A unified analysis of French causatives

Harry J. Tily
Department of Linguistics, Stanford University

Ivan A. Sag
Department of Linguistics, Stanford University

Proceedings of the 13th International Conference on
Head-Driven Phrase Structure Grammar

Linguistic Modelling Laboratory,
Institute for Parallel Processing,
Bulgarian Academy of Sciences,
Sofia,
Held in Varna

Stefan Müller (Editor)

2006

Stanford, CA: CSLI Publications

pages 339–359

Tily, Harry J. & Ivan A. Sag. 2006. A unified analysis of French causatives. In Stefan Müller (ed.), *Proceedings of the 13th International Conference on Head-Driven Phrase Structure Grammar, Varna*, 339–359. Stanford, CA: CSLI Publications. DOI: 10.21248/hpsg.2006.19.

Abstract

The treatment of French causatives and pronominal affixes outlined in Miller and Sag (1997) and Abeillé et al. (1998) is notable for its comprehensive coverage and analytic detail, but it relies on a number of *ad hoc* features and types that have little empirical justification. We sketch a new treatment of the same data set, which eliminates multiple lexical entries for the causative, as well as a number of other undesirable analytic devices. Our account builds on a long-standing observation that seeming irregularities in the system of case assignment to the "causee" of *faire* are not in fact exceptional, but determined by the general case assignment behavior of transitive verbs. This generalization, first incorporated into an HPSG analysis by Bratt (1990), was abandoned in subsequent HPSG work that sought to expand the coverage of French beyond that of Bratt's analysis. Our goal here is to show that broad coverage need not come at the expense of linguistically significant generalizations.

1 Introduction

1.1 The composition causative

The verb *faire* is the canonical French causative, exemplified by the following sentences from Miller (1991) and Abeillé et al. (1998).[1]

(1) a. Pierre fait écouter Jean à Marie
Pierre makes to.listen Jean.A Marie.D
'Pierre makes Marie listen to Jean'

b. Paul le-fera lire aux élèves de terminale
Paul it.A-will.make to.read the senior year students.D
'Paul will make the senior year students read it'

c. la chaleur a fait s'évanouir Paul
the heat.N has made SE.to.faint Paul.A
'The heat made Paul faint'

Within the lexicalist literature, a recent and successful trend in the analysis of French complex predicates has suggested that much of the internal structure assumed for (e.g.) English complex VPs is unjustified for French (Miller, 1991; Abeillé et al., 1998; Abeillé and Godard, 2000, 2002). In particular, certain verbs

[†]Thanks are due to Marie Catherine de Marneffe and Frédérique Passot for judgments; to Frank Richter, François Mouret, Olivier Bonami, Stefan Müller and others at the HPSG06 conference for their input; and especially to Danièle Godard and Anne Abeillé for generously giving their time and expertise to guide our analysis towards its final revision.

[1]In the glosses throughout, .N, .A and .D are used to differentiate between the nominative, accusative, and dative (à_1 in Miller's terms) arguments of a verb. We make no claims about the status of case in French. SE indicates a reflexive pronominal in the family *me, te, se...*

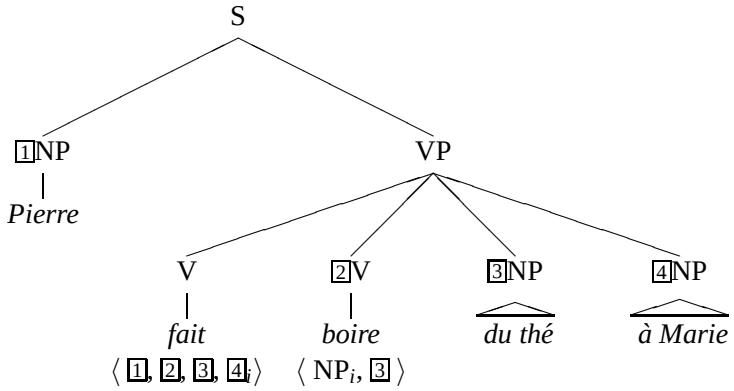


Figure 1: *'Pierre makes Marie drink tea'*

which appear to subcategorize for a dependent verb phrase have been successfully analysed as consisting of only a single VP, in which the embedded verb and all of its complements are treated as complements of the upstairs verb. This analysis applies to the causatives as well as certain other verbs, notably the tense auxiliaries. This analysis is achieved via the technique of "argument composition" (Hinrichs and Nakazawa, 1990), and results in a structure where the two verbs and all of their complements are sisters. An illustration is given in Figure 1.

One of the chief pieces of evidence for the flat VP is the placement of pronominal affixes,[2] which always appear on the finite verb, even when they are arguments of the subordinate verb. This can be seen in (1b), in which *le* ('it') is associated semantically with the downstairs verb *lire* but morphologically with the upstairs verb *fera* – it has "climbed". However, there is one exception: a class of pronominal affixes we will call *intrinsic* affixes, as well as the *se* family of reflexives, fail to climb onto the causative (although they do climb onto tense auxiliaries):

(2) a. La crainte du scandale a fait se-tuer le juge
the fear of.the scandal has made SE-to.kill the judge.A
'The fear of scandal made the judge kill himself/herself'

b. Cette décision fera en-vouloir à tout le monde à Jean
that decision will.make EN-to.want at everybody Jean.D
'Such a decision will make Jean angry at everybody'

The intrinsics are affixes that are lexically/idiomatically associated with a verb and carry no reference. For example, the verb *en vouloir* ('to get angry at someone') has an associated affix *en* which is identical in form to the general purpose pronominal

[2]Romance grammarians have often taken these dependent pronouns to be clitics. This has led to a terminological difficulty for modern lexicalist accounts, which follow Miller (1991), who argues at length that the "clitics" are in fact affixes by the criteria of Zwicky and Pullum (1983). We also follow Miller here and consistently use the term "affix", rather than "clitic".

en ('of them'), but does not contribute any independent meaning to the VP. In (2b), *en* does not climb onto the causative *fera*. Moreover, when any one intrinsic or reflexive is present on the subcategorized verb, *all* other affixal arguments of that verb must also be realized locally: they are "trapped". For example, in (3), the affix *en* is a regular indirect argument of the subcategorized verb which would usually climb and be realized non-locally:

(3) Marie a fait s'en-souvenir Jean
Marie has made SE.EN-to.remember Jean
'Marie made Jean remember it'

The presence of the reflexive *se*, however, traps it on the subcategorized verb.

A further subtlety in the behavior of composition causative verbs is the case that they assign their 'causee' argument. Curiously, the case marking of the causee seems to be dependent on properties of the embedded verb. Given an intransitive infinitive as complement, *faire* assigns accusative case to its causee; given a transitive infinitive, it assigns dative case:[3]

(4) a. Le prof fait lire l'élève
the teacher makes to.read the student.A
'The teacher makes the student read'

b. Le prof fait lire Proust à l'élève
the teacher makes to.read Proust the student.D
'The teacher makes ('to') the student read Proust'

This is equally true when the causee is realized as a pronominal affix, as it is in (5):

(5) a. Le prof le-fait lire
the teacher him.A-makes to.read
'The teacher makes him read'

b. Le prof lui-fait lire Proust
the teacher him.D-makes to.read Proust
'The teacher makes 'to' him read Proust'

However, certain verbs resist this generalization. In (6a), the subcategorized verb realizes no direct object, and yet the causee is dative. Likewise, in (6b) the causee is dative:

(6) a. Le prof lui/∗le-fait voir / comprendre
The teacher him.D/∗.A-makes to.see / to.understand
'The teacher makes him see / understand'

[3]French does not have a strong morphological case marking system. Subjects, objects and indirect objects are distinguished by differing (but syncretic) morphological forms when they are realized as pronominal affixes. When realized syntactically (as an NP), subjects and direct objects are not distinguished, but indirect objects appear with the dummy prepositional marker *à* (for a broadly compatible treatment, see Abeillé et al. (2005)).

b. Son chef lui/∗le-fait en-vouloir à tout le monde
her boss her.D/∗.A-makes EN-to.want at everyone
'Her boss makes her angry at everyone'

c. Il.fait se.les-laver aux/∗les enfants
he.makes SE.them.A-to.wash the children.D/∗.A
'He makes the children wash them (their hands)'

Even if we consider the intrinsic affix *en* to be a direct object, it is realized downstairs before argument composition occurs. Therefore, no unsaturated argument of *vouloir* can be visible when *faire* selects it, and so we would expect the intransitive behavior. The same problem is illustrated in (6c): the direct object of the downstairs verb is "trapped" on the subcategorized infinitive because of the presence of the reflexive *se*, and yet the causee case marking remains dative. If affixal realization suppresses an argument, as all analyses of which we are aware suggest, then one would expect an infinitive that has realized its direct object to pattern as an intransitive. Given these exceptions, it does not seem possible to rely on the simple generalization that the transitivity of the subcategorized verb determines the causee's case marking.

1.2 Miller and Sag 1997

In the first section, we listed three important facts that an analysis of the French composition causative should incorporate:

- the causative verb must compose the arguments of its subcategorized verb to give rise to the flat VP;
- pronominal affixes associated with the subcategorized verb must climb onto the causative, except where any among them is intrinsic or reflexive;
- the causee argument must be assigned the appropriate case, taking into account the transitivity generalization and its apparent exceptions.

The analysis presented by Miller and Sag (1997) (henceforth MS97) is the most comprehensive account of the causative that we are aware of, and will serve as our starting point.[4] Hence, we will assume a degree of familiarity with this analysis, including its type hierarchy (which will be similar to our own) and its basic treatment of morphology.

[4]A number of other authors have presented analyses influenced by Miller and Sag. Notable among them are Calcagno and Pollard (1999), who consider a broader range of causatives than we discuss here, and focus on providing a more elaborate and nuanced theory of argument realization and structural case, but not the details of pronominal placement. Crysmann (2003) attempts both to eliminate the type *aff* and to remove ad-hoc book-keeping features (an issue that we also address), while providing a uniform treatment for Italian. However, Crysmann's analysis does not address the issue of having multiple lexical entries for the causative.

MS97 succeeds in accounting for the above facts, though it does so at a cost – it posits additional "bookkeeping" features and types to capture the empirical data, rather than finding a parsimonious generalization. For example, MS97 subtypes words into *clitic-words* which have realized pronominal affixes, and *plain-words* which have not, despite there being little evidence that this distinction is ever selected for. It also imposes an almost equivalent distinction between *basic-verb* and *reduced-verb* to identify those verbs that have suppressed arguments by realizing them affixally. However, to state the selection restrictions of the causative and capture the trapping effect described above, it is necessary to assume that verbs that have realized arguments affixally as well as having intrinsic affixes are *basic-verbs* and not *reduced-verbs*. This stipulation reduces the distinction to an *ad hoc* descriptive solution.

Another expedient but undesirable device relied on by MS97 is the binary feature TRANS, used to stipulate the transitivity of a verb. In order to account for the apparent failure of some intransitives to respect the causee case marking generalization described above, transitivity is stipulated on a verb-by-verb basis, ignoring the actual length of the argument structure list. This is linguistically unnatural, reducing the notion of transitivity to an arbitrary distinction unrelated to ARGUMENT-STRUCTURE length. Moreover, in order to make this work, one has to posit two lexical entries for composition causative *faire*: one which selects for a [TRANS +] verb and assigns its causee dative case; and another which selects for a [TRANS −] verb and assigns its causee accusative case.

In this paper, we present a treatment of the composition causative *faire* based on Miller and Sag (1997) that equally well captures the facts described above, but which dispenses with the feature TRANS, the types *basic-verb*, *reduced-verb*, *plain-word* and *clitic-word* (though we instead make use of a book-keeping feature comparable to the former dichotomy), and the need for multiple lexical entries for *faire*. Instead, we suggest a principled lexical semantic basis for the phenomenon of differential causee case, and from this we derive a more parsimonious treatment.

2 Re-evaluating the data

As the issue of causee case marking is the least well treated in the existing literature, we will focus on it here. The basic pattern is shown in (4), repeated here as (7):

(7) a. Le prof fait lire l'élève
the teacher makes to.read the student.A
'The teacher makes the student read'

b. Le prof fait lire Proust à l'élève
the teacher makes to.read Proust the student.D
'The teacher makes 'to' the student read Proust'

$$\left[\text{ARG-ST} \quad \langle\, \text{NP}_i \,\rangle \oplus \boxed{A} \oplus \langle\, \text{NP}_j,\ \text{V}\left[\text{ARG-ST} \quad \langle\, \text{NP}_j \,\rangle \oplus \boxed{A}\right] \rangle\right]$$

Figure 2: Simplified lexical entry for composition causative *faire*

Given an argument composition analysis of the complex VP as described above, a natural way to capture the causee case-marking facts is given by Bratt (1990). First, we assume that the first object on every verb's argument structure is constrained to be accusative, and any further objects are required to be dative or oblique. Then, achieving the correct case-marking facts is simply a matter of performing argument composition in a novel way: rather than appending the subcategorized verb's arguments to the end of *faire*'s argument structure, we insert them before the causee. This constraint is schematized in Figure 2.

Now, when the subcategorized verb is transitive, its list of objects ($\boxed{A}$) is non-empty, and so the causee (NP_j) is not *faire*'s first object, and receives dative case. Only when the subcategorized verb is intransitive is $\boxed{A}$ empty, in which case the causee ends up the first object, receiving accusative case.[5] Thus the causee's case falls out naturally from the observation that French verbs take at most one accusative object (henceforth *Bratt's generalization*).

Unfortunately, this simple treatment does not deal with irregularities like (6a), repeated here as (8):

(8) Le prof lui/∗le-fait voir / comprendre
The teacher him.D/∗.A-makes to.see / to.understand
'The teacher makes him see / understand'

MS97 deals with the irregularities in the data through stipulation: although most intransitive verbs bear the value [TRANS −], certain verbs (such as the use of *voir* in (8) above) are lexically specified to be [TRANS +].

An alternative approach that would enable us to preserve Bratt's generalization would be to suppose that there is an invisible ("null") direct object on the argument structure of those seemingly intransitive verbs that pattern like transitives. This null object, indicated *pro*,[6] will be inherited as *faire*'s direct object and result in dative marking on the causee, as sketched in Figure 2.

At first glance, positing a null argument seems no less *ad hoc* than the feature TRANS. However, there does appear to be some linguistic justification for the presence of null arguments in French. The phenomenon of *null instantiation* has been studied in some depth (Fillmore, 1986): certain verbal arguments may be omitted according to verb-specific lexical licensing restrictions, and when appropriate

[5] In this simple sketch, we ignore the possibility of non-object items on the inherited ARG-ST. The final analysis given later resolves this issue.

[6] Here we assume that *pro* is a phonologically null subtype of *sign*, choosing terminology familiar from the analysis of unexpressed subjects. In work in preparation, Fillmore, Kay and Michaelis and Sag flesh out a typology of unexpressed arguments compatible with the analysis here. We assume that the daughters list of a phrasal construction may include any number of *pros*; thus they are "silently saturated".

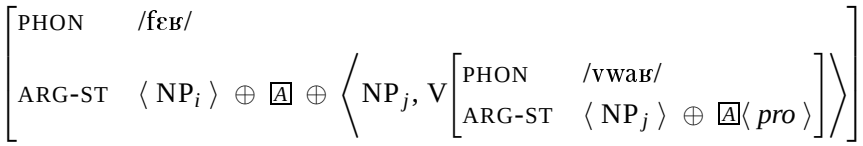


Figure 3: Deriving dative causee case marking given intransitive *voir*

pragmatic conditions are met. Lambrecht and Lemoine (2005) provide a typology of null instantiation for French based on Fillmore's work on English.

In Lambrecht and Lemoine's classification, and following Fillmore's, *indefinite null instantiation* (INI) refers to cases where the specific identity of the missing object is not and cannot be inferred from the context by the speaker. Such instances impose a "generic" interpretation of the missing argument. *Definite null instantiation* (DNI), on the other hand, is more closely related to anaphora. Missing objects whose specific referent is readily identifiable in context are classified as instances of DNI. The sentences in (9) illustrate INI, while (10) illustrates DNI:[7]

(9) a. Maman est occupée; elle.coud
mother is busy; she.sews
'Mother is busy; she is sewing'

b. Il-a encore bu
he-has again drunk
'He drank again'

(10) Je-jouais du piano. Puis nous-avons éteint
I-played piano. Then we-have turned.out
'I played piano. Then we turned out (the lights)'

We can compare this behavior with the problematic example given in (8). It appears that the intransitive use of *voir* which leads to dative case marking is an instance of (or is at least closely related to) DNI. The argument cannot receive a generic interpretation: there must be some appropriate referent that is seen for the sentence to be felicitous. Therefore, we claim that DNI missing objects are in fact present as *pro*s on the argument structure, though INI objects are truly absent.

Lambrecht and Lemoine (2005) also discuss the following contrast in acceptability between verbs with an INI object depending on their aspectual class:

(11) a. Une fois sortie de la forêt, on-voyait/#on-a vu de nouveau
once left from the forest one-saw/#one-has seen anew
'Once you were out of the forest, you could see again'

[7]These data are taken from Lambrecht and Lemoine (2005). It should be noted that these sentences were produced in natural casual speech, but are subject to quite specific contextual and pragmatic licensing and so may seem less acceptable out of context or in written form.

b. Deux heures plus tard, je-mangeais/#j'ai mangé de nouveau
two hours later, I-ate/#I.have eaten anew
'Two hours later, I was eating again'

The examples in (11) are intended as a continuation of a passage describing a situation where the speaker could temporarily not see, or was so full that they could not eat.

Both of these examples are well formed in the imperfect ('I was eating') but not the perfect ('I ate') construction. Lambrecht and Lemoine interpret this in terms of the inability of a definite referent to be the direct object of these verbs, as in these contexts, they are coerced from their default aspectual classes into a stative reading expressing the property of being able to see or eat:

> ... the perfect form *on a vu* 'you saw' would necessarily be interpreted as evoking a definite object referent (e.g. *on a vu ce qui s'était passé* 'you saw what (had) happened'). Likewise, ... the perfect form *j'ai mangé* would evoke the idea of a meal rather than some undefined edible thing. —Lambrecht and Lemoine (2005)

If this is the case, then we need not treat the fact that verbs like *voir* fail to obey the pattern of the transitive infinitive as an arbitrary lexical stipulation. Rather, it is the interpretation of the (missing) argument itself which is relevant, and the identity of the verb is only relevant insofar as it constrains the realization potential of that argument. In fact, we find occurrences of a causative combining with intransitive *voir* that *do* give rise to an accusative causee; such uses are exactly those where the interpretation of the argument carries a generic rather than specific reference:

(12) Jésus fait voir les aveugles
Jesus makes to.see the blind.A
'Jesus makes the blind see' (become able to see)

3 Analysis

3.1 The construction-based grammar

Following Sag (to appear) and Fillmore et al. (ms.), we model constructs as feature structures of the form sketched in Figure 4:[8] The immediate subtypes of *construct* are *lexical-construct* (*lex-cxt*) and *phrasal-construct* (*phr-cxt*), which form the top of the hierarchy of construct types sketched in Figure 5:[9]

[8]Note that we write type constraints as AVMs where the type specification is followed by an implication arrow, indicating that all objects of the specified type or its subtypes must obey the constraints given. We deliberately avoid writing the type name outside the AVM, as such notation is associated with the more expressive implications associated with RSRL-style HPSGs (Richter, 2004), in which an arbitrary feature structure description may be provided as the antecedent.

[9]Here *derv-cxt* abbreviates *derivational-construct*; *infl-cxt*, *inflectional-construct*; *pinfl-cxt*, *postinflectional-construct*; and *lex-item*, *lexical item*.

$$\begin{bmatrix} \textit{construct} \Rightarrow & \\ \text{MTR} & \textit{sign} \\ \text{DTRS} & \textit{list}(\textit{sign}) \end{bmatrix}$$

Figure 4: Type constraint on *constructs*

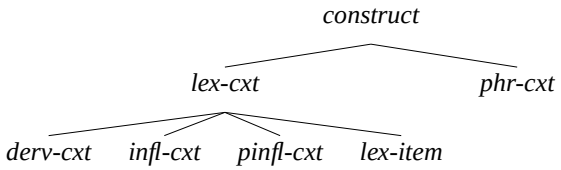


Figure 5: The construct type hierarchy

What then is a construction? According to Fillmore et al. (ms.), a construction is a constraint defining the properties that are common to all members of a family of constructs. That is, a construction is a constraint of the form shown in Figure 6, where *x-cxt* is the name of some construct type, i.e. some subtype of the type *construct*. Each construction licenses a grammatically distinctive class of constructs.

$$\begin{bmatrix} \textit{x-cxt} \\ \ldots \end{bmatrix}$$

Figure 6: A construction

Even lexical items, since they too are constructs, have a MTR and DTRS value. Lexical items are subject to a constraint requiring the DTRS value to be the empty list, which means that lexical entries will license constructs like the one in Figure 7. On this view of things, lexical entries are also constructions. That is, a lexical entry is a constraint that defines a class of lexical items. Larger signs are "constructed" from lexical items via lexical and phrasal constructions.

Much of the motivation for a construction-based analysis in HPSG has to do with delimiting the locality of selection (Sag, to appear). For this reason, the valency geometry is slightly different from Pollard and Sag (1994). In particular, the feature VAL(ENCY) is a list (of signs!) that contains all of a sign's valents that remain to be saturated, and EXTERNAL-ARGUMENT (X-ARG) is a list containing at most one privileged member of VAL (e.g. the subject of a verb).

3.2 Pronominal affixes

Rather than segregating pronominal affixes into pronominal (*p-aff*) and anaphoric (*a-aff*) types as Miller and Sag do, we instead introduce a binary feature INTRIN(SIC) on objects of type *aff*. Intrinsic affixes and reflexives carry the value [INTRIN +],

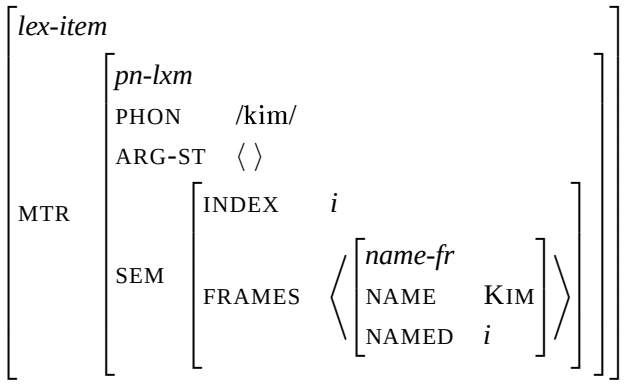


Figure 7: A lexical item

while all other affixes carry the value [INTRIN −].

In our proposal, a definite null realization of a verb's argument corresponds to the presence of a *pro* on argument structure. We claim that arguments suppressed by affixal realization have the same syntactic status as DNI arguments, and we formulate our treatment of affixal realization so as to ensure that these *pros* are present when an affix is realized. We do this via lexical constructions which remove an affixal element from ARG-ST (suppressing the argument) and which, in certain cases, insert a *pro* in its place. Rather than realizing the corresponding morphology at the moment that the argument is suppressed, we instead record the presence of an affix to be realized using a list-valued feature, PRAFS (PRONOMINAL-AFFIXES). This allows us to implement all of the morphological operations at a single point, using inflectional constructions to be described later.[10]

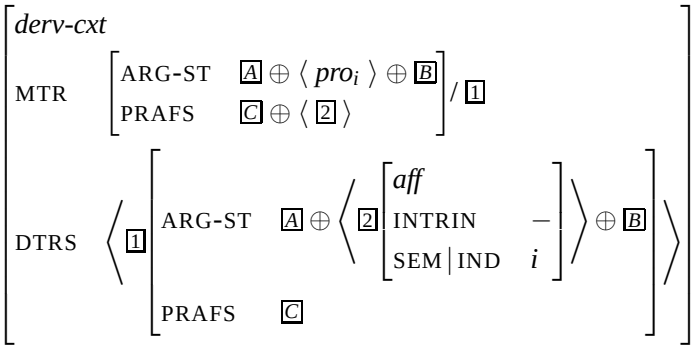


Figure 8: Nonintrinsic pronominal affixation construction

Each of the two lexical constructions in Figures 8 and 9 suppresses a single *aff*

[10] In an earlier version of this paper presented in Varna in summer 2006, we attempted to define a single construction which replaced all the relevant *affs* with *pros* in one step. However, its formulation went beyond the descriptive power of the constraint logic we assume here, and we consider the introduction of the feature PRAFS and addition of an extra construction preferable to extending the mathematical basis of the theory.

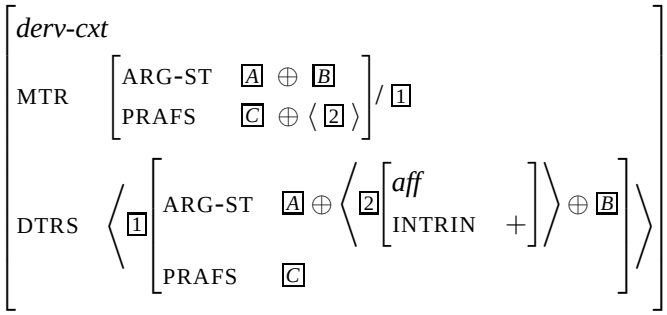


Figure 9: Intrinsic pronominal affixation construction

argument, and records the identity of that affix in the PRAFS list. Figure 8 only applies to lexemes with a [INTRIN −] affixal argument. The affix is removed from the argument structure, but a *pro* is inserted in its place. The second construction, in Figure 9, instead removes intrinsic affixes, and does not replace the removed argument. Hence, we ensure that verbs realizing a reflexive direct object have the same valency as intransitives, correctly predicting the causee's case in sentences like the following:

(13) Paul fait se.raser Figaro
Paul makes SE.shave Figaro-ACC
'Paul makes Figaro shave himself'

"Clitic climbing" and "clitic trapping" are discussed below.

The feature PRAFS performs a "bookkeeping" function – it records the feature structure's progress through a multi-step operation. It is comparable to the CLTS feature of Monachesi (1999). However, because we make PRAFS a feature of *lexeme* and not *word*, the distinction between units which have and which have not suppressed *aff* arguments is visible only to the morphology, and not to the syntax. This prevents a syntactic constraint from selecting directly for a word with certain affixes, which remains a technically available, but unattested possibility for Monachesi's grammar.

3.3 Inflectional constructions

In the construction-based grammar, *lexemes* are promoted to *words* by an *inflectional-cxt*. Subtypes of this construct correspond to the different parts of speech; verbs are handled by constructs of type *verb-infl-cxt*. There are a large number of constructions describing *verb-infl-cxt*s: one for each verb inflection class. However, as pronominal affixes are realized in basically the same way, regardless of the affixal ending, we can declare the necessity to realize pronominal affixes just once, as a constraint on all such constructions, as sketched in Figure 10.

The function `affix` takes three arguments: the syntactic category of the host, the (inflected) morphological form of the host, and and a list of pronominals to be affixed. The constraint in Figure 10 leaves the morphological form unspecified,

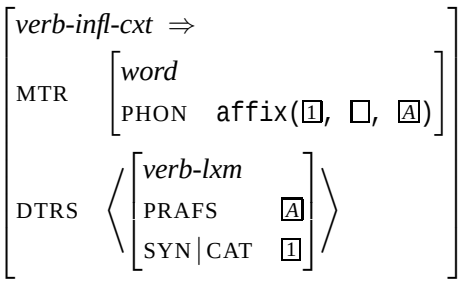


Figure 10: Type constraint on *verb-infl-cxt*s

leaving it to be filled in by the actual inflectional constructions. Space limitations preclude a detailed discussion of the morphological procedure that is implemented by `affix`, but this function behaves much like Miller and Sag's F_{PRAF}, a fairly trivial mapping between inflected verb forms and the fully affixed forms.

One property of F_{PRAF} crucial to the MS97 analysis is the requirement that the morphological realization of any affix on a past participle be null – French past participles can never host pronominal affixes. In MS97, this is guaranteed by a statement that F_{PRAF} is the identity function when given a past participle as argument, even if that participle has affixal arguments. The necessity for this stems from assumptions about structure sharing: in the case of auxiliary-participle constructions (unlike other flat complex VP constructions), *aff*s inherited from the participle appear on the argument structure of both the auxiliary and the subcategorized participle. As MS97 also states that an *aff* is always realized on the word in whose argument structure it appears, it should predict that an affixal argument of a participle is realized twice, on both the participle and the auxiliary.

In our analysis, by contrast, *aff* arguments and the morphological realization corresponding to them have complementary distribution. Only when a lexical construction has moved the *aff* to the PRAFS list will it be realized. We can therefore do without the stipulation that past participles realize their affixes covertly. Instead, we state that `inflect` is only a partial function, having no resolution given a past participle and any list of affixes other than the empty list. This is in fact a significant improvement: since the application of F_{PRAF} to any past participle yields a valid (but unaffixed) form, the MS97 account wrongly predicts that participles used outside of tense auxiliary constructions (for example, as noun modifiers) may have affixal arguments which are *never* realized.[11]

The initial value of PRAFS on all lexemes licensed directly by a lexical item is the empty list. This is simply to say that any affixes that end up realized must first

[11]The problems with MS97's definition of F_{PRAF} are even more striking when similar phenomena are considered cross-linguistically. Italian, for instance, does not prohibit affixation on past participles in all cases; in fact, it only prohibits them in auxiliary constructions (Monachesi, 1999). If we relax for Italian the statement that `inflect` cannot be resolved to an affixed past participle, our analysis goes part way towards predicting the Italian data without generating the ungrammatical "double realization" VPs that MS97 must avoid by stipulation. We leave a fuller investigation of the application of this approach to Italian for subsequent research.

have been introduced on the argument structure. We specify this with the constraint on *lex-items* in Figure 11.

$$\begin{bmatrix} \textit{lex-item} \Rightarrow & \\ \text{MTR} & \begin{bmatrix} \textit{lexeme} & \\ \text{PRAFS} & \textit{elist} \end{bmatrix} \\ \text{DTRS} & \textit{elist} \end{bmatrix}$$

Figure 11: Type constraint on lexical items

An example of an inflectional construction which produces the third-person singular form of a regular *-er* verb is given in Figure 12.

$$\begin{bmatrix} \textit{verb-infl-cxt} & \\ \text{MTR} & \begin{bmatrix} \text{PHON} & \texttt{inflect}(\square, \boxed{2}, \square) \\ \text{SYN} & \begin{bmatrix} \text{X-ARG} & \langle\ \text{NP}[\textit{3sg}]\ \rangle \\ \text{CAT}\,|\,\text{VFORM} & \textit{fin} \end{bmatrix} \\ \text{SEM} & \ldots \end{bmatrix} / \boxed{1} \\ \text{DTRS} & \left\langle \boxed{1}\begin{bmatrix} \text{STEMS}\,|\,\text{SLOT-3} & \boxed{2} \end{bmatrix} \right\rangle \end{bmatrix}$$

Figure 12: A simplified *inflectional-cxt*

We adopt the theory of inflection presented by Bonami and Boyé (this volume), which assumes that morphological information on the lexeme is encoded as a *stem space*, with a feature for each slot in the inflectional paradigm, and where regularity is encoded as constraints on those slots. A fuller exposition of this theory as it applies to French verbs is given in Bonami and Boyé (2006). Our analysis does not depend on this, however, and is compatible with other treatments of morphology.

Most inflections are instantiated by a family of constructions much like Figure 12, and we will not spell out the details of the morphological paradigms here. For infinitives, we require a slightly more constrained construction, because we need to limit the infinitives that causative *faire* can combine with. It is the "trapping" property which is at issue here: we need to ensure that pronominal arguments of the downstairs verb obligatorily climb, unless they are accompanied by an intrinsic affix, in which case they must be realized on the infinitive. This amounts to saying that *faire* selects for either (1) an infinitive that has no intrinsic arguments and realizes no pronominal affixes or (2) an infinitive that has realized all its pronominal affixes, among which is at least one intrinsic.

Related properties of German embedded infinitivals led Bech (1955) to identify two classes of constructions: (1) the coherent constructions, where two verbs (or more) are adjacent and both verbs' arguments appear to function as arguments of the 'higher' verb, and (2) the incoherent constructions, where the verbs display the expected biphrasal behavior. With this precedent, we subtype the VFORM value

inf(*initive*) into two classes: *coh-inf* (*coherent-infinitive*) and *inc-inf* (*incoherent-infinitive*). This typing is shown in Figure 13.

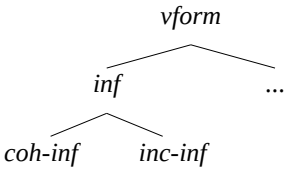


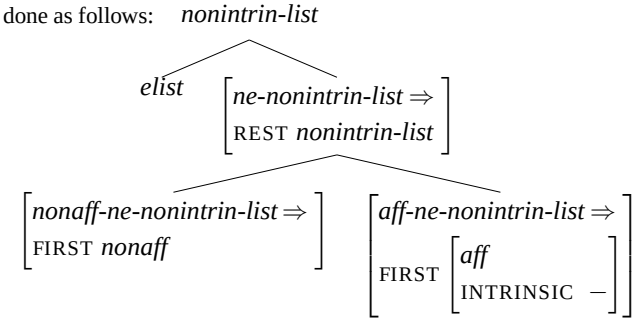
Figure 13: Subtypes of infinitive

Of course, this division is very similar to the distinction drawn in MS97 and in Abeillé et al. (1998) between *basic-verb* and *reduced-verb* – two subtypes of *head*. However, there are clear differences. Crysmann (2003) points out that MS97's type distinction should prohibit coordinations of basic verbs and reduced verbs, which he argues can be perfectly grammatical. Our system is not directly vulnerable to this criticism, since our coherence distinction does not express the actual presence or absence of pronominal affixes, but rather the suitability of the infinitive to be the subcategorized verb in an argument composition cluster. Infinitives that realize pronominals and those that don't may be coordinated, as long as their VFORM values resolve to the same coherence type (the same subtype of *inf*).

With these types in hand, we can proceed to specify inflectional constructions for infinitives that license coherent and incoherent infinitives. We place no constraints on the incoherent infinitives; any verbal lexeme may resolve to license an incoherent infinitive word. However, there are two narrow possibilities for the coherent infinitives: those which realize no affixes and have no intrinsic arguments, and those which realize all their affixes and have intrinsic arguments. This state of affairs is illustrated in Figure 14.

For an infinitive that realizes no pronominal affixes to be coherent, it must have no intrinsic affixes on its argument structure (since these must not be allowed to climb). We specify that the argument structure in this case is a *nonintrin-list*, a subtype of *list* which is guaranteed to contain no *affs* that are [INTRIN +]. This can be effected through the type inheritance system much as standard lists are implemented.[12]

[12]For example, this could be done as follows:



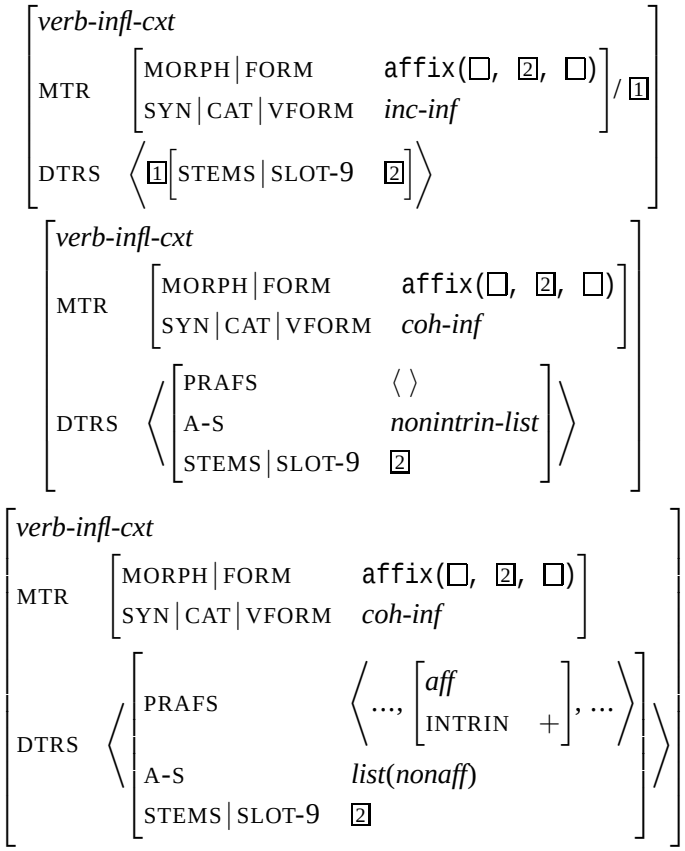


Figure 14: The infinitive inflectional constructions

3.4 Argument composition

In order to implement Bratt's generalization, we rely on a number of general structural properties of the language. First, we assume a standard obliqueness ordering of all ARG-ST lists: subjects precede direct objects, which precede indirect objects, which precede other arguments and then other adjuncts. Next, we capitalize on the fact that French verbs are either intransitive or transitive, but they never have more than a single direct object NP. In Figure 15 we sketch an appropriate simple structural case system: XP^{dir} abbreviates an unmarked direct argument (i.e. a subject or direct object) and XP^{obl} abbreviates a more oblique argument, encompassing indirect objects, prepositional and complement phrases, and predicative NPs.

We tacitly assume a theory of prepositions and oblique argument markers in the spirit of the "weak heads" of Tseng (2002), Abeillé et al. (2005) or Miller (1991). These authors suggest that certain apparent prepositions are in fact not the head of a PP, but something more like a *marker* modifying an NP, in much the same way as the treatment of complementizers given in Pollard and Sag (1994). The precise

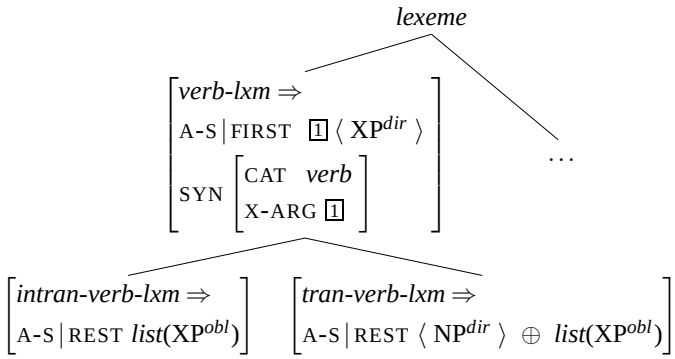


Figure 15: Constraints on lexeme types

details of this are not important for our purposes; all that is necessary is that we be able to underspecify a noun phrase so that it can resolve to be either a direct object or an indirect object – this is the status of the causee argument in our treatment.

Having set the stage in this way, the lexical entries for the complex predicate verbs are quite simple: both the causative (Figure 16) and tense auxiliary (Figure 17) compose the argument structure (A) of a subcategorized verb into their own argument structure. The causative additionally introduces a causee argument which is coindexed with the unexpressed subject of the subcategorized verb (NP_j). This argument is underspecified for case, so by constraints on *tran-verb-lexeme* in Figure 15 it must resolve to be direct (accusative) if it is the first object, or indirect (dative) if not.[13] The causee must resolve to be either a direct (accusative) or indirect (dative) object, and it is placed among the arguments inherited from the subcategorized verb. The subcategorized verb will have assigned appropriate case marking to its own arguments by the same constraints. So, if it is transitive, there will be an accusative object which must resolve as a member of A, meaning that the causee will be non-initial, and resolve to be an indirect object. If the verbal complement is intransitive, then A will resolve to the empty list, as the causee must resolve to either a direct or indirect object; *faire* is a *tran-verb-lxm* and hence requires a direct object. Thus we preserve Bratt's generalization.

It is worth noting that we only mention ARG-ST, and never VALENCE in our constraints. Previous treatments have varied in using valency or argument structure as the locus of composition. MS97 and Abeillé et al. (1998) make use of an argument structure/valency discrepancy to predict the different behavior of tense auxiliaries (which were taken to perform argument composition on ARG-ST) and the composition causative (which composed from the COMPS list). Because we do not rely on this distinction, we can retain the standard argument realization principle and assume that ARG-ST and VAL are identified in the normal case. Thus,

[13] We assume that the causee can be underspecified in such a way that it can resolve to a direct or indirect object, but nothing more oblique. It is possible to define a typing within a theory of structural case typing that allows this, but doing so is well beyond the scope of this paper.

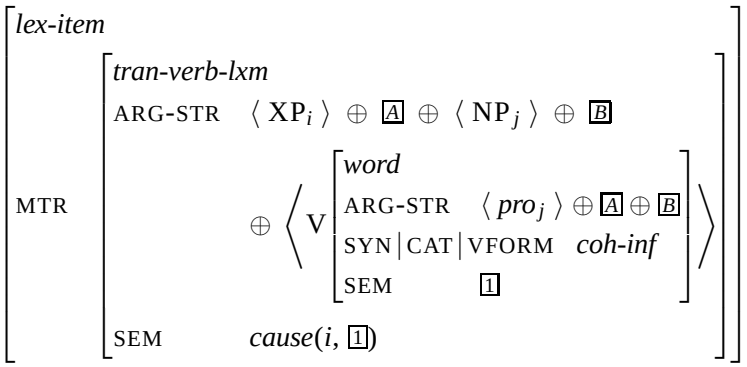


Figure 16: Lexical entry for composition causative *faire*

although we assume here that it is ARG-ST which is relevant, nothing hinges on this.

The entry for the tense auxiliary in Figure 17 simply inherits all of its participial complement's arguments. The possibility that the participle might realize any affixal arguments is ruled out by the morphological function `affix` as discussed in section 3.3.

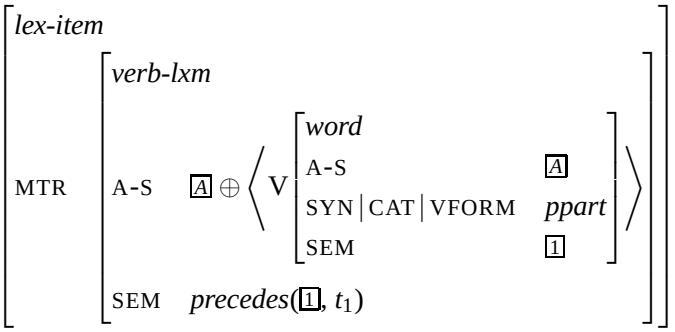


Figure 17: Lexical entry for tense auxiliary *avoir*

4 Summary

Our analysis improves on that of Miller and Sag (1997) in the following ways:

- It specifies a uniform analysis for composition *faire*, without needing multiple lexical entries, and it captures Bratt's generalization.
- It does so so in a principled way, appealing to the notion of null instantiation (Fillmore, 1986; Lambrecht and Lemoine, 2005; Fillmore et al., ms.).
- It eliminates much of Miller and Sag's partitioning of the type hierarchy, including:

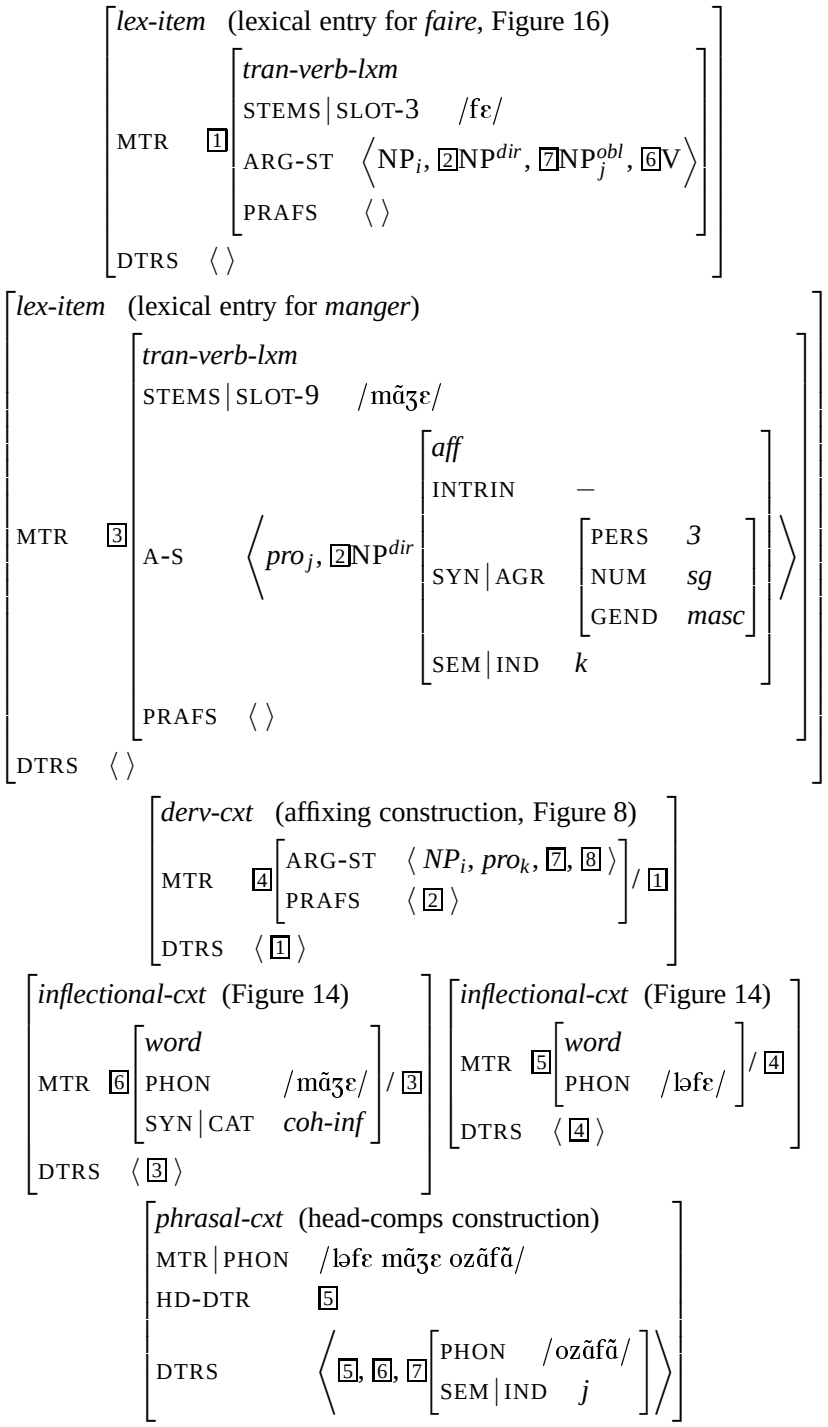


Figure 18: Partial analysis of *le fait manger aux enfants* ('make the children eat it')

  - the distinction between *plain-word* and *clitic-word*, which is now invisible to the syntax;
  - the distinction between *reduced-verb* and *basic-verb* (we retain a similar distinction in *coh-inf/inc-inf*, as discussed below)
  - the distinction between *a-aff* and *p-aff*. (Naturally, the binding theory must still have a way to discriminate between referential and anaphoric pronouns, but it can now be stated as a semantic property.)

- It dispenses with the *ad hoc* feature [TRANS $\pm$], and reverts to an empirical notion of transitivity as determined by argument structure.

The cost of these improvements are the additional features INTRIN and PRAFS, the subtyping *coh-inf/inc-inf*, and a number of new constructions.

The INTRIN feature is justified, as (1) there are nonintrinsic and intrinsic variants of every pronominal other than the reflexives, and (2) the non-existence of nonintrinsic reflexives shows that intrinsic status is a property of pronominals themselves rather than of the verbs that select them. The PRAFS feature and the constructions that move affixes from ARG-ST to PRAFS allow a complex operation (the replacement of a number of *affs* with *pros*) to be stated as several simpler steps. PRAFS is a *lexeme* feature, and so is not available to be selected syntactically. The *coh-inf/inc-inf* distinction is an improvement on MS97's *red-vb/bas-vb* in that (1) it is limited to infinitives, (2) it does not make incorrect predictions about coordination potential, and (3) it is motivated by similar phenomena cross-linguistically (Bech, 1955). Other than the two affixing constructions that process PRAFS, the only new constructions are the *inflectional-cxts* for infinitives that distinguish between coherent and incoherent infinitives. Any grammar must assume at least one *inflectional-cxt* for infinitives, so our net addition to the grammar is minimal.

We have outlined a treatment of pronominal affixes and their behavior under the composition causative and other argument composition verbs which continues an HPSG tradition of providing broad-coverage formal grammars for French. In doing so, however, we have eliminated a number of *ad-hoc* descriptive devices and have replaced them with a more principled linguistic account that appeals to null instantiation and argument composition to derive comparable coverage.

References

Abeillé, A., Godard, D. and Sag, I.A. 1998. Two kinds of composition in French complex predicates. In E. Hinrichs, A. Kathol and T. Nakazawa (eds.), *Complex Predicates in Nonderivational Syntax*, pages 1–41, NY: Academic Press.

Abeillé, Anne, Bonami, Olivier, Godard, Danièle and Tseng, Jesse. 2005. The syntax of *à* and *de*. In *The Linguistic Dimensions of Prepositions*, Dordecht: Academic Press.

Abeillé, Anne and Godard, Danièle. 2000. French word order and lexical weight. In Robert Borsley (ed.), *The nature and function of Syntactic Categories*, volume 32 of *Syntax and Semantics*, pages 325–358, New York: Academic Press.

Abeillé, Anne and Godard, Danièle. 2002. The syntactic structure of French auxiliaries. *Language* 78(3), 404–452.

Bech, Gunnar. 1955. *Studien über das deutsche Verbum infinitum. Det Kongelige Danske videnskabernes selskab*, volume 35 of *Historisk Filosofiske/filologiske Meddelelser*. Copenhagen: Munksgaard.

Bonami, Olivier and Boyé, Gilles. 2006. Remarques sur les bases de la conjugaison, ms Université Paris-Sorbonne & Universit Bordeaux 3.

Bonami, Olivier and Boyé, Gilles. this volume. Deriving inflectional irregularity.

Bratt, Elizabeth O. 1990. The French Causative Construction in HPSG, ms Stanford University.

Calcagno, Mike and Pollard, Carl. 1999. The role of argument structure in the interaction between reflexives and causatives in French. In *Proceedings on Lexical Semantics and Linking in Constraint-Based Theories. Workshop held at the 11th ESSLLI*, Utrecht University.

Crysmann, Berthold. 2003. Clitic climbing revisited. In *Proceedings of HPSG03*, pages 67–89, Stanford, CA: CSLI Publications.

Fillmore, Charles J. 1986. Pragmatically controlled zero anaphora. In *Proceedings of the 12th Berkeley Linguistics Society*, pages 95–107.

Fillmore, Charles J., Kay, Paul, Michaelis, Laura and Sag, Ivan A. ms. Sign-Based Construction Grammar, unpublished ms.: Berkeley, Boulder, and Stanford.

Hinrichs, Erhard and Nakazawa, Tsuneko. 1990. Subcategorization and VP structure in German. In *Proceedings of the 3rd Symposium on Germanic Linguistics*, Amsterdam: Benjamins.

Lambrecht, Knud and Lemoine, Kevin. 2005. Definite Null Objects in (Spoken) French: A Construction Grammar Account. In *Grammatical Constructions: Back to the Roots*, Amsterdam, The Netherlands: Benjamins.

Miller, Philip H. 1991. *Clitics and Constituents in Phrase Structure Grammar*. Ph. D.thesis, Free University of Brussels.

Miller, Philip H. and Sag, Ivan A. 1997. French clitic movement without clitics or movement. *Natural Language and Linguistic Theory* 15, 573–639.

Monachesi, Paola. 1999. *A Lexical Approach to Italian Cliticization*. Stanford, CA: CSLI Publications.

Pollard, Carl J. and Sag, Ivan A. 1994. *Head-Driven Phrase Structure Grammar*. Stanford, CA: CSLI Publications, distributed by University of Chicago Press.

Richter, Frank. 2004. *A Mathematical Formalism for Linguistic Theories with an Application in Head-Driven Phrase Structure Grammar*. Phil. dissertation (2000), Eberhard-Karls-Universität Tübingen.

Sag, Ivan A. to appear. Remarks on Locality. In Detmar Meurers and Robert D. Levine (eds.), *Ohio State Working Papers in Linguistics*, Ohio State University.

Tseng, Jesse. 2002. Remarks on marking. In *Proceedings of the 8th International Conference on HPSG*, pages 267–283, Stanford, CA: CSLI Publications.

Zwicky, Arnold and Pullum, Geoffrey K. 1983. Cliticization vs. inflection: English *n't*. *Language* 59, 502–523.