"Polymorphic" Causatives: Complex Predicates in French

Annie Zaenen
Rank Xerox Research Centre, Meylan, France
annie.zaenen@xerox.fr

Mary Dalrymple
Xerox PARC, Palo Alto, CA USA
dalrymple@parc.xerox.com

January 17, 1995

1 Introduction

The French causative construction is a complex predicate construction, in which multiple predicative elements combine to form a single syntactic predicate. In example (1), the two verbs faire 'cause' and venir 'come' combine to give rise to a syntactically monoclusal structure with a complex semantic representation:

(1) Pierre a fait venir Paul.

In this case, it is the intransitive verb venir 'come' which combines with causative faire to form a new two-place syntactic predicate. In fact, causative faire can combine with a verb of any arity to form a new syntactic predicate; the resulting predicate has as many arguments as the original verb, with the addition of a new argument for the causer. We propose an approach to syntactically-derived complex predicates which produces the correct syntactic structure and correctly specifies the relation between the syntax and the meaning without relying on an ambiguity in the causative verb to permit combination with verbs of varying arities.

2 The nature of complex predicates

Much work has been devoted to the analysis of complex predicate constructions (Manning 1992, Alsina 1993a, Butt 1993, and works cited therein, among many others), and in particular to the syntactic features of these constructions. The analysis presented here for the French causative construction closely resembles the analysis proposed for Urdu complex predicates by Dalrymple et al. (1993), employing a resource logic, linear logic, to guide the composition of meanings. The approach gives rise to a clean characterization of completeness and coherence (Dalrymple, Lamping, and Saraswat, 1993), and meshes well with the treatment of other semantic phenomena, such as quantification, modification, and intensional verbs (Dalrymple et al., 1994a; Dalrymple et al., 1994b).

2.1 Permissives: Syntactically biclausal

The French permissive construction is exemplified in (2):

(2) Pierre a laissé Paul venir
    let come
    'Pierre let Paul come.'

French permissives may be syntactically biclausal, in contrast with the normally monoclausal causative construction. The f-structure for (2) is given in (3); the complement clause headed by venir 'come' bears the XCOMP relation to the main clause verb laisser 'let'.

(3) [PRED LAISSER
    SUBJ [PRED PIERRE]
    OBJ [PRED PAUL]
    XCOMP [PRED VENIR]
]

This stands in contrast to the situation with causatives, which form complex predicates and are monoclausal.

2.2 Causatives: Complex predicates

The French causative construction is exemplified in (4), with its f-structure given in (5):

(4) Pierre a fait venir Paul
    cause come
    'Pierre caused Paul to come.'

(5) [PRED FAIRE(VENIR)
    SUBJ [PRED PIERRE]
    OBJ [PRED PAUL]
]

Evidence for the differing status of these constructions comes from clitic climbing. A clitic object appears adjacent to its verb. For biclausal constructions, the clitic object of the subordinate clause verb appears next to the subordinate verb:

Due to lack of space, we will not discuss the monoclausal permissive construction exemplified in (a):

(a) Pierre a laissé venir Paul.
(6) Marie veut [la manger]
    wants it to eat
    'Marie wants to eat it.'

However, in the case of complex predicates the clitic object appears adjacent to the complete verb complex of which it is an object:

(7) Pierre l'a fait venir
    him-has cause come
    'Pierre has caused him to come.'

(8) Marie la lui a fait manger
    it him has cause eat
    'Marie has caused him to eat it.'

Much other evidence has been presented that differentiates these two constructions and evidences the monoclausality of the complex predicate construction; see, for example, Abeille and Godard (1994) and the references cited there.

3 Semantic composition as deduction in linear logic

Dalrymple, Lamping, and Saraswat (1993) present an approach to semantic composition within the LFG framework which makes use of a fragment of linear logic\textsuperscript{2} (Girard, 1987) to state constraints on the composition of meanings; see also Dalrymple et al. (1993), Dalrymple et al. (1994a), and Dalrymple et al. (1994b). We briefly review the assumptions of this framework here.

A key principle of the approach is to distinguish meanings from instructions about how to assemble meanings. The language of meaning can be any suitable logical language for expressing truth-conditional meanings of words and phrases; for current purposes, first-order logic will suffice. The language of assembling meanings or glue language is a fragment of linear logic (the tensor fragment) used to express constraints on the composition of meanings. That is, the glue language expresses how the meanings of words and phrases can combine.

We will go through a simple example before turning to the more complicated case of complex predicates. The relevant lexical entries for sentence (9) are given in (10):

(9) Marie a mangé la pomme
    ate the apple

(10) Marie NP (↑ PRED) = MARIE
    ↑ ⊸ Marie
    pomme N (↑ PRED) = POMME
    ↑ ⊸ apple
    manger V (↑ PRED) = MANGER
    \( \forall X, Y. (↑_\sigma AGENT) \sim X \otimes (↑_\sigma THEME) \sim Y \)
    \( \to ↑_\sigma \sim \text{eat}(X, Y) \)

Since we present this simple example merely in order to illustrate our formal assumptions, we ignore the syntax and semantics of the other words in the sentence, as well as issues of tense, aspect, definiteness, and so on.

Our analysis requires a means of specifying the meanings of particular syntactic objects and the relations between these meanings. We take advantage of the projection architecture of LFG (Kaplan, 1987; Halvorsen and Kaplan, 1988), which enables a characterization of meanings associated with f-structures. Notationally, the subscript \( \sigma \) represents the projection function from f-structures to semantic representations. The expression \( ↑_\sigma \) represents the semantic projection of the f-structure \( ↑ \), for example. We assume that semantic projections are associated via the \( \sim \) relation with meanings, and so the expression \( ↑_\sigma \sim \text{Marie} \) indicates that the meaning \text{Marie} is recorded as the meaning of \( ↑_\sigma \). Semantic projections can also be structured with attributes and values, just like f-structures, and the values of these attributes can be associated with meanings. Thus the expression \( (↑_\sigma AGENT) \sim X \) means that the semantic projection \( ↑_\sigma \) of the f-structure \( ↑ \) has an attribute AGENT, and that \( X \) is recorded as the meaning of the value of \( ↑_\sigma AGENT \).

The lexical entries in (10) give rise to the f-structure and associated semantic information in (11):

(11) \[
\begin{array}{l}
\text{PRED MANGER} \\
\text{f: [SUBJ g: [PRED MARIE] ]} \\
\text{OBJ h: [PRED POMME] ]} \\
\text{g_\sigma \sim Marie} \\
\text{h_\sigma \sim apple} \\
\forall X, Y. (↑_\sigma AGENT) \sim X \\
\otimes (↑_\sigma THEME) \sim Y \\
\to f_\sigma \sim \text{eat}(X, Y) \\
\end{array}
\]

Lexical items like \text{Marie} and \text{pomme} have meanings like \text{Marie} and \text{apple}. Verbs are more complicated. The verb \text{manger} makes the following semantic contribution:

\( \forall X, Y. (↑_\sigma AGENT) \sim X \)
\( \otimes (↑_\sigma THEME) \sim Y \)
\( \to f_\sigma \sim \text{eat}(X, Y) \)

This rule states that if the agent of \text{manger} has meaning \( X \), and (\( \otimes \)) the theme of \text{manger} has meaning \( Y \), then (\( \to \)) those meanings are consumed, and a meaning for the f-structure for the sentence, \( f \), is produced. For further details, see Dalrymple, Lamping, and Saraswat (1993).

Note that the verb \text{manger} 'eat' does not subcategorize for grammatical functions such as SUBJ and OBJ, as is assumed in classical LFG analyses. Instead, it specifies semantic/thematic information about its arguments. The grammatical functions that express these arguments are given by mapping rules, which specify a relation between arrays of thematic roles and the grammatical functions that realize them. For the example at hand, the following mapping rule is relevant:

(12) \( !(\forall f, X, Y. ((f \text{SUBJ})_\sigma \sim X) \otimes (f \text{OBJ})_\sigma \sim Y) \)
\( \to (f_\sigma AGENT) \sim X \otimes (f_\sigma THEME) \sim Y \)
This mapping rule states a general relation between subjects and agents, and objects and themes: given a subject \( X \) and an object \( Y \), an agent \( X \) and a theme \( Y \) can be posited.

The mapping rule in (12) is a general rule about possible thematic role/grammatical function associations for f-structures in general. We can instantiate this mapping rule for the particular f-structure that we are concerned with, the f-structure labeled \( f \):

\[
(13) \forall X, Y. (g_o \sim X) \otimes (h_o \sim Y) \\
\quad \rightarrow (f_o \text{AGENT}) \sim X \otimes (f_o \text{THEME}) \sim Y
\]

Having done this, we begin the semantic deduction with the following premises:

- Marie : \( g_o \sim \text{Marie} \)
- apple : \( h_o \sim \text{apple} \)
- mapping1 : \( \forall X, Y. (g_o \sim X) \otimes (h_o \sim Y) \\
\quad \rightarrow (f_o \text{AGENT}) \sim X \otimes (f_o \text{THEME}) \sim Y \)
- eat : \( \forall X, Y. (f_o \text{AGENT}) \sim X \otimes (f_o \text{THEME}) \sim Y \\
\quad \rightarrow f_o \sim \text{eat}(X, Y) \)

From these premises, the following deduction is possible, giving the desired result:

- \( \text{Marie} \otimes \text{apple} \otimes \text{mapping1} \otimes \text{eat} \)
- \( \vdash (f_o \text{AGENT}) \sim \text{Marie} \\
\quad \otimes (f_o \text{THEME}) \sim \text{apple} \\
\quad \otimes \text{eat} \)
- \( \vdash f_o \sim \text{eat}(\text{Marie}, \text{apple}) \)

The use of linear logic instead of classical logic as the 'glue language' provides clear advantages which we have discussed in other work (Dalrymple, Lamping, and Saraswat, 1993; Dalrymple et al., 1993): it allows us to capture the intuition that lexical items and phrases contribute uniquely to the meaning of a sentence, thus enabling a clean semantic definition of the LFG requirements of completeness and coherence. This requirement is what ensures the selection of the correct mapping rule. In a semantic derivation, the wrong mapping rule might be incorrectly chosen, for example one which relates experiencers to subjects, and themes to objects:

\[
(14) !((\forall f, X, Y. ((f \text{OBJ})_o \sim X) \otimes ((f \text{OBJ})_o \sim Y) \\
\quad \rightarrow (f_o \text{EXPERIENCER}) \sim X \\
\quad \otimes (f_o \text{THEME}) \sim Y)
\]

The use of this mapping rule in the derivation of the sentence above would not allow a well-formed derivation. First, the meaning for a clause containing the verb manger 'eat' can only be obtained in the presence of an agent, and this mapping rule does not provide one. Second, this rule asserts the presence of an experiencer which is not consumed during the derivation; the result is incoherence, since the premises of the sentence do not lead to a single meaning associated with the sentence, with no premises remaining.

Alternatively, a mapping rule such as the following might have been incorrectly selected:

\[
(15) !(\forall f, X, Y. ((f \text{OBL-agent})_o \sim X) \\
\quad \otimes ((f \text{SUBJ})_o \sim Y)) \\
\quad \rightarrow (f_o \text{AGENT}) \sim X \otimes (f_o \text{THEME}) \sim Y
\]

This rule relates themes to subjects, and requires the agent to be realized as an oblique phrase. Again, a well-formed derivation would not be possible, since the rule allows the derivation to proceed only in the presence of an OBL-agent phrase. In this and in the above case, the interaction of the mapping rules with completeness and coherence constraints ensures that no other derivation produces a well-formed output.

It is important to note that we intend no claims about the correctness of the specific details of these mapping rules; rather, our claim is that mapping rules should be of the general form we have illustrated, specifying possible relations between thematic roles and grammatical functions. In particular, no theoretical significance should be attached to the choice of thematic role labels used here; for the verb eat, for example, labels such as 'eater' and 'eaten-thing' would do as well. Mapping rules may specify thematic/semantic information as we have given it, as assertions about roles such as agent and theme. Alternatively, such information can be specified in terms of structures of Conceptual Semantics (Jackendoff, 1990), as in Butt's approach (Butt, 1993), in terms of Proto-Roles (Dowty, 1991), as in Alsina's approach (Alsina, 1993a), or by some other means. Grammatical function information can be specified in terms of grammatical function labels such as SUBJ and OBL, as is done here; alternatively, such information can be given in terms of a feature decomposition cross-classifying grammatical functions (Levin, 1986; Bresnan and Kanerva, 1989; Alsina, 1993a). Our claims relate to the formal properties of mapping rules, and thus our approach could be adopted as a part of any one of a number of specific approaches to mapping theory.

4 Complex predicates, mapping rules, and deduction

We are now ready to consider how our approach applies to complex predicates. Our task is to explain how causative faire can combine with a verb of any arity — here, with the intransitive verb venir 'come' — to license a syntactically monoclausal argument structure. We also want to ensure that the syntactic representation for this sentence is correctly related to its (complex) meaning. We will consider the following example:

\[
(16) \text{Pierre a fait \ venir Paul} \\
\text{cause come}
\]

'Pierre caused Paul to come.'

The first issue we must address is what the semantic structure of this sentence should look like. Semantically, is the French causative verb two-place or three-place?

\[
(17) \text{a. cause(Pierre, come(Paul))} \\
\text{b. cause(Pierre, Paul, come(Paul))}
\]

Our framework enables us to produce either meaning unproblematically; for discussion of the linguistic issues involved in
this choice, see Alsina and Joshi (1991). We lack space to address this issue in detail, and will make the assumption that the first alternative is the correct one.

The f-structure and lexical entries for sentence (16) are:

\[(18) \quad \begin{align*}
    \text{PRED} & \text{FAIRE(VENIR)} \\
    f : \quad \text{SUBJ} & \ r \text{PIERRE} \\
    \text{OBJ} & \ r \text{PAUL}
\end{align*}\]

(19) Pierre NP (↑ PRED) = PIERRE
Paul NP (↑ PRED) = PAUL

\[\begin{align*}
    \text{VENIR} & \ r \text{COME} \\
    \forall X. (↑ \text{AGENT}) & \ r X \rightarrow ↑ \text{come}(X) \\
    \text{FAIRE} & \ r X, P. (↑ \text{CAUSER}) \ r X \\
    \quad & \quad \quad \quad \theta (↑ \text{AGENT}) \ r P \\
    \quad & \quad \quad \quad \quad \rightarrow ↑ \text{cause}(X, P)
\end{align*}\]

The semantic contribution of causative faire is one of the two crucial ingredients in our analysis:

\[\forall X, P. (↑ \text{CAUSER}) \ r X \cap (↑ \text{AGENT}) \ r P \rightarrow ↑ \text{cause}(X, P)\]

Causative faire requires an argument with the role of CAUSER, the role filled by Pierre in the sentence above. We reiterate that the exact choice of thematic role label is not the important issue here; rather, our approach requires merely that the thematic/semantic information referred to by mapping rules and lexical entries be expressed clearly and unambiguously.

Besides a CAUSER, causative faire also requires another argument: a preliminary meaning for the f-structure that it is associated with. This meaning will be obtained when the verb venir is provided with its arguments. Causative faire requires, then, that the CAUSER X and the preliminary meaning P be consumed, and a new meaning produced: cause(X, P). This preliminary meaning may be provided by a verb of any arity; thus, we need only one lexical entry for causative faire, no matter what the arity of the verb with which it combines.

It is here that one of the advantages of the use of linear logic for guiding semantic composition can be clearly seen: classical logics would not allow a step such as this one, where two different assignments of meaning to ↑ are made on the left and right hand sides of the → linear implication; the resource orientation of linear logic makes such an analysis possible.

The second crucial ingredient in our analysis is the following mapping rule:

(20) \[(\forall f, X, Y, Z. ((f \text{SUBJ}) \ r X) \cap ((f \text{OBJ}) \ r Y) \rightarrow (f \text{CAUSER}) \ r X \cap (f \text{AGENT}) \ r Y))\]

This rule maps a SUBJ X and an OBJ Y to the thematic roles CAUSER and AGENT; crucially, this mapping takes place whether these roles are both associated with a single verb (the usual case) or with more than one verb (the case of complex predicates). In other words, we claim that in the case of complex predicates, the SUBJ and OBJ (and other grammatical functions) of a single f-structure can be associated with thematic roles of multiple verbs.

The semantic contributions from the lexical entries of the words in the sentence and the instantiated mapping rule give the following premises:

(21) Pierre : (g σ ∼ Pierre)
Paul : (h σ ∼ Paul)

\[\text{mapping2} : (\forall X, Y, Z. (g σ ∼ X) \cap (h σ ∼ Y) \rightarrow (f σ \text{CAUSER}) ∼ X \cap (f σ \text{AGENT}) ∼ Y)\]

\[\text{come} : (\forall X. (f σ \text{AGENT}) ∼ X \rightarrow f σ ∼ \text{come}(X))\]

\[\text{cause} : (\forall X, P. (f σ \text{CAUSER}) ∼ X \cap f σ ∼ P \rightarrow f σ ∼ \text{cause}(X, P))\]

From these premises, the deduction proceeds as follows:

(22) \[\begin{align*}
    \vdash (f σ \text{CAUSER}) ∼ Pierre \cap (f σ \text{AGENT}) ∼ Paul \\
    \quad \cap \text{come} \cap \text{cause} \\
    \vdash (f σ \text{CAUSER}) ∼ Pierre \\
    \quad \cap f σ ∼ \text{come}(Paul) \cap \text{cause} \\
    \vdash f σ ∼ \text{cause}(Pierre, \text{come}(Paul))
\end{align*}\]

5 Semantic forms and resource-accounting

Many syntactic frameworks implicitly or explicitly assume that the array of grammatical functions licensed by a syntactic predicate is immutable and cannot be altered in the course of a derivation. Complex predicates seem to violate this assumption, however; verbs such as causative faire seem to operate on syntactic predicates so as to increase their arity by one.

In the context of LFG, certain assumptions have traditionally been made about the nature and function of semantic forms, the values of the attribute PRED. Kaplan and Bresnan (1982) assume PRED values for verbs of the following form:

(23) \[\text{eat< (↑ SUBJ)} \quad (↑ OBJ) >\]

agent theme

According to Kaplan and Bresnan (1982), these semantic forms encode four types of information:

1. Specification of the semantic relation
2. Mapping of grammatical functions to semantic roles
3. Subcategorization information (the governed grammatical functions)
4. Instantiation to indicate semantic distinctness (predicate uniqueness)
Encoding these kinds of information by means of a single formal device permits the syntactically relevant aspects of meaning to be confined to a single place in the f-structure. But we have seen that subcategorization information must be allowed to change in the course of a syntactic derivation, and so some change is necessary to allow for the formation of complex predicates while still allowing for the above four kinds of information to be properly encoded.

Kaplan (personal communication) observes that the effect of our approach and the approach taken in much other recent LFG literature (Alsina 1993a, Butt 1993, among others) is to treat the different kinds of information encoded by semantic forms with separate and independent mechanisms. Specification of the semantic relation is accomplished in the lexicon: a verb like ‘eat’, for example, specifies that its meaning is $eat(X, Y)$ when given an agent $X$ and a theme $Y$. We assume that the mapping of grammatical functions to semantic roles, as well as the particular inventory of governed grammatical functions, are given by the mapping rules: the mapping rules might specify that for a verb like ‘eat’, the subject bears the thematic role of agent and the object bears the role of theme.

Given these new mechanisms, complex predicates and simpler constructions are handled identically. If a rule exists to map a set of grammatical functions licensed by a complex predicate to a set of thematic roles, then the result will be a monoclausal syntactic structure, whether this array of thematic roles arose in association with a single lexical item or with multiple predicative elements. This aspect of the analysis of complex predicates is thus unproblematic for any approach which assumes that verbs are lexically associated not with an array of grammatical functions, but with an array of thematic roles; mapping rules will associate the appropriate array of grammatical functions with these thematic roles in all cases, for complex predicates as well as for simpler cases.

On these assumptions, only one function of the semantic form remains: instantiation to indicate semantic distinctness. On the classical LFG view, different PRED values are incompatible; this disallows the presence of multiple fillers of a single syntactic argument slot. For example, clitic doubling is disallowed in some languages because the clitic pronoun and the full noun phrase contribute incompatible PRED values, as in the following example (Grimshaw, 1982):

(24) a. Je cherche Pierre
   I am looking for Pierre
   'I am looking for Pierre.'

b. Je le cherche
   I him am looking for
   'I am looking for him.'

c.*Je le cherche Pierre
   I him am looking for Pierre

Note, however, that in the case of complex predicates the same situation seems to arise. Each of the verbs that make up the complex predicate construction would be expected to contribute a PRED with a unique value.

We believe that the syntactic function of predicate uniqueness is an important one, and so we continue to permit semantic forms in the lexical entries of semantically contentful lexical items to specify the values of PRED attributes in the f-structure. However, we also believe in a necessary separation between syntax and semantics, so that which particular semantic form appears is irrelevant for syntactic purposes. That is, the syntax is concerned only with the presence or absence of a semantic form, and not with its internal structure. Thus, if the only remaining function of the PRED is to ensure predicate uniqueness, it would do as well to assume that the PRED value for a sentence with a complex predicate is contributed by the main verb (for example, VENIR), and that the function of FAIRE is to modify the argument structure but not to contribute to or change the PRED value of the construction. It is for this reason that we do not take a position on how a semantic form such as 'FAIRE<VENIR>' is formed, since we believe that the particular syntactic shape of this predicate is not an issue.

6 Conclusion

The use of linear logic enables a treatment of complex predicates which produces the correct syntactic structure and correctly specifies the relation between the syntax and the meaning without relying on an ambiguity in the causative verb to permit combination with verbs of varying arities. Our deduction framework enables us to use linear logic to state such operations in a formally well-defined manner.

7 Acknowledgments

We would like to thank John Lamping and Vijay Saraswat for valuable input and comments.

References

Abeille, Anne and Danielle Godard. 1994. French causatives in HPSG. MS, University of Paris VII.


