A Note on Complex Predicate Formation: Evidence from Auxiliary Selection, Reflexivization, and Past Participle Agreement in French and Italian

Anette Frank
Institute for Computational Linguistics
University of Stuttgart
Azenbergstr.12, D–70174 Stuttgart
Germany[1]

We propose a analysis of well–known contrastive data of complex predicate formation in French and Italian where we distinguish, with Rosen(1989), between two types of argument composition: *full* and *partial merger* of argument structure. Two alternative ways are investigated to integrate this analysis into the LFG framework: complex predicate formation in syntax or by lexical rule.

1 A puzzle in complex predicate formation

French and Italian confront us with an interesting puzzle in the analysis of complex predicates. In both languages we find constructions – involving e.g. causative verbs and so–called restructuring verbs (in Italian[2]) – for which various authors have motivated an analysis in terms of complex predicate formation.[3] The most compelling criterion in support of this analysis is (object) clitic climbing – obligatory with *faire/fare* (1a) and optional with restructuring verbs (1b) – and which we also find with auxiliaries in both of these languages (1c).[4] Another argument comes from causative constructions, which induce complex relation changes on the embedded verb (1a).[5]

(1) a. fr. Marie lui a fait réparer la voiture.
it. Maria gli a fatto riparare la macchina.
'Mary made him repair the car.'

b. fr. [no restructuring verbs]
it. Mario lo vuole leggere.
'Mary wants to read it.'

c. fr. Marie l'a mangé.
it. Maria l'ha mangiato.
'Mary has eaten it.'

Both of these phenomena are best explained, in an LFG analysis, in terms of a monoclausal *functional* structure, to be obtained via some sort of predicate composition,[6] while still positing, for all of (1), a rightward branching VP–embedding c–structure.[7]

[1]I am grateful for valuable comments by Miriam Butt, Christian Rohrer and Jürgen Wedekind – which does not imply that they necessarily subscribe to the view layed out in the following.

[2]There are no (intransitive) restructuring verbs in French: *vouloir, devoir*, etc. do not admit clitic climbing.

[3]See e.g. Abeillé et al(1996), Alsina(1993), Butt(1995a/b), Frank(1989/1990), Monachesi(1993), Manning(1992/1996), Rosen(1989), as well as earlier work in the framework of Relational Grammar.

[4]While auxiliaries are in general not considered as inducing complex predicate constructions, we will consider them as constituting a special class of restructuring verbs (in Italian). Manning(1996) takes a similar position.

[5]Another criterion for complex predicate formation is *tough*–movement (Manning(1996), Abeillé et al(1996)).

[6]See e.g. Alsina(1993). Alternative approaches have been proposed by Manning(1992/1996), Dalrymple et al(1992) and Kaplan/Wedekind(1993).

[7]See in particular the arguments in Manning(1992/1996).

Yet, once we look at further properties of complex predicate constructions, as e.g. auxiliary selection, passivization, past participle agreement and reflexivization, it turns out that argument composition somehow comes "in different degrees" in these two languages. This is not explained by current LFG analyses of complex predicates in Romance languages.[8]

Let us look at these phenomena in turn.

Auxiliary selection In both French and Italian the selection of auxiliary verbs is governed by the thematic *unaccusativity* property of (intransitive) verbs.[9] By and large, intransitive verbs whose argument is associated with the thematic role THEME select *être*/*essere*.

So, since the subject of *volere* is not a THEME, it selects *avere* (2a), while *andare* selects *essere*. Yet, complex predicate constructions involving restructuring verbs, such as (2b), induce a change in auxiliary selection, depending on the auxiliary selection on the embedded verb: in (2b) *essere* is selected, which is predicted by the embedded *andare*, but not by the matrix predicate *volere* (see e.g. Manning(1996)).

(2) a. it. Mario avrebbe proprio voluto andarci.
b. it. Mario ci sarebbe proprio voluto andare.
'Mario would have really wanted to go there.'

In French there are no corresponding intransitive verbs that allow for complex predicate formation. Nevertheless it can be shown, by recurring to auxiliary constructions (3), that French behaves differently in this respect: passivized *accuser* selects *être*, yet the (putatively) complex predicate formed by *être* and *accusée* does not in turn select *être*, which is to be expected if French were to follow the pattern of Italian restructuring verbs. The Italian complex auxiliary construction (3a), by contrast, is in accordance with the general behaviour of restructurig verbs: *stare accusare* selects an *essere* auxiliary, as imposed by the passivized *accusare*.

(3) a. it. Maria è stata accusata.
b. fr. Marie a été accusée.
'Mary has been accused.'

Passivization Italian complex predicate constructions allow for long passivization. They do so for some restructuring verbs (4a–b) and uniformly with causative verbs (5a)/(6a), where it is either the (logical) subject of the embedded predicate (5) or else its object (6) that can become the functional subject of the complex construction.

French causative constructions do not allow for long passivization, as shown by (5b) and (6b). Nor is this possible for non–causatives (*entendre*, etc.) in complex construction.

(4) a. it. Il palazzo fu cominicato a costruire sotto Carlo V.
'The palace was begun to be built under Charles V.'
b. it. ?(?) L'affitto fu continuato a pagare fino alla fine dell'anno.
'The rent was continued to be payed till the end of the year.'

[8] And it isn't by current HPSG analyses either (Abeillé et al(1996), Monachesi(1993)). Rosen(1989) offers a solution to the problem in a GB–framework. See below. See also Zubizarreta(1985) for an earlier GB–account resorting to coanalysis.

[9] In fact, auxiliary selection is not only subject to thematic unaccusativity (with various exceptions to be stated in the lexicon), but is also governed by morphosyntactic processes of argument reduction, such as reflexivization and passivization (see below).

(5) a. it. I prigionieri furono fatti lavorare nelle minieri.
b. fr. * Les prisoniers ont été fait travailler dans les mines.
'The prisoners have been made working in the mines.'

(6) a. it. Questo libro è stato fatto leggere a Mario da Giovanni.
b. fr. * Ce livre a été fait lire à Mario (par Jean).
'This book has been made reading to/by Mario by John.'

Past Participle Agreement In Italian, complex predicate constructions license past participle agreement of the governing predicate (7a)/(8a), again contrasting with French (7b)/(8b). See (12) for past participle agreement in reflexive complex constructions involving restructuring verbs. Again the contrast carries over to auxiliary constructions (8a–b).

(7) a. it. Giovanni li ha volut*i* leggere.
'John wanted to read them.'
b. fr. Jean les lui a entendu/**s* parler.
'John heared them talking to him.'

(8) a. it. Le tavole che Gianni ha fatt*e* riparare
b. fr. Les tables que Jean a fait/**es* réparer
'The tables that John made repair'

(9) a. it. Maria è stat*a* accusat*a*.
b. fr. Marie a été/**e* accus*ée*.
'Mary has been accused.'

Reflexivization Finally, while both Italian and French allow for reflexivization of the complex causative predicate (10) where an argument of the "embedded" predicate is bound to the functional subject of the complex predicate, Italian does not admit *si*–reflexivization of the embedded predicate (11a), as opposed to French *se*–reflexivization (11b).[10]

(10) a. it. Maria si est fatta accusare da Giovanni.
'Mary made herself (got) accused by John.'
b. fr. Jean s'est fait écraser par une voiture.
'John made himself (got) run over by a car.'

(11) a. it. * Mario ha fatto accusarsi Piero.
b. fr. Marie a fait s'accuser Pierre.
'Mario/Mary made Peter accuse himself.'

Italian restructuring verbs admit long *si*–reflexivization (12a) and long object preposing involving "middle" *si* (12b), where clitic climbing is obligatory (see aux–selection in (12a')).

(12) a. it. I ragazzi si sarebbero volut*i* vedere piu spesso. (essere-aux)
a'. it. I ragazzi avrebbero voluto vedersi piú spesso. Burzio(1986)
'The kids will see each other more often.'
b. it. Queste case gli si vogliono vendere a caro prezzo.
'These houses are wanted to be sold to him at a high price.'

[10] Embedded reflexives only appear with *se stesso* in Italian: *Maria farà accusare se stesso à Giovanni.*

Besides clitic climbing, *tough*–movement and relation changes in causative constructions, the data involving auxiliary change (2a–b), passivization (4)/(5/6a) and long object preposing (12) are in general taken to provide strong arguments for a monoclausal analysis of complex predicates. Therefore the contrasting data of French call for an analysis to capture differences in "degree" of complex predicate formation. Our discussion of this problem will lead us to a more principled question, namely whether complex predicate formation can in fact be argued to take place in syntax, as claimed by e.g. Alsina(1993), or else whether it should be viewed of as a lexical process, to be located in the lexicon component.

2 Complex predicate formation in syntax

Alsina(1993) presents an analysis of causative constructions for the Bantu language Chicheŵa and for Romance languages, developed on the basis of Catalan data. In Chicheŵa the formation of a complex causative predicate is assumed to take place in the lexicon, by combination of a verb stem and a causative affix. The causative construction in Romance languages, however, is argued to be essentially syntactic: it allows for coordination not only of V^0 constituents in the embedded clause, but also of VP–constituents (13).[11]

(13) a. Jean a fait rire et pleurer ses enfants.
'John made his children laugh and cry.'

b. Jean les a fait manger des gâteaux et boire du vin.
'John made them eat cake and drink wine.'

But complex predicate formation in syntax is not easily obtained in the LFG framework:[12] Both clitic climbing and complex relation changes in Romance causative constructions call for a monoclausal functional structure for the two verbal predicates, where the PRED value is composed from the PRED values of the governing and the embedded verb. Yet, in a *monoclausal* structure the PRED values of these respective predicates would have to, but cannot unify.

Alsina therefore introduces the notion of *composition*, as opposed to *unification*, to apply to PRED values. Identity of feature structures (=) is redefined as $=_H$ in (14), where the function F in (15) defines the *composition* of the PRED value of the node annotated by $\uparrow =_H \downarrow$ and the PRED value of its sister node (referred to by $\rightarrow$ PRED). This yields an analysis of complex causative constructions as displayed in (16) (see Alsina(1993:298)).[13]

(14) $\uparrow =_H \downarrow \equiv_{def} (\uparrow \alpha) = (\downarrow \alpha)$ for all $\alpha \neq$ PRED
$(\uparrow$ PRED$) =$ F($(\downarrow$ PRED), $(\rightarrow$ PRED)) Alsina(1993:297)

(15) a. F(x, $\emptyset$) = x Alsina(1993:297)

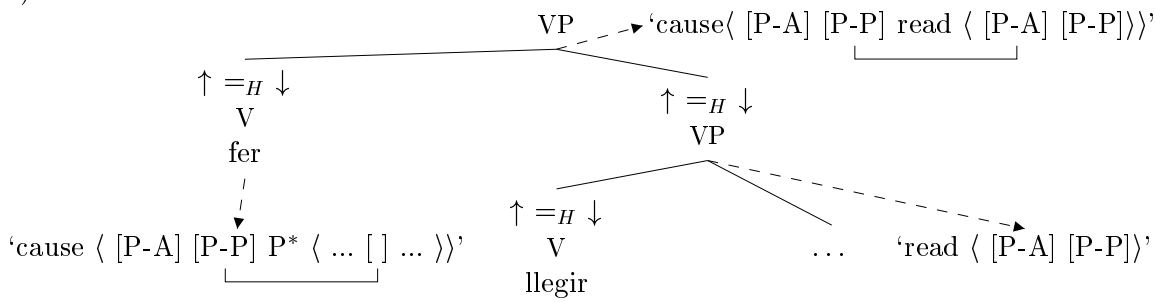
b. F('$P_1\langle a\rangle$', '... $P^*\langle b\rangle$ ...') = '... $P^1\langle c\rangle$ ...'
where "P^* is an unspecified predicator and "c" is the unification of "a" and "b".

[11] See Alsina(1993) for further motivation: separability of the verbal heads that enter the complex construction, and non–availability of morphological derivations (nominalization).

[12] See also Dalrymple et al(1992), Manning(1992/1996).

[13] The use of Proto–Roles (*Proto–agent* [P-A] and *Proto–patient* [P-P]) that is made in Alsina(1993), as opposed to thematic roles like AGENT, THEME, etc. is not essential for our present concerns.

(16)



While Alsina's analysis seems to be well suited for the analysis of causatives in Chichewa and Catalan, there are two problems to it.

First, there are many features other than PRED that are distinct for the predicates that enter into the complex construction and therefore lead to unification failure. Alsina discusses the problem of conflicting inflectional features, which can be circumvented – following Kroeger(1991) and King(1993) – by encoding finiteness in the category INFL. But one cannot argue in a similar way for other features marking, e.g. past participle agreement, reflexivization and auxiliary selection (see below for more detailed discussion).

But more importantly, as it stands the analysis does not account for the contrasting data of Italian and French with regard to auxiliary selection, passivization, past participle agreement and reflexivization. How come, e.g. that the argument structure resulting from argument composition in (16) allows for passivization in Italian, but not in French? Why does it license agreement of the past participle with a cliticized object in Italian, but not in French? How to account for the fact that the "embedded" argument structure can undergo reflexivization in French (with the clitic remaining downstairs), but not in Italian? Finally, auxiliary changes (2)/(12) can only be accounted for if auxiliary selection is not defined within a verb's lexical entry, but defined "on the fly", during syntactic derivation, *after* predicate composition has taken place.

3 Auxiliary selection, reflexivization and PP–agreement

In the following we will try to give an answer to these questions, by sketching an analysis of complex predicate formation in terms of a lexical rule. We build on an analysis of past participle agreement in French (and Italian) in Frank(1989/1990).[14] Further we distinguish, with Rosen(1989), between two variants of predicate composition: *full* vs. *partial* merger of argument structure.

Lexical Mapping Theory and Morphological Structure The analysis is based on the architecture of Lexical Mapping Theory LMT (see e.g. Bresnan(1990)), which defines a mapping from argument structure (a–structure) to f–structure which constrains the assignment of grammatical functions to argument positions. Thematic roles in a–structure are assigned functional features [± o] [± r] by rules of *intrinsic role classification*. These determine underspecified grammatical relations which are further specified either by (morpho)syntactic operations (passive, locative inversion, etc.) or by *default* classification rules.

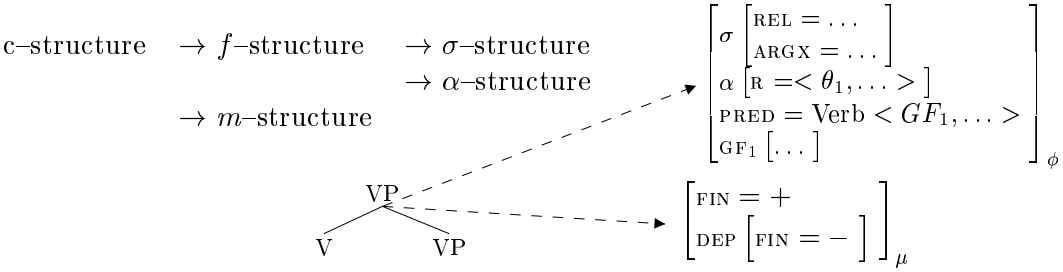
Further the analysis benefits from integration of a recent proposal by Butt et al(1996), to

[14] There are some minor differences in PP–agr in Italian (where reflexivization of ditransitive verbs triggers subject agreement of the past participle) that are not captured by this analysis, but it turned out correct for the relevant facts of Italian complex predicate constructions involving causative and restructuring verbs.

separate off morphological information in a distinct "module" of the overall architecture, the morphological or m–structure. This move is i.a. motivated by a crosslinguistic perspective on the analysis of auxiliaries, which favours a *flat* analysis of auxiliaries at f–structure, as opposed to the "received wisdom", which treats them as raising verbs. Since both c– and m–structure define an embedded structure for auxiliary constructions, it is possible to define selectional constraints by annotating the lower VP–node in (17) by the equation $(\uparrow_\mu \text{DEP}) = \downarrow_\mu$.

Following Butt et al(1996) the m–structure is an immediate projection of the c–structure. The a–structure and the semantic structure (σ–structure) are projected from the f–structure (17).[15] For reference to f–structure vs. m–structure attributes we use the somewhat simplifying notation $\uparrow_\phi$, $\downarrow_\phi$ vs. $\uparrow_\mu$, $\downarrow_\mu$ instead of the more accurate notation $\mu(M(*))$, $\mu(*)$ and $\phi(M(*))$, $\phi(*)$, where $*$ refers to the actual c–structure node. Thus,$\uparrow$ and $\downarrow$ are to be read to refer to the mother and actual c–structure nodes, respectively.

(17)



We briefly state the relevant parts of the analysis of reflexivization, auxiliary selection and past participle agreement in French (and Italian) on the basis of LMT.[16]

Auxiliary Selection We distinguish between lexically determined auxiliary selection (18), which is governed by the thematic unaccusativity property of intransitive verbs, and auxiliary change/selection induced by morphosyntactic processes of argument reduction, such as passivization and reflexivization (see below). This distinction is motivated by reflexive verbs which derive from ditransitive or non-transitive verbs (*acheter*, *parler*). Here the subject has no THEME characteristics, yet we still have selection of *être*.[17] The rules in (18) must be constrained to apply only to lexical entries that do not state (exceptional/intrinsic) auxiliary selection. The AUX feature is characterized as a feature of the m–structure.

(18) Lexically determined auxiliary selection: (L-aux)

a. $(\uparrow_{\phi\alpha} \text{R}) = \text{R}^* \langle \text{ theme } \rangle \rightarrow (\uparrow_\mu \text{AUX}) = \text{être}$

b. $(\uparrow_{\phi\alpha} \text{R}) = \text{R}^* \langle \text{ agent } \rangle \rightarrow (\uparrow_\mu \text{AUX}) = \text{avoir}$

c. $(\uparrow_{\phi\alpha} \text{R}) = \text{R}^* \langle \_, \_ .. \rangle \rightarrow (\uparrow_\mu \text{AUX}) = \text{avoir}$

Reflexivization We distinguish various reflexive constructions: (lexicalized) *intrinsic* reflexive verbs, *se–ergatif*, *se–moyen*, and truly reflexive or reciprocal verbs (see (19) for ditransitive verbs). We assume (with Grimshaw(1982)) that the clitic *se* does not have argument status. The realization of the reflexive clitic (which bears the feature $\uparrow_\mu$ REFL = +) is enforced by the

[15] This decision is not in any way meant to be well founded. There may be good reasons to choose another architecture, in particular with regard to σ–structure. (see in particular Dalrymple et al(1992) for discussion).

[16] These rules have been developed in Frank(1989/1990) to account for a wide range of data of past participle agreement, including reflexivization, (stylistic) inversion, psych–verbs, impersonal constructions, and clitic climbing, which is not achieved by alternative theories (Baker(1983), Burzio(1986), Kayne(1985/1989), Waite(1986)). We have adapted the rules to the architecture in (17).

[17] See also German, which has the unaccusative distinction for *sein/haben*–selection, but no *sein*–selection with reflexive verbs. See also the English passive construction (i) where the subject bears the BENEFICIENT role.
(i) *The children were cooked supper (by the parents)* (Bresnan(1990)).

constraining equation $(\uparrow_\mu \text{REFL}) =_c +$ on the verb. Thus reflexive marking is characterized as a morphological feature. Reflexivization with *se* is characterized as a closed predicate in the semantic form, which accounts for (20) (see Sells et al(1987)). This induces an argument reduction process which is characterized, in the f–structure by assigning the bound argument role the function NULL. Semantic argument binding makes use of the Kleene star on ARG in (19) to account for "long reflexivation" in complex constructions (see below).

(19) Reflexivization (Rvz): (Kleene star on ARG set to 0 for ordinary ditransitive verbs)

$$\begin{array}{ll} (\uparrow_{\phi\sigma} \text{REL ARG}^* \text{ ARG1}) = \text{x}^* & (\uparrow_{\phi\sigma} \text{REL ARG}^* \text{ ARG2}) = (\uparrow_{\phi\sigma} \text{REL ARG}^* \text{ ARG1}) \\ (\uparrow_{\phi\sigma} \text{REL ARG}^* \text{ ARG2}) = \text{y}^* & (\uparrow_{\phi\alpha} \text{R}) = \text{R}^* \langle \text{ agent, ben, theme} \rangle \\ (\uparrow_{\phi\alpha} \text{R}) = \text{R}^* \langle \text{ agent, ben, th} \rangle \quad \rightarrow & (\uparrow_{\phi} \text{PRED}) = \text{se-Verb}^* \langle \_ , (\text{NULL}), \_ \rangle \\ (\uparrow_{\phi} \text{PRED}) = \text{Verb}^* \langle \_\_, \_\_, \_ \rangle & (\uparrow_{\mu} \text{REFL}) =_c + \\ (\uparrow_{\mu} \text{AUX}) = \text{avoir} & (\uparrow_{\mu} \text{AUX}) = \text{être} \end{array}$$

(20) Jean se défend mieux que Pierre.
(a)= Jean se défend mieux que Pierre se défend. (sloppy)
(b)≠ Jean se défend mieux que Pierre le défend. (strict)

Past Participle Agreement The lexical rule of past participle agreement (23) is based on the notions of TRANSitivity (21) and an abstract notion of UNACCusativity (22). We characterize as TRANSitive verbs where both the SUBJ and OBJ function are associated with thematic arguments. In LMT this is captured by (21). UNACCusativity is then divided into (22a) *thematic* unaccusativity of non–TRANSitive verbs which select *être* – and which induces agreement of the past participle with the SUBJect in (23a) – and a *structural* notion of unaccusativity for TRANSitive verbs (22b), which obtains when the OBJect is not derived in the canonical VP–position (by cliticization, wh–construction, etc.) and thus leads to some sort of intransitivization of the VP–structure – and which induces OBJect agreement in (23b).[18, 19]

(21) Transitivity:

$$\begin{array}{lll} \text{a.} & (\uparrow_{\phi\alpha} \text{R}) = \text{R}^* \langle \ \theta_x \ \ldots \ \theta_y \ \ldots \rangle & \rightarrow \ (\uparrow_{\phi} \text{TRANS}) = + \\ & \quad\ \ (\uparrow_{\phi} \text{PRED}) = \text{V}^* \langle \ (\text{SUBJ}) \ldots (\text{OBJ}) \ldots \rangle & \\ \text{b.} & \text{otherwise} & \rightarrow \ (\uparrow_{\phi} \text{TRANS}) = - \end{array}$$

(22) a. Thematic unaccusativity:

$$\begin{array}{lll} (\uparrow_{\phi} \text{TRANS}) = - & & \\ \text{a.} \ \begin{array}{l} (\uparrow_{\phi\alpha} \text{R}) = \text{R}^* \langle \text{ .. theme} \rangle \\ (\uparrow_{\mu} \text{AUX}) = \text{être} \end{array} & \rightarrow & (\uparrow_{\phi} \text{UNACC}) = + \\ \text{b.} \ \text{otherwise} & \rightarrow & (\uparrow_{\phi} \text{UNACC}) = - \end{array}$$

b. Structural unaccusativity:[16]

$$(\uparrow_{\phi} \text{TRANS}) = + \ \rightarrow \ \begin{array}{l} \{/ \ (\uparrow_{\phi} \text{UNACC}) = + \\ \ / \ (\uparrow_{\phi} \text{UNACC}) =_c - \ /\} \end{array}$$

(23) Past participle agreement:

$$\begin{array}{l} (\uparrow_{\mu} \text{PART}) = + \\ (\uparrow_{\phi} \text{UNACC}) = + \\ \text{a.} \ (\uparrow_{\phi} \text{TRANS}) = - \rightarrow \ \begin{array}{l} (\uparrow_{\mu} \text{NUM}) = (\uparrow_{\phi} \text{SUBJ NUM}) \\ (\uparrow_{\mu} \text{GEN}) = (\uparrow_{\phi} \text{SUBJ GEN}) \end{array} \\ \text{b.} \ (\uparrow_{\phi} \text{TRANS}) = + \rightarrow \ \begin{array}{l} (\uparrow_{\mu} \text{NUM}) = (\uparrow_{\phi} \text{OBJ NUM}) \\ (\uparrow_{\mu} \text{GEN}) = (\uparrow_{\phi} \text{OBJ GEN}) \end{array} \end{array}$$

[18] The relevant phrase structure interacting with (22b) rule is:

$$\begin{array}{lll} \text{VP} \rightarrow & \text{V} & [\text{NP}] \\ & \uparrow = \downarrow & (\uparrow_{\phi} \text{OBJ}) = \downarrow_{\phi} \\ & & (\uparrow_{\phi} \text{UNACC}) = - \end{array}$$

[19] Following Butt et al (1996) verbal inflection is represented in m–structure, while inflectional features NUM, PERS, etc. of nominals are stated in the f–structure. Otherwise one has to make use of the inverse function ϕ^- in (23): $(\uparrow_\mu \text{NUM}) = ((\uparrow_\phi \text{SUBJ})_{\phi^-\mu} \text{NUM})$. See Dalrymple et al(1992).

4 A lexical rule for complex predicate formation

Let us now, on the basis of these rules, try to account for the contrastive data of French and Italian complex predicates. To this end we adapt to the LFG framework a proposal of Rosen(1989), to distinguish between *full merger* vs. *partial merger* of argument structures.[20]

4.1 Full vs. partial merger of argument structure

Both types of *merging* involve some sort of *incorporation* of the full argument structure of a "dependent" predicate into the argument structure of the "light verb" (causative, restructuring verb, etc.), which we assume to host an argument role standing proxy for an eventuality–type (the caused/etc. situation or event) and thus licenses incorporation of an argument structure that corresponds to this type of thematic role.

Now, *full merger* $\oplus_{full}$ defines this incorporation of argument structure to yield an atomic, fused argument structure, where the argument structure of the dependent predicate not only replaces $\theta_{ev-type}$, but is defined to concatenate with the (remaining) argument positions of the incorporating predicate. By contrast, *partial merger* $\oplus_{par}$ defines this incorporation in a way such that the argument list to replace the eventuality–type role is not concatenated with the governing argument list. While being incorporated into the governing predicate's argument structure, it preserves its status as an (embedded) atomic argument structure.

(24) a. Full merger: $\langle\ \theta_1, \ldots, \theta_n, \theta_{ev-type}\rangle \oplus_{full} \langle \text{A–list}\rangle =_{def} \langle\ \theta_1, \ldots, \theta_n \mid \text{A–list}\ \rangle$

b. Partial merger: $\langle\ \theta_1, \ldots, \theta_n, \theta_{ev-type}\rangle \oplus_{par} \langle \text{A–list}\rangle =_{def} \langle\ \theta_1, \ldots, \theta_n, \langle \text{A–list}\rangle\ \rangle$

On the basis of these two types of argument structure incorporation, which define different "degrees" of complex predicate formation, it is possible to explain the observed contrasts between French and Italian, by postulating a language specific "parameter" that constrains complex predicate formation to make use of *full merger* $\oplus_{full}$ in Italian, as opposed to *partial merger* $\oplus_{par}$ in French (cf. Rosen(1989)). We will first lay out the basic ideas and then investigate how this view on argument composition could be integrated into an LFG analysis of complex predicate formation – in syntax or by lexical rule.

In (25) we state a function *CPF*,[21] which applies to a pair of functional specifications, corresponding to the predicates that enter into complex constructions, here a causative construction. The function applies uniformly to French and Italian, yet with an underspecified *merger* operation $\oplus$ for composition of argument structures, which is to be set to $\oplus_{par}$ and $\oplus_{full}$, respectively. The function defines, at the level of the σ–projection, the composition of the semantic structures of the two predicates, to yield the *cause*–relation ${}_i\uparrow_{\phi\sigma} = {}_g\uparrow_{\phi\sigma}$ where $({}_i\uparrow_{\phi\sigma}$ ARG2) is defined as reentrant with ${}_h\uparrow_{\phi\sigma}$ of the second argument.

At the level of *a*–structure the R values of ${}_g\uparrow_{\phi\alpha}$ and ${}_h\uparrow_{\phi\alpha}$ are (fully or partially) *merged*, according to (24), to yield ${}_i\uparrow_{\phi\alpha}$.[22] While the merging of argument structures is subject to different "degrees of incorporation" in French and Italian, the PRED value of the outcome ${}_i\uparrow_{\phi}$ is defined

[20] In Frank(1989/1990) we drew a different line of distinction, where for Italian we assumed full merger of argument structure (yet represented as in (24b)) as opposed to a sentential function COMP for causative verbs in French, which allowed for (controlled) climbing of clitics in coherent (CLX) constructions.

[21] To avoid immediate inconsistencies the arrows come with distinguishing (prefixed) function indices.

[22] Contrary to other approaches (see e.g. Alsina(1993), Butt(1995a/b)) we do not represent a PATIENT role for the causative predicate, to correspond to the causee. This decision might well turn out to be shortsighted, since we didn't (and will not) go into the lexical semantics of causatives. But we assume that the semantics of

as an ordinary, atomic list of (underspecified) grammatical functions. This characterizes complex predicate formation to yield a monoclausal *functional* structure, which accounts for its characteristic properties of both clitic "climbing" and complex relation changes in Romance causative constructions. The latter characteristics is obtained by assigning the functional feature [+ o] to the highest argument role θ_1 of the embedded predicate (see below). Finally, the ${}_g\uparrow_\mu$ value of the causative (which besides its local μ values constrains the μ values of the second argument in terms of its DEP value) defines the μ value of the complex predicate ${}_i\uparrow_\mu$.

$$
\begin{array}{lll}
(25) & \text{CPF} \Big(& \begin{array}{l} ({}_g\uparrow_{\phi\sigma} \text{ REL}) = \text{cause} \\ ({}_g\uparrow_{\phi\sigma} \text{ ARG1}) = \text{x} \\ ({}_g\uparrow_{\phi\sigma} \text{ ARG2}) = \text{P}^* \\ ({}_g\uparrow_{\phi\alpha} \text{ R}) = \text{cause}\langle \text{agent, event} \rangle \\ ({}_g\uparrow_{\phi} \text{ PRED}) = \text{faire/fare}\langle \text{ \_\_, \_\_ } \rangle \\ ({}_g\uparrow_{\mu} \text{ CLX}) = + \\ ({}_g\uparrow_{\mu} \text{ DEP}) = [2] \end{array} \quad , \quad \begin{array}{l} ({}_h\uparrow_{\phi\sigma} \text{ REL}) = \text{Rel}^* \\ ({}_h\uparrow_{\phi\sigma} \text{ ARG1}) = \text{y}^* \quad \ldots \\ ({}_h\uparrow_{\phi\sigma} \text{ ARGN}) = \text{z}^* \\ ({}_h\uparrow_{\phi\alpha} \text{ R}) = \text{R}^* \langle \theta_1 \mid \text{TL} \rangle \\ ({}_h\uparrow_{\phi} \text{ PRED}) = \text{Verb}^* \text{ GFL}_2^* \\ {}_h\uparrow_{\mu} = [2] \end{array} \Big) = \\
 & & \begin{array}{l} {}_i\uparrow_{\phi\sigma} = {}_g\uparrow_{\phi\sigma} \\ ({}_i\uparrow_{\phi\sigma} \text{ ARG2}) = {}_h\uparrow_{\phi\sigma} \\ ({}_i\uparrow_{\phi\alpha} \text{ R}) = \text{cause-R}^* \langle \text{ agent } \rangle \oplus \langle \theta_1 \mid \text{TL} \rangle \\ ({}_i\uparrow_{\phi} \text{ PRED}) = \text{faire/fare-Verb}^* \text{ GFL}_3^* \\ {}_i\uparrow_{\mu} = {}_g\uparrow_{\mu}. \end{array}
\end{array}
$$

How does *CPF* – in conjunction with (24) – account for the data outlined in Section 1?

Clitic climbing Given that *CPF* yields a monoclausal f–structure "climbing" of (object) clitics is straightforwardly explained: the grammatical functions corresponding to arguments of the incorporated predicate are now defined within the functional clause of the complex construction and therefore can cliticize to the "higher" verb. To prohibit cliticization to the lower V,[23] one could define the embedded VP in the phrase structure rule for *complex* predicates (annotated by $(\uparrow_\mu \text{ CLX}) = +$) not to license the derivation of object clitics. But see Section 4.3 for an more principled solution.

Relation Changes The relation changes that are characteristic for French and Italian causative constructions are captured by assigning θ_1 in $({}_i\uparrow_{\phi\alpha} \text{ R})$ in (25) the functional feature [+ o] if θ_1 is AGENT. This is illustrated in (26) for composition with an intransitive vs. transitive predicate in French. The two AGENT roles that cooccur in a complex argument structure are somehow to be distinguished. If irrespective of the particular type of merging the "lower" AGENT is assigned the feature [+ o], this yields what in traditional accounts was covered by the term *subject demotion* (Comrie(1976)). Application of the rules of intrinsic and default classification[24] maps the lower AGENT in (26a) to the OBJ function, while in (26b) the AGENT surfaces as an OBJ_θ and the THEME as an OBJ.[25]

causation *implies* that the agent of the caused event – if there is one – is directly or indirectly the patient, or causee of the causation. This more sparse representation is then also open to causative constructions that do not involve a patient or causee, such as e.g. *Dieu a fait pleuvoir*, where there is clearly a caused event, that of raining, while no a patient, or causee. Also, the status of "fused" thematic roles in argument structure (e.g. by internal linking of a *patient* and an *agent* role, cf. (16)) as regards intrinsic role classification is not clear.

[23] But see Manning(1996) who argues for more sophisticated constraints, to leave room for downstairs clitics in some Romance complex constructions.

[24] In Frank(1989/1990) we assumed the rules in (i), which are equivalent to those of Bresnan(1990).

(i) a. Intrinsic role classification: AGENT [− o] THEME/PATIENT [− r]

b. Default Classification: Assign default values from left to right. Choose the least "marked" value possible from the hierarchy of functional features: [− r] < [− o] < [+ o] < [+ r]

[25] We assume there to be different sorts of "demotion" of one or the other of the two logical subjects in a single argument structure: Besides the assignment of [+ o] to the "lower" agent (26), we get a passive–like

(26) a.
$$\begin{array}{lll} (\uparrow a\ \textsc{r}) = \text{cause–Intrans–R}^*\langle & \text{agent}, & \langle\ \text{agent}\rangle\ \rangle \\ & [-\ \text{o}] & [+\ \text{o}] \\ (\uparrow \textsc{pred}) = \text{faire–Intrans–V}^*\langle & (\textsc{subj}) & (\textsc{obj})\ \rangle \end{array}$$

b.
$$\begin{array}{lll} (\uparrow a\ \textsc{r}) = \text{cause–Trans–R}^*\langle & \text{agent}, & \langle\ \text{agent}, \ \text{theme}\rangle\ \rangle \\ & [-\ \text{o}] & [+\ \text{o}] \quad [-\ \text{r}] \\ (\uparrow \textsc{pred}) = \text{faire–Trans–V}^*\langle & (\textsc{subj}) & (\textsc{obj}_\theta)\ (\textsc{obj})\ \rangle \end{array}$$

Passivization The distinction between full vs. partial merger of argument structure immediately accounts for the French–Italian contrast as regards passivization of complex predicates. In LMT passivization is constrained in terms of the predicate's argument structure,[26] where we assume, for French and Italian, that only transitive verbs (that host a thematic object) undergo passivization. With full merging of argument structures the Italian complex causative predicate will qualify as transitive, such that the passive rule can apply. As mentioned in fn. 25, if assignment of [+ o] to θ_1 in (25) is optional in case of a matrix passive, our linking rules predict the two types of passivization in (5a) and (6a). In French, however, as indicated in (26), complex predicate formation results in a *partially* merged argument structure, such that the passive rule can in principle apply to either one of both "levels" of the argument structure. Since the "higher" level does not qualify as transitive, long passivization is ruled out.[27]

Past Participle Agreement It is predicted that causative constructions in Italian trigger past participle agreement, as opposed to French. In Italian the fully merged argument structure qualifies as TRANSitive, such that according to (22b) and (23b) a dislocated OBJ will trigger object agreement of the participle. Again the French causative does not qualify as TRANSitive, such that according to (18c) and (22a) past participle agreement is ruled out (23).

Reflexivization For reflexivization we assume that the binder and bindee can in principle pertain to different propositional domains in σ–structure (see (19) with Kleene star on ($\uparrow_{\phi\sigma}$ REL ARG* ARG1) and ($\uparrow_{\phi\sigma}$ REL ARG* ARG2). Yet, in French and Italian binder and bindee must be realized within the same (functional) clause nucleus.[28] An additional restriction for *se/si*–reflexivization is that the "higher" argument must be linked to an *external argument*, which is the highest thematic role within an argument structure. This accounts for "matrix" reflexivization of causatives in Italian and French, where an argument pertaining to the embedded relation is bound to the causative agent (10). This involves binding of an argument to another argument outside of its propositional domain, yet the causative agent still qualifies as an *external argument* of the (composite) argument structure within the functional clause nucleus of the bound element. But we predict a difference with regard to reflexivization of the embedded predicate. Since in French, due to partial merging, the embedded predicate still qualifies as a complete, incorporated argument structure, reflexivization as in (11b) is licit: the THEME of *accuser* can be bound to the lower AGENT, which qualifies as an external argument in the local functional clause nucleus. In the Italian fully merged argument structure, however, the lower AGENT is not an external argument, which rules out (11a). Reflexive

effect for the "embedded" predicate by assigning it the feature [+ r]. Another (alternative) way for demotion of one of the agents is matrix passivization (see below), where the "higher" agent role is mapped to an oblique function by assignment of [+ r]. In this case, assignment of [+ o] to the lower agent is optional, which yields the different passive structures in (5a) vs. (6a). We cannot detail this for lack of space.

[26] Passivization: $(\uparrow_{\phi\alpha}\ \textsc{r}) = \text{R}^*\langle\ \theta_x, \ldots, \text{theme}\rangle$, $(\uparrow_{\mu}\ \textsc{part}) = +$ $\rightarrow$ $(\uparrow_{\phi\alpha}\ \textsc{r}) = \text{R}^*\langle\ \theta_x\ [+\ \text{r}], \ldots, \text{theme}\rangle$, $(\uparrow_{\mu}\ \textsc{aux}) = \text{être}$.

[27] Yet, passivization may well apply to the lower level of the argument strucure, if transitive as in (26b), by assigning the lower agent the feature [+ r] instead of the feature [+ o] for subject demotion.

[28] For a wellfounded analysis of locality constraints for anaphoric binding see Dalrymple(1990).

marking must be tied to the μ–projection in (10) to enforce the clitic to "climb", while in (11b) reflexive marking must affect the DEPendent μ projection, to enforce the realization of the clitic downstairs (to be discussed in detail below).

Auxiliary Selection Finally, complex predicate formation with Italian restructuring verbs is to be defined in such a way that the complex predicate's argument structure is identified with the argument structure of the embedded predicate (see *light merger* in Rosen(1989)).[29] This is motivated by the fact that restructuring verbs are (optionally complex) subject raising or control verbs, where the "matrix" and "embedded" subject map to the same individual. It is then predicted that the auxiliary selection of the complex predicate is determined by the thematic structure of the embedded predicate (18). Also, if reflexivization applies to the complex argument structure, auxiliary change to *essere* and climbing of *si* is predicted (12).

Finally, for complex restructuring predicates that either have undergone reflexivization (12a), qualify as thematically unaccusative (9a) or structurally unaccusative (7a)/(8a) we predict past participle agreement. For the French auxiliary (see (9b)) we then have to assume that, contrasting to Italian, it does not involve complex predicate formation, but must be analyzed as a purely morphological element, along the lines of the *flat* analysis in Butt et al(1996).

In sum, Rosen's distinction between full and partial merger of argument structures in complex predicate formation – which defines a *monoclausal* argument structure in Italian, as opposed to a *hierarchical/complex* argument structure in French – accounts for the contrastive data outlined above, while still capturing the common "basic" syntactic characteristics that motivate a monoclausal *functional* structure: clitic climbing and relation changes.

We will now investigate two different ways to integrate the function *CPF* into an LFG grammar. The obvious way is to follow Alsina's analysis and locate the function in syntax. But it will turn out that, besides principled difficulties arising, again, from conflicting features in complex constructions, this account makes strong predictions as to the interaction of syntax, lexicon and morphosyntactic lexical rules that lead to wrong results. We will then, in 4.3, propose an alternative approach, to characterize complex predicate formation as a lexical rule.

4.2 Complex predicate formation in syntax?

While following Alsina's analysis in spirit, we use a slightly different rule for complex predicate formation, by functional definition of the $\uparrow$ value of the node that directly dominates the complex construction (27). Here *CPF* (adapted in (28) to application in (27)) applies to the functional structures referred to by the function variables ${}_{g}\downarrow$ and ${}_{h}\downarrow$, which we take to be coreferential with ${}_{g}\uparrow$ and ${}_{h}\uparrow$ of the left and right daughter node in (27), respectively.[30]

[29] The relevant changes to CPF_{caus} (25) are sketched below for CPF_{res}:

$$CPF_{res}\Big(\begin{array}{l} ({}_{g}\uparrow_{\phi\alpha}\ \text{R}) = \text{restrR}^{*}\langle(\theta_x),\text{event}\rangle \\ ({}_{g}\uparrow_{\phi}\ \text{PRED}) = \text{restrV}^{*}\ \text{GFL}_1 \end{array} , \begin{array}{l} ({}_{h}\uparrow_{\phi\alpha}\ \text{R}) = \text{R}^{*}\ \text{TL} \\ ({}_{h}\uparrow_{\phi}\ \text{PRED}) = \text{V}^{*}\ \text{GFL}_2 \end{array} \Big) = \begin{array}{l} ({}_{i}\uparrow_{\phi\alpha}\ \text{R}) = \text{restrR}*\text{-R}*\ \text{TL} \\ ({}_{i}\uparrow_{\phi}\ \text{PRED}) = \text{restrV}*\text{-V}*\ \text{GFL}_3 \end{array}$$

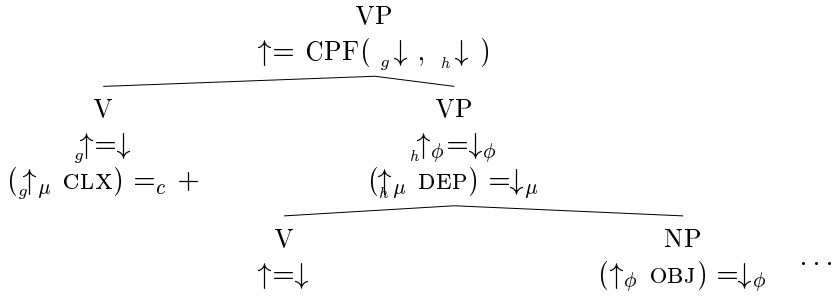
[30] To make this work, the standard definition of $\uparrow$ and $\downarrow$ must be abandoned, where $\uparrow$ refers to the f–structure of the mother node: $\uparrow =_{def} \phi(M(*))$ and $\downarrow$ refers to the f–structure of the actual node: $\downarrow =_{def} \phi(*)$.

Instead, $\uparrow$ must be defined to refer to the f–structure of the actual node: $\uparrow =_{def} \phi(*)$, while $\downarrow$ refers to the union of the f–structures of the daughter nodes: $\downarrow =_{def} \phi(\text{D}_1(*)) \cup \phi(\text{D}_2(*)) \cup ... \cup \phi(\text{D}_n(*))$.

As before, to make the definitions more readable, we use the simplifying notation $\uparrow_{\mu}$, $\downarrow_{\mu}$ and $\uparrow_{\phi}$, $\downarrow_{\phi}$ to refer to the μ– and ϕ–values of the actual and daughters' c–structure nodes, respectively.

Since we will finally not adopt the syntactic account to complex predicate formation in (27)/(28), we will, in Section 4.3., return to the standard definition where $\uparrow$ refers to the mother node.

(27)



It is now evident how the conflicting σ– and α–values of the two predicates can be "reconciled" by application of *CPF* (28): the σ–structure $_i\uparrow_{\phi\sigma}$ is projected from the light verb, with incorporation of the dependent predicate into the corresponding argument position ARG2. The α–projection $_i\uparrow_{\phi\alpha}$is computed from $_g\downarrow_{\phi\alpha}$ and $_h\downarrow_{\phi\alpha}$ in terms of (full or partial) *merging* of the respective R–lists. Now, while we can in a similar way redefine the PRED value of $_g\downarrow_\phi$ to yield a new, composite predicate name (and list) for $_i\uparrow_\phi$, there will be additional features of the respective functional projections $_g\downarrow_\phi$ and $_h\downarrow_\phi$ that are (potentially) incompatible: e.g., the ϕ–features TRANS and UNACC,[31] but also ϕ–features like NEG. On the other hand, all of the nonconflicting ϕ–features defined by *both* daughter constituents must project to $_i\Uparrow_\phi$, such as nominal or cliticized GF's and ADJuncts. Moreover, since α and σ are represented as projections within the f–structure, the values $\downarrow_{\phi\alpha}$ and $\downarrow_{\phi\sigma}$ of g and h must not be projected to $_i\Uparrow_{\phi\alpha}$ and $_i\Uparrow_{\phi\sigma}$.

The only way out of this problem is to resort to the restriction operator \ introduced in Dalrymple et al(1992) and Kaplan/Wedekind(1993), which in (28) defines the value of $_i\Uparrow_\phi$ in terms of the partial functional structure $_g\downarrow_\phi \setminus \{$PRED, TRANS, UNACC, α, $\sigma\}$ that is identical to $_g\downarrow_\phi$ except for being undefined for the features within the set following \, and similarly for $_i\Uparrow_\phi = {}_h\downarrow_\phi \setminus \{$PRED, TRANS, UNACC, α, $\sigma\}$.[32]

(28) CPF $\Big($ $(_g\downarrow_{\phi\sigma}$ REL$) =$ cause
$(_g\downarrow_{\phi\alpha}$ R$) =$ cause$\langle$agent, event$\rangle$
$(_g\downarrow_\phi$ PRED$) =$ faire/fare$\langle$ __, __ $\rangle$
$(_g\downarrow_\mu$ CLX$) = +$
, $(_h\downarrow_{\phi\alpha}$ R$) =$ R* $\langle\theta_1 \mid$ TL $\rangle$
$(_h\downarrow_\phi$ PRED$) =$ Verb* <__ | L> $\Big) =$

$_i\Uparrow_{\phi\sigma} = {}_g\downarrow_{\phi\sigma}$
$(_i\uparrow_{\phi\sigma}$ ARG2$) = {}_h\downarrow_{\phi\sigma}$
$(_i\uparrow_{\phi\alpha}$ R$) =$ cause-R$^*\langle$ agent $\rangle \oplus \langle\ \theta_1 \mid$ TL $\rangle$
$(_i\uparrow_\phi$ PRED$) =$ faire/fare-Verb* GFL*
$_i\Uparrow_\phi = {}_g\downarrow_\phi \setminus \{$ PRED, TRANS, UNACC, α, σ $\}$
$_i\Uparrow_\phi = {}_h\downarrow_\phi \setminus \{$ PRED, TRANS, UNACC, α, σ $\}$
$_i\uparrow_\mu = {}_g\downarrow_\mu$.

While the problem of conflicting ϕ–features can be circumvented by resorting to the notion of *restriction*, the *syntactic* account to argument composition makes strong predictions as to the interaction of syntax, lexicon and morphosyntactic lexical rules. In particular, the rules of

[31] While one could argue that TRANS is to be characterized as a feature of the a–structure (cf. (21)), this will not carry over to UNACC, which is dependent on the c–structural realization of the OBJect function for object agreement. Also, UNACC cannot be represented as a feature of the μ–projection: it will then not account for OBJect agreement of the past participle (see below).

[32] See Kaplan/Wedekind(1993):
If f is an f–structure and a is an attribute: $f\backslash a = f\mid_{Dom(f)-\{a\}} = \{< s, v >\in f \mid s \neq a\}$.
Above we extend the usage of the restriction operator to a *set* of attributes:
If f is an f–structure and A is a set of attributes: $f\backslash A = f\mid_{Dom(f)-A} = \{< s, v >\in f \mid s \notin A\}$.

reflexivization, passivization and PP–agr must be applied "on the fly", during syntactic analysis. The same holds for the linking rules from argument structure to grammatical functions.

This is most evident for the linking rules, since the composed argument structure is in place only at the level of ${}_i\uparrow_{\phi\alpha}$. The mapping rules cannot, therefore, in general be precompiled in the lexicon, where the argument structures of the involved predicates are isolated.

Also, the lexical rule of passivization must apply *after* argument composition has taken place in order to allow for long passive in Italian (4)/(5/6a)), and similarly the rule for past participle agreement, since both of them are dependent on the feature TRANS = + that is to be determined anew for the argument structure resulting from application of *CPF*.[33] Similarly the lexical rule of reflexivization must apply "on the fly". Otherwise we could not have "matrix" reflexivization with causatives (10), or long *si*–reflexivization (12) with restructuring verbs. But this complicates matters as to embedded reflexivization in Italian.

If the rule of reflexivization is free to apply in the process of syntactic analysis, this should predict embedded reflexivization with causative structures to be ok in Italian: in (27), at the level of the embedded V, the argument structure of the embedded predicate is still atomic, such that binding to the highest argument role involves an external argument. In order to detect that this constraint on reflexivization is violated *after CPF* has been applied, the composition of argument structures must to the least involve a wellformedness constraint on reflexive binding. But this will be difficult, since the merged argument structure does now not go along with reflexive marking: the reflexive feature is closed up within ${}_i\uparrow_{\mu}$ DEP. It is then not easy to distinguish between *si/se* reflexivization and reflexivization with *se stesso*, which does not require binding to an external argument and which is grammatical under *fare* (see fn.10).

Another problem for "free" application of reflexivization shows up with restructuring verbs. If the embedded verb is transitive, as in (12a), reflexivization could well apply *before* or *after* argument composition in syntax. If it applies *before* composition takes place, reflexive marking is assigned to the DEPendent μ–feature and the clitic is predicted to remain downstairs. Yet, if complex predicate formation takes place (as evidenced by long object preposing or long reflexivization going along with auxiliary change), clitic climbing is obligatory (cf. (12a–a')).[34]

So while attractive from a theoretical perspective, it turns out that "free" application of both linking rules and morphosyntactic rules in a syntactic account to complex predicate formation not only is quite complex from a computational perspective, but also leads to unwarranted predictions.

We therefore want to give a brief sketch of an alternative approach to complex predicate formation, in terms of a lexical rule, which avoids the computational complexities of the syntactic approach to argument composition, and moreover yields the right predictions as to the observed contrasts between French and Italian complex predicate constructions.

4.3 Complex predicate formation by lexical rule

The idea is simple, starting from LFG's basic insight, that relation changing rules (passivization, dative shift, etc.) are to be defined in terms of lexical redundancy rules that apply to a finite set of lexical entries, as opposed to the application of rules during syntactic derivation.[35]

[33] Since UNACC is represented as a ϕ–feature, it will project to the higher predicate if the OBJect is processed in its base position, to – correctly – rule out agreement with the OBJect.

[34] Similar problems that arise for auxiliary selection with restructuring verbs can be avoided if the AUX values of ${}_i\uparrow_{\mu}$ AUX and ${}_i\uparrow_{\mu}$ DEP AUX are identified in the *CPF*–rule for restructuring verbs. But a similar move does not help for the problem of embedded *si*–cliticization.

[35] "It is important to note that these relation changing rules [dative shift, passive] are not applied in the syntactic derivation of individual sentences. They merely express patterns of redundancy that

Now one could argue that in the framework of Lexical Mapping Theory lexical entries with *underspecified* grammatical functions can enter the syntactic component to yield distinct *specified* functional assignments by interaction with morphosyntactic rules and syntactic analysis. This seems to be the basic idea underlying the approach of argument composition in syntax.

But as we have made efforts to show in the previous Section, the data to be covered in the analysis of complex predicate formation, in particular passivization and reflexivization, can only be accounted for if applied on the fly in syntactic analysis, *after* the individual predicates have undergone argument structure composition. I.e. these rules would somehow have to be restricted not to apply *freely*, in particular not *before* argument composition has taken place.

In sum, the syntactic account to complex predicate formation is committed to application of morphosyntactic lexical rules in syntax. This not only involves a high degree of computational complexity, but will either undermine basic tenets of LFG in that the analysis is forced to impose ordering constraints for application of syntactic rules, or else will yield wrong results.

These considerations naturally lead us to a view of argument composition as a lexically driven process, triggered by lexical rule. This rule must, however, differ in various respects from other relation changing lexical rules, such as passivization, reflexivization, etc. For one, it cannot apply to a single (lexical) element, but must involve *two* verbal predicates. Yet, upon closer inspection, many lexical rules must be considered as triggered by *two* elements, a verbal stem and some morphological marking, which *together* define a new, morphologically derived lexical item: Passivization, e.g., can be viewed of as induced by the language particular morphological passive marking and a verbal stem, to yield a (morphologically) derived verb form that is marked for the appropriate relation changes and further morphosyntactic features. *se/si* reflexivization can be taken to instantiate another class of lexical rule, triggered by the reflexive clitic and a verbal stem, to yield *two* discontinuous lexical items, the clitic and the verb stem, marked for appropriate relation changes and morphosyntactic features.

Our lexical rule of complex predicate formation (29) then constitutes just another class of lexical rule, which applies to *two* verb stems, to yield two discontinuous verb stems, where the changes induced to the different levels of grammatical representation are roughly as in (28).[36]

As before, the functional descriptions of the input arguments, here V_1 and V_2, are indexed by variables g, h, and now yield two distinct functional descriptions indexed g', h' for the derived lexical items $V_{1'}$ and $V_{2'}$. Further we must assume that the variables (indicated by *) are globally defined (for g, h, g', h') within the scope of the lexical rule.

The definition of the σ–, α–, PRED–, and μ–values of g' of $V_{1'}$ is equivalent to (28). What differs is the use of a new ϕ–feature CO–PRED, which takes as value the predicate name Verb* of the dependent predicate V_2. Together with the definition of CO–PRED in $V_{2'}$ the constraining equation in the definition of $V_{1'}$ ensures that – though discontinuous – the two lexical items that result from application of *CPF* contribute to the formation of the appropriate complex predicate in syntax. The embedded verb is thus characterized as a CO–PREDicator of the

obtain among large but finite classes of lexical entries. [...] Indeed [...] our formalism [...] embodies a [...] prohibition against syntactic manipulations of function assignments and function/argument mappings:

(12) *Direct Syntactic Encoding* No rule of syntax may replace one function name by another.

[...] The principle of direct syntactic encoding sharpens the distinction between two classes of rules: rules that change relations are lexical and range over finite sets, while syntactic rules that project onto an infinite set of sentences preserve grammatical relations." Bresnan/Kaplan(1995:35/36)

[36] Given that the lexical rule of *CPF* does not involve a morpological item, the problems raised in Alsina(1993), that Romance causatives do not trigger nominalization, can be rejected: the rule only applies to verb stems. Also, VP–coordination (13b), which is mentioned as a problem for a lexical rule account, can be captured by resorting to the analysis of coordination in Kaplan/Maxwell(1988), as it is also required for a flat auxiliary analysis of *Jean a mangé des gâteaux et bu du vin.* (13a) must be viewed as an instance of right node raising.

complex argument structure that is mapped to the PRED value carried by $V_{1'}$.

Even though the functional description of a lexical entry is fixed, we prefer to define ${}_{g'}\uparrow_{\phi}$ in *CPF* by way of the restriction operator. *CPF* may then freely apply to different verb types.

(29)

V_1: $({}_{g}\uparrow_{\phi\sigma}\ \text{REL}) = \text{cause}$
$({}_{g}\uparrow_{\phi\sigma}\ \text{ARG1}) = x^*$
$({}_{g}\uparrow_{\phi\sigma}\ \text{ARG2}) = P^*$
$({}_{g}\uparrow_{\phi\alpha}\ \text{R}) = \text{cause}\langle \text{agent, event}\rangle$
$({}_{g}\uparrow_{\phi}\ \text{PRED}) = \text{faire}\langle\ \_,\ \_\_\ \rangle$
$({}_{g}\uparrow_{\mu}\ \text{CLX}) = +$
$({}_{g}\uparrow_{\mu}\ \text{DEP}) = M^*$

$\rightarrow_{CPF_{caus}}$

$V_{1'}$: ${}_{g'}\uparrow_{\phi\sigma} = {}_{g}\uparrow_{\phi\sigma}$
$({}_{g'}\uparrow_{\phi\sigma}\ \text{ARG2}) = {}_{h}\uparrow_{\phi\sigma}$
$({}_{g'}\uparrow_{\phi\alpha}\ \text{R}) = \text{cause-R}^*\langle\ \text{agent}\rangle \oplus \langle\ \theta_1 \mid \text{TL}\ \rangle$
$({}_{g'}\uparrow_{\phi}\ \text{PRED}) = \text{faire-Verb}^*\ \text{GFL}_{1'}$
$({}_{g'}\uparrow_{\phi}\ \text{CO-PRED}) =_c \text{Verb}^*$
${}_{g'}\uparrow_{\phi} = {}_{g}\uparrow_{\phi} \setminus \{\ \text{PRED, TRANS, UNACC, } \alpha, \sigma\ \}$
${}_{g'}\uparrow_{\phi} = {}_{h}\uparrow_{\phi} \setminus \{\ \text{PRED, CO–PRED, TRANS, UNACC, } \alpha, \sigma\ \}$
${}_{g'}\uparrow_{\mu} = {}_{g}\uparrow_{\mu}$

V_2: $({}_{h}\uparrow_{\phi\sigma}\ \text{REL}) = \text{Rel}^*$
$({}_{h}\uparrow_{\phi\alpha}\ \text{R}) = \text{R}^*\ \langle\ \theta_1 \mid \text{TL}\rangle$
$({}_{h}\uparrow_{\phi}\ \text{PRED}) = \text{Verb}^*\ \text{GFL}_2$
$({}_{h}\uparrow_{\mu}) = M^*$

$V_{2'}$: $({}_{h'}\uparrow_{\phi}\ \text{CO-PRED}) = \text{Verb}^*$
${}_{h'}\uparrow_{\mu} = {}_{h}\uparrow_{\mu}$

How does the lexical rule approach in (29) improve over the syntactic approach sketched in (28)? The basic facts discussed in Section 4.1 carry over to (29) without modification. Yet we make two further assumptions: First, morphosyntactic lexical rules are applied *after* complex predicate formation *CPF* has taken place. Secondly, we assume that application of morphosyntactic rules (passive, reflexivization, PP–agr) and auxiliary selection to a *fully merged* argument structure assigns μ–features to $\uparrow_{\mu}$, while in a *partially merged* argument structure μ–features are assigned to $\uparrow_{\mu}$ if the lexical rule applies to the higher level of argument structure, and to $\uparrow_{\mu}$ DEP if the rule applies to the incorporated argument structure.

For the problematic cases of Italian restructuring verbs in (12), where the complex argument structure is projected from the embedded verb, this yields assignment of REFL to $\uparrow_{\mu}$, which predicts clitic climbing. Similarly, auxiliary selection with restructuring verbs will induce AUX–marking at the level of $\uparrow_{\mu}$. Embedded *si*–reflexivization in Italian causative constructions is ruled out since reflexivization is constrained to apply *after* complex predicate formation.

Finally, attachment of object clitics to the higher verb in complex constructions is captured by assigning the subcategorizing (PRED–bearing) verb ($V_{1'}$ in (29)) μ–features for (optional) cliticization:[37] $\{(\uparrow_{\mu}\ \text{OBJ–CL CASE}) =_c \text{acc} \mid \neg(\uparrow_{\mu}\ \text{DEP}^*\ \text{OBJ–CL})\}$. Since cliticization is in no way dependent on *a*–structure, these equations will always be assigned to the local μ–projection. The constraining equation can only be satisfied by cliticization to the higher verb.

Summary

We discussed well–known contrasts of French and Italian complex predicate constructions that are not captured by current LFG analyses. Rosen's(1989) distinction between full and partial merging of argument structures accounts for the basic facts considered. Our main concern was to investigate how this idea can be put to use in the LFG framework. It turned out that the syntactic approach to argument composition meets difficulties not only with regard to computational complexity, but also makes incorrect predictions. We proposed a lexical rule account to complex predicate formation that is implementationally simpler than composition in syntax and makes empirically correct predictions. Yet, complex predicate formation by lexical rule requires a more elaborate theory of lexical rules and lexicon organization, as well as an efficient architecture for the lexicon–to–syntax interface to reduce lexical redundancy.

[37] See Miller(1991), Abeillé et al(1996), who argue for an analysis of clitics as morphological affixes.

References

[1] Abeillé, A./ Godard, D./ Miller, P./ Sag, I. 1996: "French Bounded Dependencies", Ms.

[2] Alsina, A. 1993: *Predicate Composition. A Theory of Syntactic Function Alternations.* Doctoral Dissertation, Stanford University.

[3] Baker, M. 1983: "Objects, Themes and Lexical Rules in Italian" in: Levin, L./ Rappaport, M./ Zaenen, A. (eds): *Papers in Lexical–Functional Grammar*, 1–45.

[4] Bresnan, J. 1990: "Monotonicity and the Theory of Relation Changes in LFG." Ms.

[5] Burzio, L. 1986: *Italian Syntax. A Government–Binding Approach.* Reidel, Dordrecht.

[6] Butt, M. 1995a: *The Structure of Complex Predicates in Urdu.* CSLI Publications, Stanford.

[7] Butt, M. 1995b: "Clause Union – then and now", Ms. University of Tübingen.

[8] Butt, M./ Fortmann, Ch./ Rohrer, Ch. 1996: "Syntactic Analyses for Parallel Grammars: Auxiliaries and Genitive NPs", in: Proceedings of *COLING 96.*

[9] Comrie, B. 1976: "The Syntax of Causative Constructions: Cross–Language Similarities and Divergences", in: *Syntax and Semantics 6*, 261–312.

[10] Dalrymple, M. 1990: *Syntactic Constraints on Anaphoric Binding.* Doctoral Dissertation, Stanford University. CSLI Publications. CSLI Lecture Notes No.36.

[11] Dalrymple, M./ Halvorsen, K./ Kaplan, R./ Manning, C./ Maxwell, J./ Wedekind, J./ Zaenen, A. 1992: "Relating Projections", ms, Xerox PARC, Palo Alto.

[12] Frank, A. 1989: "Argumentstruktur, grammatische Relationen und grammatische Regeln", in: Rolshoven, J./ Seelbach, D. (Hrsg): *Romanistische Computerlinguistik*, Linguistische Arbeiten 266, Tübingen, 19–75.

[13] Frank, A. 1990: *Eine LFG–Analyse zur Kongruenz des Participe Passé im Kontext von Auxiliarselektion, Reflexivierung, Subjektinversion und Clitic–Climbing in kohärenten Konstruktionen.* Master's thesis, University of Stuttgart.

[14] Grimshaw, J. 1982: *On the Lexical Representation of Romance Reflexive Clitics.* in: Bresnan J. / Kaplan, R. M. (eds): The Mental Representation of Grammatical Relations, 87–148.

[15] Kaplan, R./ Bresnan, J. 1995: "Lexical–Functional Grammar: A Formal System for Grammatical Representation", in: Dalrymple, M./ Kaplan, R./ Maxwell III, J.T./ Zaenen, A. (eds.): *Formal Issues in Lexical Functional Grammar*, CSLI Lecture Notes No.47, Stanford.

[16] Kaplan, R./ Maxwell, J.T. 1988: "Constituent coordination in Lexical-Functional Grammar', in: *Proceedings of COLING'88*, Budapest, 303–305.

[17] Kaplan, R./ Wedekind, J. 1993: "Restriction and correspondance–based translation. in: *Proceedings of the EACL 6*, Utrecht.

[18] Kayne, R.S. 1985: "L'accord du participe passé en français et en italien", in: *Modèles linguistiques*, 7, 73–89.

[19] Kayne, R.S. 1989: "Facets of Romance Past Participle Agreement", in: Paola Benincà (ed): *Dialect Variation and the Theory of Grammar.* Proceedings of the GLOW Workshop in Venice, 1987, 85–103.

[20] King, T.H. 1993: *Configuring Topic and Focus in Russian*, Ph.D. dissertation, Stanford University. CSLI Publications.

[21] Kroeger, P. 1991: *Phrase Structure and Grammatical Relations in Tagalog.* Ph.D. dissertation, Stanford University. CSLI Publications.

[22] Manning, C.D. 1992: "Romance is so complex", Technical Report, CSLI.

[23] Manning, C.D. 1996: "Romance complex predicates: phrasal and functional structure", Ms, University of Syndey.

[24] Miller, P. 1991: *Clitics and Constituent Structure*, Doctoral Dissertation, University of Utrecht.

[25] Monachesi, P. 1993: "Object clitics and clitic climbing in Italian HPSG Grammar", in: Proceedings of the Sixth Conference of the EACL, Utrecht.

[26] Rizzi, L. 1982: *Issues in Italian Syntax.* Foris, Dordrecht.

[27] Rosen, S.T. 1989: *Argument Structure and Complex Predicates.* Doctoral Dissertation, Brandeis University.

[28] Sells, P. / Zaenen, A. / Zec, D. 1987: "Reflexivization Variation: Relations between Syntax, Semantics, and Lexical Structure." in: Iida, M. / Wechsler, S. / Zec, D. (eds): *Working Papers in Grammatical Theory and Discourse Structure.* CSLI Lecture Notes 11, 169–238.

[29] Waite, J. J. 1986: *Grammaire Lexicale Fonctionnelle. Dislocation, inversion stylistique et constructions impersonnelles.* Thèse, Montreal.

[30] Zubizarreta, M.L. 1985: "The Relation between Morphophonology and Morphosyntax: The Case of Romance Causatives", in: *Linguistic Inquiry 16*, 247–289.