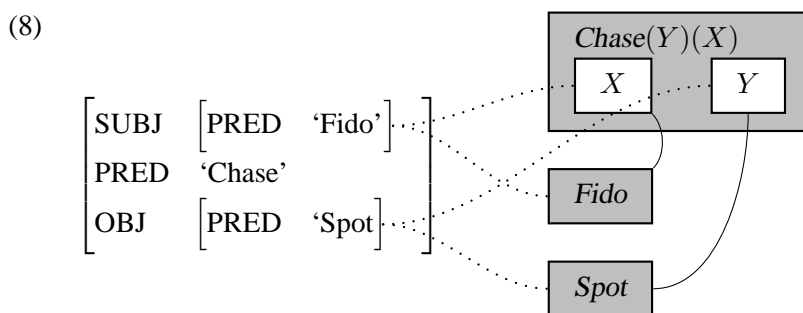
<sup>5</sup>This can be convenient if you actually want to write the completed semantic values into the boxes, as noted by Andrews (2024).

to unshaded boxes, which are called ‘axiom-links’, because they represent application of the axiom in the Gentzen Sequent Calculus, and implicit ones we can think of as already there, connecting an unshaded box to its immediately containing unshaded box. These links form a system of paths, which must obey the following rules:

- (7) a. Each unshaded box must be linked to one and only one shaded box.
- b. There is one and only one shaded box (the ‘root’) with no axiom-link to an unshaded box.
- c. From every shaded box, there is a unique path to the root (the path consisting of zero steps, for the root). This path will in general include implicit links as well as axiom links.
- d. In order to be axiom-linked, a shaded and unshaded box must be connected to the same f-structure.

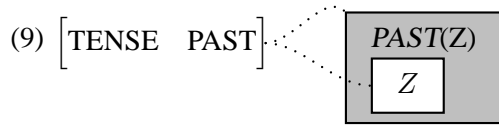
Condition (a) prevents an unshaded box from being linked to nothing (ie, unfilled argument positions), (b) requires all the contributions to link into a single structure (rather than, say, two shaded boxes representing name meanings, with no relationship indicated between them), while (c) rules out a variety of configurations that don’t produce sensible meanings. Finally, (d) is how the syntax constrains the semantic assembly.

So, given (6), there is only one way of adding the axiom-links that satisfies all the constraints:

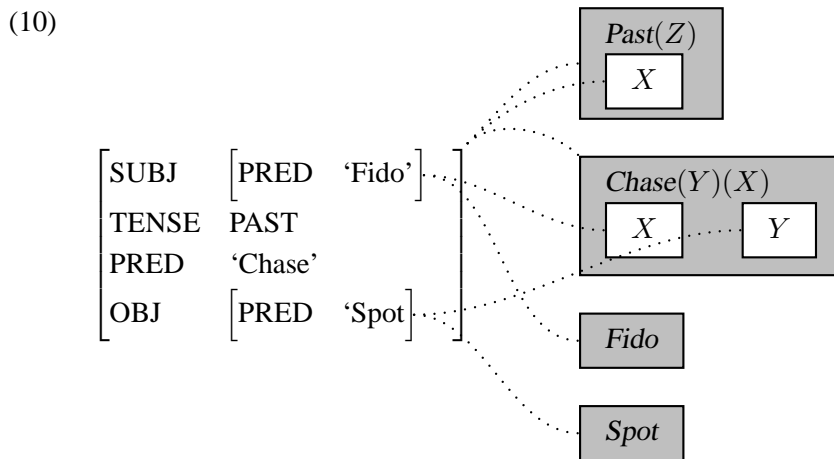


Given the axiom links, if we substitute in (8) the semantic values of the shaded boxes for all of the occurrences the variables in the unshaded boxes they are linked to (the one inside the box and the repeated one outside of it), the reading we get for the whole is  $Chase(Fido)(Spot)$ , which is what we want.

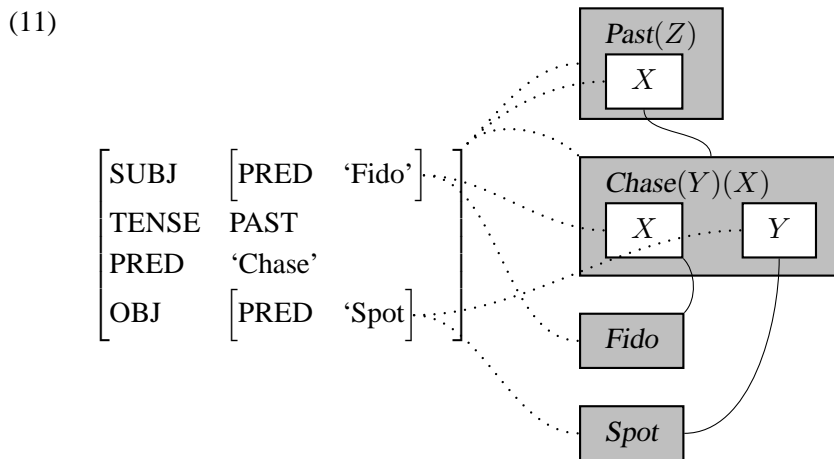
The use of (c) as well as the capacity of glue to build hierarchical structure that is not present in the f-structure can be seen by including a meaning for the PAST feature. An SLE for it might be:



Now the f-structure below and the SLEs will produce the following collection of semantic contributions with their association links to f-(sub)structures:



And, as before, there is only one way to hook them up:



This can build semantic structure that is hierarchically richer than the f-structure because the meanings are not in any clear sense ‘located’ at the f-structure nodes; the associations to f-structure merely constrain the ways in which unshaded and shaded boxes can be connected by links.<sup>6</sup> Finally, (7c) rules out

<sup>6</sup>Perhaps a bit like a rather officious dating agency which prescribes which dates are possible on the basis of some compatibility criteria, but has no involvement in whether they happen, or, if they do, any interest in what happens on them.

another way of hooking up the boxes that satisfies all the others; finding that could be an exercise for the reader.

Now, repeated substitution as with (9) produces  $Past(Chase(Fido)(Spot))$ , which is what we want. However there is another way of linking up, obeying all the rules of (7) except for (c), for which the substitution procedure does not yield a single formula, and there is a variable which never gets anything (other than itself) substituted for it. Can you find this structure?

What we have seen so far is very close the Construction Grammar framework of Kay & Fillmore (1999); in fact it might not be wrong to say that without knowing it, they invented a version of proof-nets for the easy half of intuitionistic implicational linear logic, representing implication Left in the Gentzen sequent calculus, or Implication Elimination/Modus Ponens in natural deduction. But they did leave something out, and so far, so have we, which is types. Formal semantics normally assumes a classification of semantic elements into types, basic and composite. Usually, at least two basic type are assumed,  $e$  for ‘entity’, and  $t$  for (something like) ‘proposition’, essentially, something that can be true or not.<sup>7</sup> Composite types are constructed with type-constructors; the only one we will use is ‘ $\rightarrow$ ’, function formation. The type  $a \rightarrow b$  is the type of a function into which you drop in something of type  $a$ , and get out something of type  $b$  (a machine that returned a candy bar when you feed it a \$5 bill might be described as being of type  $\$5\ bill \rightarrow\ candy\ bar$ ).<sup>8</sup> In a very (almost certainly too) simple system, names would be of type  $e$ , while common nouns, adjectives and intransitive verbs would be of type  $e \rightarrow t$ . That is, something that when supplied with an entity, tells you whether it is true that it has whatever property the linguistic item designates.

To add types to our system, we want every box to have a type, and then add to the hookup rules (7) the additional requirement that the axiom links can only connect boxes with the same type. The types were essentially to early glue, but various ways have been found to avoid using them, and they tend to be omitted in notation. But we need the ideas to understand the treatment of quantifiers such as *every*. This is that *every* designates a relation between two properties (type  $e \rightarrow t$  where by the entities for which the first property returns ‘true’ are a subset of those for which the second does. Therefore, the type of *every* is  $(e \rightarrow t) \rightarrow (e \rightarrow t)$ , usually written as  $(e \rightarrow t) \rightarrow e \rightarrow t$ , using a convention of omitting rightmost parentheses, including outermost.

Montague’s analysis of quantification requires a further idea, which is bound variables and lambda-calculus. In a simple use of a quantifier such as *every*,

---

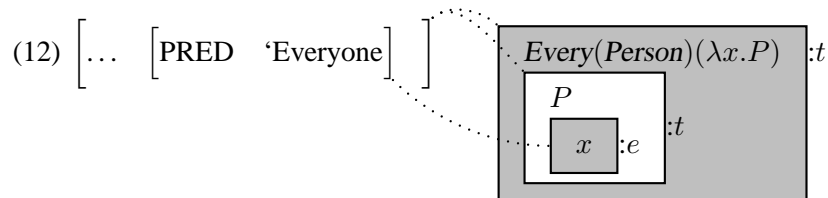
<sup>7</sup>Casadio (1988) is an interesting discussion of the history of type theory, while Partee (2006); Partee & Borschev (2004) provide further comment.

<sup>8</sup>In most formal semantics literature, angle brackets are used instead of the arrow, so  $a \rightarrow b$  would be written  $\langle a, b \rangle$ .

it is applied to a common noun, so we can reasonably write something like  $Every(Dog)$  for *every dog*, or  $Every(Person)$  for *everybody*. But then these combination appear inside a proposition, as in *Rover loves everybody*. The first step in managing this is to create a property from a proposition, which is what lambda-abstraction does. Put a variable, say,  $x$ , in the position of *everybody*, and put ‘ $\lambda x.$ ’ in front of the resulting formula, to get  $\lambda x.Loves(x)(Rover)$ . The intuitive understanding of this is that you get a proposition by substituting something concrete for the variable.

To put this into our box notation, we need to go substantially beyond Fillmore and Kay, by allowing an unshaded box to contain one or more a shaded ones, each of which must be labelled with a different lower case variable. These are different from the upper case substitutional variables that we have used so far, and cannot be easily eliminated.<sup>9</sup>

To manage this, we go beyond anything that could reasonably be attributed to Kay and Fillmore, which is to allow shaded boxes inside unshaded ones. These are interpreted as producers of lambda calculus variables (in lower case, to distinguish them from the naive substitutional variables we have been used so far), that wind up getting bound by lambdas. Then, in the semantic formula for the whole, these variables lambda-bind whatever gets substituted for the containing unshaded box’s upper case variable. So the SLE for *everyone* can be written as follows, where we have also included the types to the upper right of the boxes, preceded by a colon:



where the ... means that the outer brackets represent some f-structure that contains the one whose PRED is ‘Everyone’.

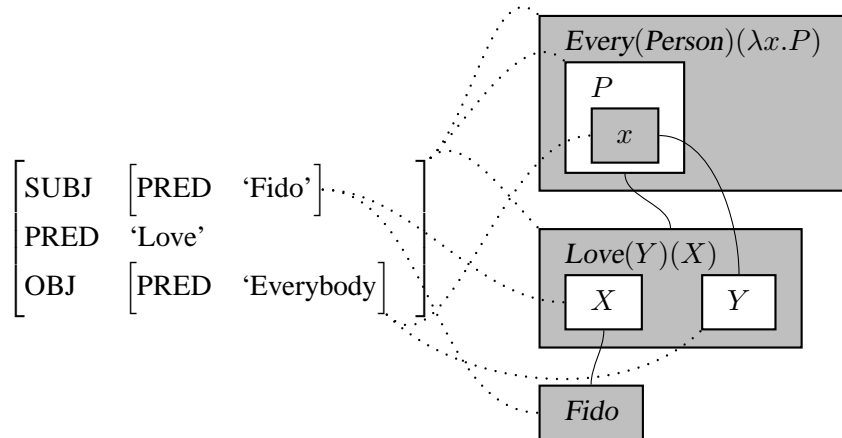
The overall effect is to inject the variable  $x$  into a semantic position associated with the NP *everybody*, collect a meaning from some larger, containing, f-structure, and wrap in a lambda-formula in  $x$ , and supply this as argument to *everybody*:

---

<sup>9</sup>I think it might be possible to do natural language semantics with only substitutional variables, rather than lambda calculus variables, but so far as I can make out, it is messy and dubious.



(13)



Here, by the given links,  $P$  would evaluate to  $Love(x)(Fido)$ , which gives the right result when we plug in into the  $P$  position in the *Everybody* formula.

But to assure only links that yield formulas where the variables are properly bound (and the diagrams represent proper proofs in linear logic) we need one more condition:

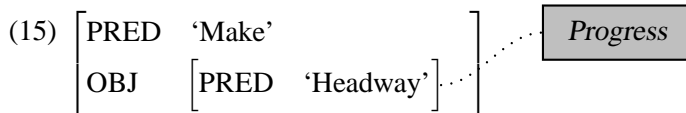
- (14) From a shaded box that is inside an unshaded box, there must be a path to its immediately containing unshaded box (by what we've already said, this path will have to be unique).

Note that there is *no* implicit path from an inner shaded box to its immediately containing unshaded box; the required path must be constructed using axiom links.

This has been a very quick introduction to a glue notation which, although limited in what it can do, is hopefully easier to understand than the usual ones, which use either explicit deductions, or, sometimes, proof nets in different formats.

## 2 Analysis of Some Idioms

Now that we have presented the basics of glue, we can look at some idioms and their analysis. Consider first *make headway*. Here, I propose that *headway* means 'progress' when it occurs as direct object of the verb *make*, which takes its ordinary meaning, or one of them, as in locutions such as *make progress*.<sup>10</sup> The proposed SLE is (15) below:



<sup>10</sup>We probably need to extend this a bit to deal with examples such as *the headway they made was insufficient*, but this involves complex issues of exactly how relative clauses work.

This associates a meaning with the direct object which therefore, as is well-known, can take various kinds of modification (*they made impressive headway*). No meaning however is provided to the whole f-structure, which means that even though the PRED-value ‘Make’ has been checked off, it still, must license the SLE for its normal meaning in order for semantic assembly to be successful.

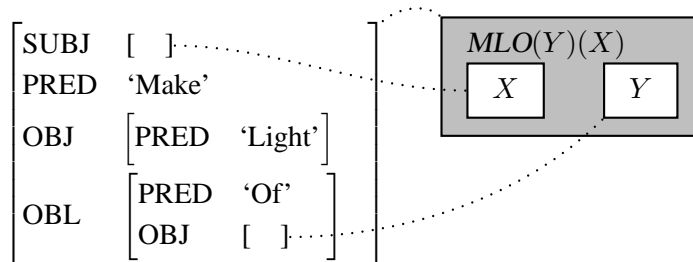
This idiom presents a problem for the very interesting proposal by Egan (2008) that idiomatic meanings are based on literal meanings via pretense. The problem is that for many speakers, including this author before he consulted a dictionary, *headway* has no literal meaning at all, and therefore nothing for a pretense to be based on. It also exposes a problem with Andrews’ earlier formulations, since, without some additional notation, the PRED ‘Make’ would be checked off, and so unable to introduce an additional meaning contribution, requiring (14) to be expanded to re-specify the normal meaning of *make*. We deal with this problem here by allowing (an occurrence of) a feature-value to license more than one SLE, but each SLE only once.

For a different situation, consider the idiom *make light of*. Here we have no possibility of modifying the (presumed) noun *light*, and *make* does not appear to have any of its ordinary meanings:

(16) Fred made (\*considerable) light of the situation

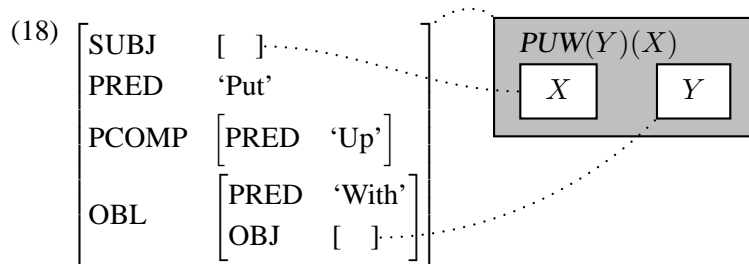
Here, I suggest, the meaning of *make light* is a predicate meaning ‘SUBJ communicates an attitude that the situation referred by the object of *of* is not serious’. This can be accomplished with the following SLE:

(17)



Another issue, which we will consider in the section on passives, is whether *light* is associated with some meaning-contribution supplied by the SLE, or is not, as depicted here. associated with any meaning contribution at all, as depicted here.

This approach extends naturally to the verb-preposition-particle constructions we discussed earlier. So the entry for *put up with* might be:



Here we have somewhat arbitrarily proposed that the grammatical function of the particle is PCOMP; the question of what the grammatical functions of particles and prepositional particles is complex, and will not be considered here.<sup>11</sup>

This analysis, unlike the previous one of *make headway*, checks off many features that do not introduce additional meaning-contributions; if they did, they would not fit into a well-formed hookup.

We will now proceed to our next topic, the *nose* idiom.

### 3 ‘Nose’ Idioms

What makes *X gets up Y’s nose* and several other idioms with *nose* interesting is that they provide concrete evidence that the literal meanings of the components are used in the construction of the idiomatic meaning:

- (19) a. Penelope’s paper got up the reviewer’s nose.  
 b. \*Penelope’s paper got up two of the reviewers’ nose.  
 c. Penelope’s paper got up two of the reviewers’ noses.

(a) has an idiomatic interpretation along the lines of ‘annoy severely’, and so does (c), where the noun ‘nose’ in the idiom is pluralized, in a sort of apparent agreement with its possessor. (b), on the other hand, without pluralization of ‘nose’, has no plausible interpretation whatsoever, idiomatic or literal. But English has no syntactic number agreement between possessors of NP and their heads, so the explanation for the required pluralization is presumably semantic, along the lines than given the way human anatomy works, human reviewers cannot share a single nose. But this cannot explain anything about the idiomatic meaning if the literal meaning is not used in some way to construct the idiomatic interpretation.

The analysis will make use of the literal meanings of the component words, assembled in the usual way, but adding the information that the possessor of the nose is reacting to the subject in a manner similar to somebody affected

<sup>11</sup>In Lexical Mapping Theory, they would be varieties of  $OBL_{\theta}$ , where  $\theta$  indicates some semantic role, but what the semantic roles are and why is not carefully worked out.

by something actually getting up their nose. We need to make reference to the actual words used because the idiomatic reading of nose idioms is triggered by specific lexical items in specific positions: if we substitute *climb up* or *crawl up* for *get up*, the idiomatic reading vanishes, and likewise if we substitute *nostril* for *nose*:

- (20) a. # Penny's paper climbed/crawled up the reviewer's nose  
 b. # Penny's paper got up the reviewer's (left) nostril

Observe that (a) would be fine if the grammatical subject was a small insect, (b) if it were some kind of dust or powder, as well as a small insect.

Therefore, the functional description needs to specify this material as follows:

- (21) 
$$\left[ \begin{array}{l} \text{SUBJ} \quad [ \quad ] \\ \text{PRED} \quad \text{'Get'} \\ \text{OBL} \quad \left[ \begin{array}{l} \text{PRED} \quad \text{'Up'} \\ \text{OBJ} \quad \left[ \begin{array}{l} \text{PRED} \quad \text{'Nose'} \\ \text{POSS} \quad [ \quad ] \end{array} \right] \end{array} \right] \end{array} \right]$$

Now if a patch of f-structure that looks like this gets interpreted in the normal way, we get a situation that can be depicted like this, basically as a 'derived' meaning constructor:

- (22)
- 
- $$\left[ \begin{array}{l} \text{SUBJ} \quad [ \quad ] \\ \text{PRED} \quad \text{'Get'} \\ \text{OBL} \quad \left[ \begin{array}{l} \text{PRED} \quad \text{'Up'} \\ \text{OBJ} \quad \left[ \begin{array}{l} \text{PRED} \quad \text{'Nose'} \\ \text{POSS} \quad [ \quad ] \end{array} \right] \end{array} \right] \end{array} \right]$$
- Get\_up\_nose(Y)(X)  
 (literally)  
X Y

On the right hand side we have not represented the actual assembly of the meanings provided by the left, only provided a convenient representation of the result, which is to be interpreted literally. Note that the dotted lines are simply the ones from the original meaning-constructors that have not yet been utilized to guide assembly. This demonstrates a similarity between glue semantics and Construction Grammar (which does not depend on the present notation): an assembled combination of units functions in further combinations in the same way as a basic one.

So on with the proposed analysis, which is going to be that the idiomatic sense is basically a simile, for which a first attempt at an account might be:

(23) *Y reacts to X like X is something small and [X has gotten up Y's nose].*

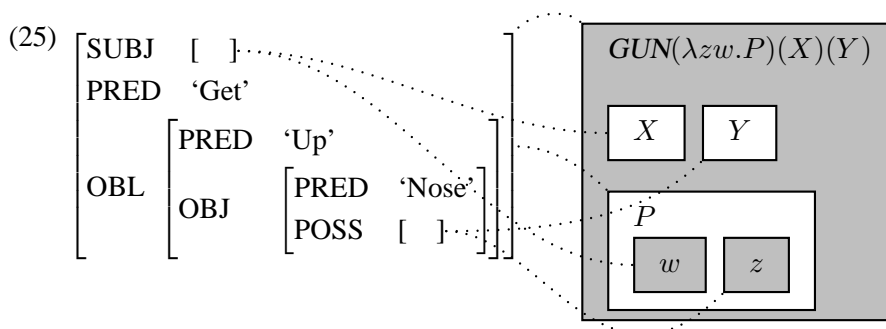
where the portion in brackets is the literal meaning, constructed in the usual way. However, I think it is interesting to try to state the meaning using Natural Semantic Metalanguage, of which 'react' is certainly not an element, so here is a revision:

(24) *Y thinks that X is bad.*

Because of this, *Y* does things like (as if) *X* has gotten up *Y*'s nose

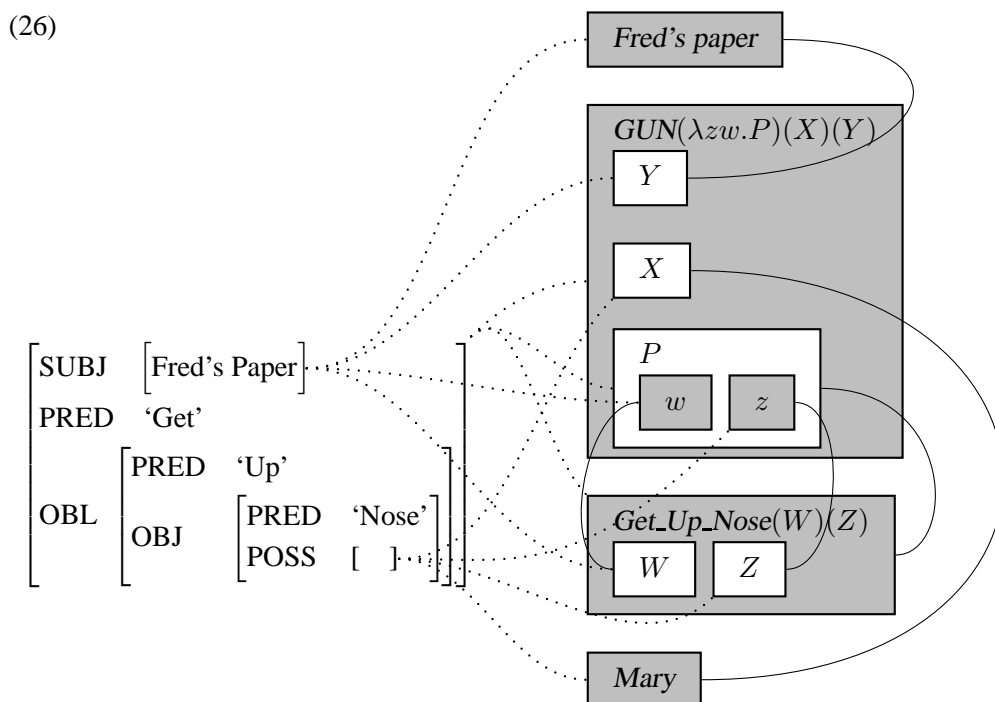
This explication undoubtedly needs more work, but it's a start. It's also an explication only of the desired result, not an account of how we are to get it. This, unfortunately, requires a somewhat complex application of lambda-calculus.<sup>12</sup> But before we embark on this, I will comment on some inherently confusing aspects of the semantic roles involved. For the literal meaning of *X* getting up *Y*'s nose, if for example, *X* is a fly, then *X* is taking the more active role, so we would want to apply it last, according to the most usual (but not universal) convention, at least somewhat supported by the considerations adduced in Marantz (1984). But in the idiomatic sense, the possessor of the object might be regarded as the more active argument, since awareness and some kind of activity is attributed to it. but the subject is still in some sense the causer, and so also has a claim to be the more active participant. In an attempt to reduce the confusion a bit, we will apply the subject last, same as in the literal meaning.

The meaning contribution that produces the idiomatic sense needs to take three arguments, the subject, the object' possessor, and the literal meaning, packaged as a function with two arguments. We'll represent this as *GUN*, introduced by this SLE:



<sup>12</sup>Perhaps somebody will be able to come up with a better way to do this, but lambda-calculus is well-understood and worked out.

To see how this works, here is a somewhat simplified example, where the subject NP and the literal meaning of the *get up Y's nose* combination are treated as blocks, and we ignore tense:

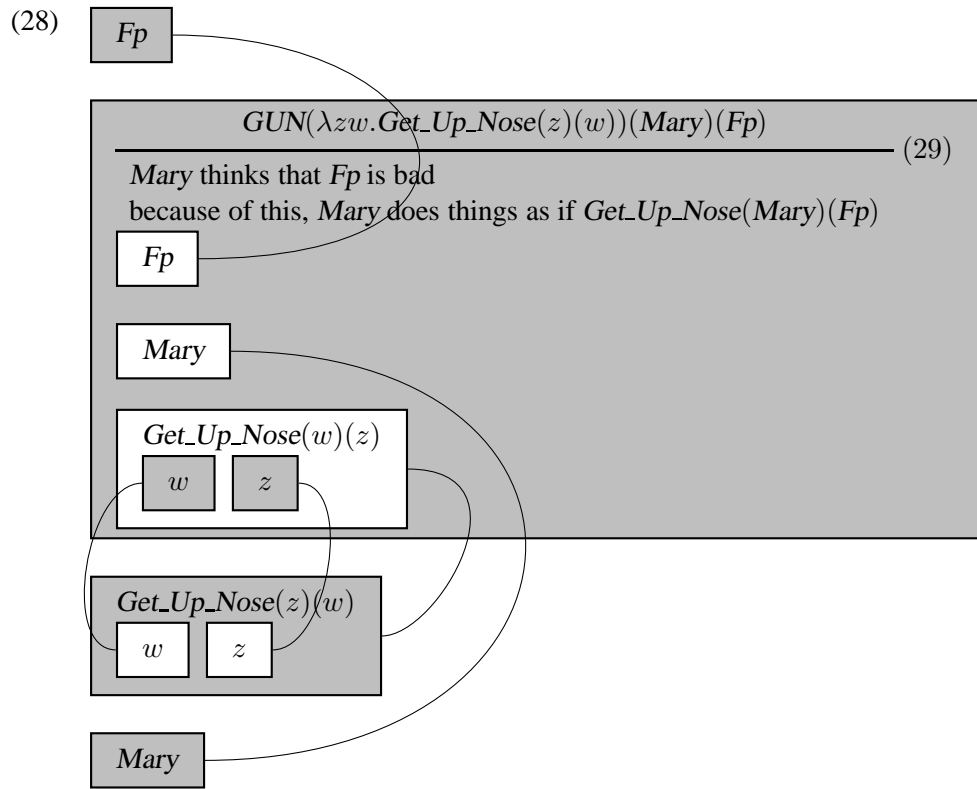


The axiom links as given are the only ones that yield a well-formed assembly.

A first step towards putting the meaning in a more conventional format is to provide a semi-formal definition of *GUN*, a combination of NSM and lambda-calculus as suggested in Andrews (2006):

$$(27) \text{ GUN} =_{def} \lambda P y x. y \text{ thinks that } x \text{ is bad} \\ \text{because of this, } y \text{ does things as if } P(y)(x)$$

Now we can derive the delivered meaning in a more conventional format by performing the indicated substitutions in (28), with a lambda-calculus derivation in (29):



The lambda-calculus derivation is:

$$\begin{array}{c}
(29) \quad \frac{GUN(\lambda zw. Get\_Up\_Nose(z)(w))(Mary)(Fp)}{[\lambda Pyx. y \text{ thinks that } x \text{ is bad} \\ \text{because of this, } y \text{ does things as if } P(y)(x)] \\ (\lambda zw. Get\_Up\_Nose(z)(w))(Mary)(Fp)} \text{ def } GUN \\
\hline
[\lambda yx. y \text{ thinks that } x \text{ is bad} \\ \text{because of this, } y \text{ does things as if } (\lambda zw. Get\_Up\_Nose(z)(w))(y)(x)] \\ Mary(Fp) \quad \beta \\
\hline
[\lambda yx. y \text{ thinks that } x \text{ is bad} \\ \text{because of this, } y \text{ does things as if } (\lambda w. Get\_Up\_Nose(y)(w))(x)] \\ (Mary)(Fp) \quad \beta \\
\hline
[\lambda yx. y \text{ thinks that } x \text{ is bad} \\ \text{because of this, } y \text{ does things as if } Get\_Up\_Nose(y)(x)] \\ (Mary)(Fp) \quad \beta \\
\hline
[\lambda x. Mary \text{ thinks that } x \text{ is bad} \\ \text{because of this, } Mary \text{ does things as if } Get\_Up\_Nose(Mary)(x)] \\ (Fp) \quad \beta \\
\hline
Mary \text{ thinks that } Fp \text{ is bad} \\ \text{because of this, } Mary \text{ does things as if } Get\_Up\_Nose(Mary)(Fp) \quad \beta
\end{array}$$

It would be a reasonable complaint from a cognitive point of view that a lambda-calculus derivation as complex as this one seems unlikely to have much in the way of psychological reality, and perhaps someone will find a simpler way to get the desired meaning to emerge from the parts in a formally. But meanwhile, lambda-calculus provides a precise and very well understood means to describe what needs to be done.

So now we have an account of how an idiom can use previously assembled meanings of its parts. A limitation of a pure glue semantic analysis is that all we can do is use the used assembly as is, or, possibly, just discard it. That is because glue says nothing about the facilities of the language in which meanings are expressed, other than that it can be manipulated by the lambda calculus. This however does out the presence of further facilities that might dissect meanings, using some parts and discarding others, but says nothing about them. Theories of meaning representation, such as those of Jackendoff or Wierzbica, might be able to provide facilities to do this.

There are an interesting range of further *nose* idioms which would certainly be worth looking through, for example:

- (30) a. Griselda turned up her nose at Alfonse's quiche.  
b. The botched roof repair got Elmer's nose out of joint.



- c. Janice kept sticking her nose into Brett’s romantic life.
- d. John told the investigators to follow their noses.

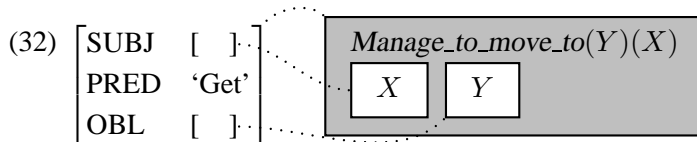
However we will not look at these here, but say more about the basis of the literal meaning of the *get [up NP]* construction that we are using in the construction of the idiomatic reading.

The idiom involves the use of *get* as a motion verb, which it seems able to function as in combination with quite a number of combinations, but not all:

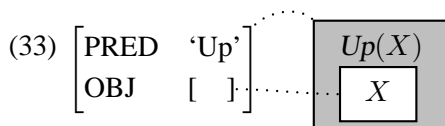
- (31) a. John got up/down/past the hill  
 b. John get into/out of the gully  
 c. \*John got along the road  
 d. John got to/\*from/\*by the town

The acceptable examples mostly appear to involve successful motion to a specific endpoint, under circumstances where success is not assured. In particular, we can perhaps explain the badness with *along*, *from*, and *by* by the absence of an implied endpoint (although the difference between *by* and *past* could use additional discussion). An exception to this however is ‘out of’: *John got out of the room* (as opposed to into it) does not seem to imply that there might have been any issue in managing this, but rather if he didn’t do it, something bad was likely to happen to him.

I there propose that *get* with a prepositional complement can be a motion verb, with the semantic conditions presented above, and an SLE that looks like this:



Given this much, we also need a literal meaning for *up*. We won’t try to analyse it, but just name it as *Up*, and propose the following SLE:



Given these constructors and one for *nose* in the environment of (25), the literal meaning might construct as  $Manage\_to\_move\_to(Up(Nose(y)))$ , which seems reasonable as a first pass.

However, it should turn out that our needed meaning for *get up NP* should also be treated as an idiom (as *get out of NP* has to be, according to what we

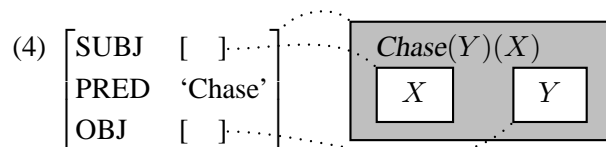
have observed above), that would not be a problem: there is no reason why a given word can't be involved in the invocation of 3 meaning contributions instead of just 1 or 2, as long as they all fit together, and none of them are applying to the same constellation features twice.

We will now take on the question of how to treat 'voice' alternations such as the passives in our approach.

## 4 The Passive

For some time, the dominant theory of voice alternations in LFG has been the Lexical Mapping Theory (LMT), in various versions, surveyed by Findlay et al. (2023). I believe that it should be possible to apply LMT to the present treatment of idioms, but it proves to be rather complex, and involve setting up an additional level of structure, 'semantic structure', along lines set out by Asudeh et al. (2013), requiring a considerable amount of additional explanatory material to be added to this paper. I will therefore take a different approach, which can be seen as a variant of the treatment in Findlay (2019).

This is a natural move, since the present proposal resembles Findlay's in associating lexical items with descriptions of expanses of structure that can be of arbitrary size, the lexical trees of Tree-Adjoining Grammar for Findlay, SLEs here. Consider the SLE for *chase*, repeated here for convenience:

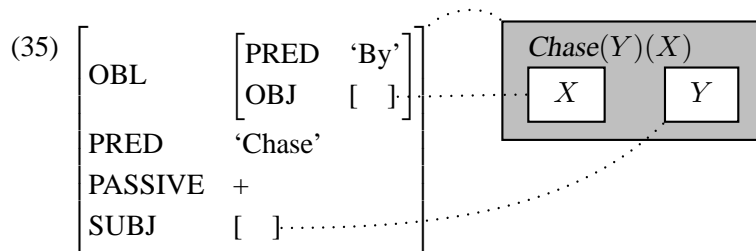


From this, we can derive a passive SLE for the PRED-feature 'Chase' the following steps:

- (34) a. Change the SUBJ grammatical function to OBJ of OBL with the preposition *by*.
- b. change the OBJ grammatical function to SUBJ.
- c. Add a constraint that the feature VOICE must have the value +.

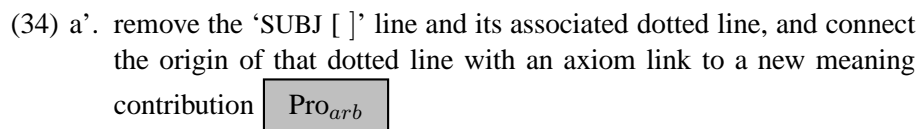
Note that a basic SLE can block the application of passive by imposing a constraint that the passive feature have no value.

The result of these changes on (4) is:

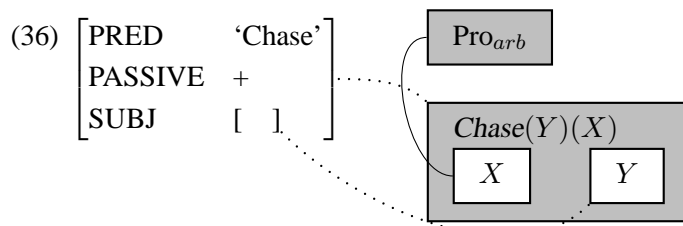


If one wanted to add some additional meaning contributions saying something along the lines of ‘this is saying something about SUBJ’, that would be possible, but I will not pursue this idea here.

There is also of course the agentless passive, as in *Fido was chased*, which can be formulated with the variant a’ of step a:

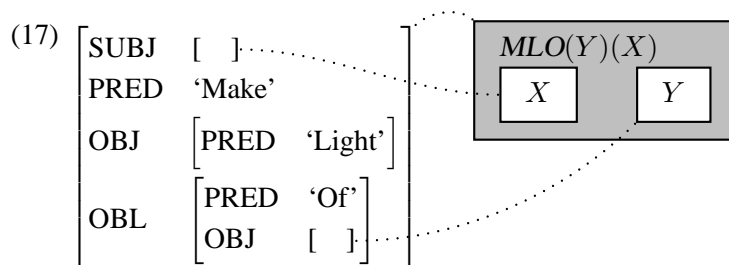


This yields the following result:



This is an oversimplification of the interpretation of unexpressed agents of passives; Collins (2024) argues that there are a number of distinct elements representing this agent. Dealing with these arguments in LFG would constitute a considerable challenge, which I will not undertake here.

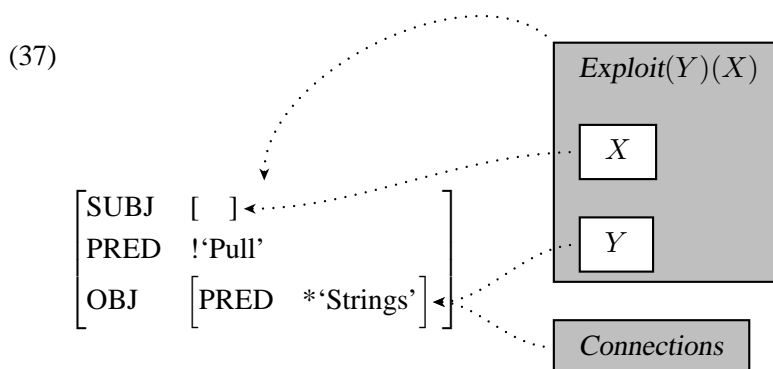
The next topic to consider is how this formulation will apply an idiom, such as for example *make light of*, whose SLE (17) is repeated here for convenience:



There is clearly nothing in the passive rule (34) that would prevent it from applying here, but there are two issues, one concerning exactly how, and the other, whether it actually should apply to (17) as formulated.

The how issue is how the grammatical functions to be changed are located, given that the SLE is introduced by the OBJ-value. The answer is that we can locate them with respect to the f-structure (description) that the final output of the meaning-constructor is linked to.

The ‘whether’ issue arises from work summarized and extended in Nunberg et al. (1994), with considerable subsequent followup, which shows that there is a very strong tendency for the passivizability of an idiom to depend on it being possible to provide metaphorical interpretations for the nominal and verbal components. A classic example would be *pull strings*, where *pull* seems metaphorically interpretable as ‘exploit/use’, and *strings* as ‘(personal connections)’. We can include independent meanings for the components of idioms like this:



The idea that these meanings would be motivated by the metaphor, and the link from the clear box labelled *Y* to the OBJ-value would allow Passive to apply.

On the other hand, for idioms such as *shoot the breeze* or *kick the bucket*, it is hard to see what a metaphorical (or any other kind of) interpretation for the object would be, so no link, and therefore Passive cannot apply. Although appealing, the proposal that passivization of idioms is controlled by decomposability in such a way has been challenged, for example in Wierzba et al. (2023), and we will not take a position on this issue here. But we do make two relevant observations:

- (38) a. In principle, a language could specify whether or not a link to the OBJ was required in order for Passive to apply.
- b. Individual speakers could vary in whether or not an idiom was decomposable, by having slightly different SLEs, or could even have two or more, varying in the degree to which they were preferred.

It might also be possible for people to construct metaphorical readings on the fly to understand idiomatic passives that they don’t normally use.

## 5 Double and Prepositional Passives

We now come to a substantial challenge, how to manage the ‘double passives’ of certain idioms, discussed in Bresnan (1982b, 60-62), but not, to my knowledge, in the Lexical Mapping Theory literature. An example from Bresnan is:

- (39) a. Noone took advantage of her talents.  
b. Not much advantage was taken of her talents.  
c. Her talents were not taken advantage of.

(a) is the active, (b) is unproblematic on the basis of what we have done so far, while (c) offers two challenges:

- (40) a. If *advantage* is the apparent object, why isn’t it being passivized?  
b. How is the ‘prepositional passive’, a typologically unusual feature of English, to be implemented?

The first question can be answered at least from a descriptive/implementational point of view by claiming that *advantage* here is not actually an OBJ, but something else. This idea is supported by the fact that it seems rather resistant to modification:

- (41) ??Her talents were taken very little/almost no/hardly any advantage of

All of these quantity modifiers are fine in the (39b) version, and also in the active (for example, if we replace *noone* with *they* in (a)). It therefore seems plausible that the idiomatic noun in (c) is an N, rather than an NP, bearing some grammatical relation other than OBJ to the verb.

More challenging is dealing with the apparent ‘preposition stranding’ in (39c). This is a characteristic of the prepositional passive of English (also called the pseudo-passive), discussed in LFG by Bresnan (1982b, 51-59). Bresnan’s analysis basically works, but is not entirely compatible with either the present proposal or recent LFG practice, requiring some discussion. The lexical rule is stated as follows:<sup>13</sup>

- (42) V – P *Incorporation*:  
Operation on Lexical Form: (P OBJ)  $\mapsto$  (OBJ)  
Morphological Change: V  $\mapsto$  V P

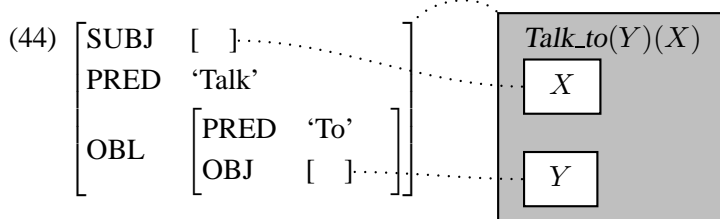
---

<sup>13</sup>The original rule writes the morphological change with a double-headed arrow, but I assume that this is a typo.

To render this into the present approach, it will be useful to start with a simple, non-idiomatic example, such as:

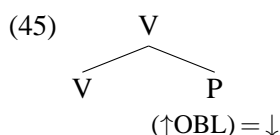
(43) John has been talked to

An SLE for this use of *talk* can be written as:



It is perhaps worth discussing why we might want a 2 place predicate here, rather than a 1-place predicate with an adjunct, since *talk* does not require a *to*-PP. The reason is that 'talking to' somebody implies that they can and probably do make an active contribution to the conversation; in fact, we can change the preposition to *at* in order to signal that the speaker is not making any effort to give them an opportunity to encourage them to, and is indeed probably discouraging them.

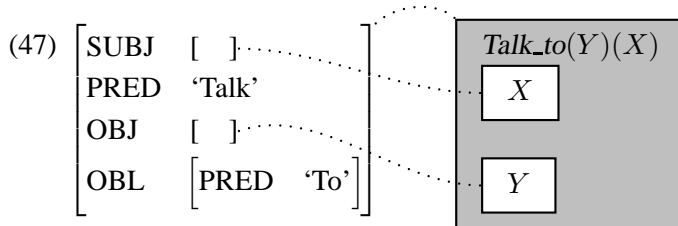
So now, as a feed-in for the prepositional passive, I propose an SLE extension that effectively incorporates the preposition into a verb, and reanalyses its object as an object of the verb (this is of course very similar to the treatment of applicatives proposed by Baker (1988), although with a rather different from of implementation). One component will be a V-expansion rule that introduces a P as the OBL of its clause:



This is a c-structure (or possibly morphological) rule that sets the stage for the SLE modification. What this does is very close to what Bresnan originally proposed, and should be regarded as a different implementation of the same idea:

(46) OBL OBJ  $\mapsto$  OBJ

The effect of this is to transform (44) into:



This is now set up for the Passive rule to apply.

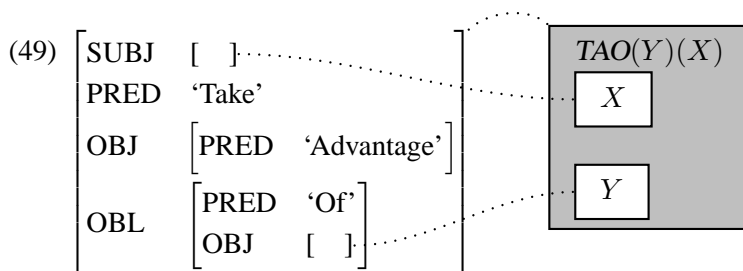
This analysis can also be described as reimplementing in LFG+Glue of the treatment proposed in Baker (1988, 259-263), who also provides an extensive discussion of typological phenomena along with factors that promote or inhibit it.

All three versions of the analysis lead to a prediction that the ‘complex NP shift’ phenomenon discussed in Ross (1967) should be able to apply to apparent objects of prepositions as long as these are adjacent to the verb. The contrast between (a) and (b) below suggests that this is correct to at least some degree:

- (48) a. ?\* John talked to yesterday the billionaire who is funding his project.  
 b. \*\* John explained the problem to yesterday the billionaire who is funding his project.

If the decision is to class (a) as ungrammatical, this could be achieved by adding to (46) a restriction allowing this to apply only when the passive morphology is present, while the intelligibility of (a) could be explained by allowing hearers to ignore this restriction if necessary to get an interpretation. The problem with (b) is considerably worse, and can’t be overcome just by ignoring a restriction on an SLE.

We are now ready to present an analysis of the (39c) passive, called the ‘outer passive’ in Nunberg et al. (1994). Consider the following basic SLE for *take advantage of*, not attempting to decompose the meaning to get a separate contribution for *advantage*:



Applying (46) to this is clearly either not going to be possible at all, or will lead to an unwanted result, but this can be fixed by two moves:

- (50) a. Add an optional nominal with some GF, for which we will use INC, to (45), between the V and the P  
 b. Have an SLE transformation rule changing OBJ to INC

## 6 Conclusion

I have shown how the version of the syntax-semantics interface developed in Andrews (2007, 2008, 2019) can deal with a number of problems in the analysis of idioms, including a small change to accommodate ones in which the literal meaning appears to be used in the construction of the idiomatic meaning. The account has a substantial resemblance to Construction Grammar, as shown by its ability to use an extended version of a popular Construction Grammar notation, but preserves the extensively developed and computationally implemented framework of LFG syntax.

An important feature of the analysis is its heavy reliance on the constraints on assembly imposed by linear logic itself, expressed as the proof-net rules, so that the syntax can be seen as serving in an advisory capacity to constrain the possible reasons. Indeed, Andrews (2007) proposes to use glue as a ‘generative’ level, in the form of the input to Optimality Theoretic LFG syntax.

Relevant to the approach are the results of Markantonatou & Samaridi (2018), who find that a variety of idioms Modern Greek do not undergo Passive, but can be reordered in various ways in the highly flexible ordering system in Modern Greek.<sup>14</sup> This can be easily handled in the present system, if Passive involves changes in f-structure, but numerous other reorderings don’t.

This general approach can be applied to at least some versions of Minimalism, as shown in Andrews (2024), where it is applied to the Hopf-algebraic formulation of (some aspects of) recent Minimalism launched in Marcolli et al. (2023) and more fully developed in Marcolli et al. (2024). Andrews (2024) also uses the box notation, and contains an appendix explaining its relationship to the standard notation for proof nets.

## References

- Andrews, Avery D. 2006. Semantic composition for NSM, using LFG + Glue. In Keith Allan (ed.), *Selected papers from the 2005 conference of the Australian Linguistic Society*, <https://als.asn.au/Conference/Proceedings/ALS200502> Feb 2025.
- Andrews, Avery D. 2007. Generating the input in OT-LFG. In J. Grimshaw, J. Maling, C. Manning & A. Zaenen (eds.), *Architectures, rules, and preferences: A festschrift for Joan Bresnan*, 319–340. Stanford CA: CSLI Publications. URL: <http://AveryAndrews.net/Papers> (Recent Publications (Refer-  
eed)).

---

<sup>14</sup>Which has the results is that it is extremely difficult to motivate any word order as ‘basic’ or underlying (Kechagias, 2011).



- Andrews, Avery D. 2008. The role of PRED in LFG+glue. In Miriam Butt & Tracy Holloway King (eds.), *The Proceedings of LFG08*, 46–76. Stanford, CA: CSLI Publications. <http://csli-publications.stanford.edu/LFG/13/lfg08.html> (accessed 19 Feb 2010).
- Andrews, Avery D. 2019. A one-level analysis for Icelandic Quirky Case. In Miriam Butt & Tracy Holloway King (eds.), *The Proceedings LFG19*, 24–27. Stanford, CA: CSLI Publications.
- Andrews, Avery D. 2024. Glue semantics for Hopf-algebraic Minimalism. <https://ling.auf.net/lingbuzz/007871>.
- Asudeh, Ash. 2023. Glue semantics. In Mary Dalrymple (ed.), *Handbook of Lexical-Functional Grammar*, 651–697. Berlin: Language Science Press.
- Asudeh, Ash, Mary Dalrymple & Ida Toivonen. 2013. Constructions with lexical integrity. *Journal of Language Modelling* 1. 1–51.
- Baker, Mark C. 1988. *Incorporation*. University of Chicago press.
- Belayev, Oleg. 2023. Introduction to LFG. In Mary Dalrymple (ed.), *Handbook of lexical functional grammar*, Berlin: Language Science Press. <https://langsci-press.org/catalog/book/312>.
- Bresnan, Joan Wanda (ed.). 1982a. *The mental representation of grammatical relations*. Cambridge MA: MIT Press.
- Bresnan, Joan Wanda. 1982b. The Passive in Lexical Theory. In Joan Bresnan (ed.), *The mental representation of grammatical relations*, 3–86. MIT Press.
- Casadio, Claudia. 1988. Semantic categories and the development of categorial grammar. In Emmon Bach R.T. Oehrle & Deirdre Wheeler (eds.), *Categorial grammar and natural language semantics*, 95–123. Reidel.
- Collins, Chris. 2024. *Principles of argument structure*. Cambridge, MA: MIT Press.
- Dalrymple, Mary (ed.). 1999. *Syntax and semantics in Lexical Functional Grammar: The resource-logic approach*. Cambridge, MA: MIT Press.
- Dalrymple, Mary. 2001. *Lexical Functional Grammar*. Academic Press.
- Dalrymple, Mary, Ronald M. Kaplan, John T. Maxwell & Annie Zaenen (eds.). 1995. *Formal issues in Lexical-Functional Grammar*. Stanford, CA: CSLI Publications.
- Egan, Andy. 2008. Pretense for the complete idiom. *Noûs* 42. 381–409. Original ms. 2004.
- Findlay, Jamie. 2019. *Multiword expressions and the lexicon*: Oxford University dissertation.
- Findlay, Jamie, Roxanne Taylor & Anna Kibort. 2023. Argument structure and mapping theory. In Mary Dalrymple (ed.), *Handbook of lexical-functional grammar*, Langsci Press.
- Halvorsen, Per-Kristian. 1983. *Semantics for Lexical-Functional Grammar*.

- Linguistic Inquiry* 14. 567–615.
- Halvorsen, Per-Kristjan & Ronald M. Kaplan. 1988. Projections and semantic description in Lexical-Functional Grammar. In *Proceedings of the International Conference on Fifth Generation Computer Systems*, 1116–1122. Tokyo: Institute for New Generation Computer Technology. Also published in Dalrymple et al. (1995), pp 279-292.
- Kay, Paul & Charles J. Fillmore. 1999. Grammatical constructions and linguistic generalizations: The what’s X doing Y? construction. *Language* 75. 1–33.
- Kechagias, Axiotis. 2011. *Regulating word order in modern greek: Verb initial and non-verb initial orders & the conceptual- intentional interface*. London: University College London dissertation.
- Marantz, Alec. 1984. *On the nature of grammatical relations*. Cambridge MA: MIT Press.
- Marcolli, Matilde, Noam Chomsky & Robert Berwick. 2023. Mathematical structure of syntactic merge. <https://arxiv.org/abs/2305.18278>.
- Marcolli, Matilde, Noam Chomsky & Robert Berwick. 2024. Mathematical structure of syntactic merge. [https://www.its.caltech.edu/~matilde/MergeMCB-MITPress-LI.pdf?fbclid=IwY2xjawE8jOlleHRuA2FlB...aem\\_Bq8ceyDiIQPK91tf2FCKvw](https://www.its.caltech.edu/~matilde/MergeMCB-MITPress-LI.pdf?fbclid=IwY2xjawE8jOlleHRuA2FlB...aem_Bq8ceyDiIQPK91tf2FCKvw).
- Markantonatou, Stella & Niki Samaridi. 2018. Revisiting the grammatical function “ $\hat{I}$ bject” (obj and obj $_{\theta}$ ). In Manfred Sailer & Stella Markantonatou (ed.), *Multiword expressions: Insights from a multi-lingual perspective*, 187–213. Berlin: Language Science Press.
- Moot, Richard. 2002. *Proof-nets for linguistic analysis*: University of Utecht dissertation. Retrieved Sept 9, 2024, from <http://www.labri.fr/perso/moot/>.
- Nunberg, Jeff, Ivan Sag & Thomas Wasow. 1994. Idioms. *Language* 70. 491–538.
- Partee, Barbara H. 2006. Do we need two basic types. In Hans-Martin Gaertner, Regine Eckardt, Renate Musan & Barbara Stiebels (eds.), *Puzzles for Manfred Krifka*, Berlin. URL: [www.zas.gwz-berlin.de/40-60-puzzles-for-krifka/](http://www.zas.gwz-berlin.de/40-60-puzzles-for-krifka/).
- Partee, Barbara H. & Vladimir Borshev. 2004. Genitives, types, and sorts. In Ji yung Kim, Yury A. Lander & Barbara H. Partee (eds.), *Possessives and beyond: Semantics and syntax*, 29–43. Amherst MA: UMass GSLA. URL: <http://people.umass.edu/partee/Research.htm>.
- Ross, John Robert. 1967. *Constraints on variables in syntax*. Cambridge MA: MIT dissertation. Distributed by the Indiana University Linguistics Club; revised version published as Ross (1986).
- Ross, John Robert. 1986. *Infinite syntax!* NJ: Norwood. Revision of Ross (1967).
- Wierzba, Marta, J.M.M. Brown & Gisbert Fanselow. 2023. Sources

of variability in the syntactic flexibility of idioms. *Glossa* 8(1).  
doi:<https://doi.org/10.16995/glossa.8502>.