Christoph Schwarze (University of Konstanz)

On the representation of French and Italian clitics

6th International LFG Conference, Hong Kong, June 27, 2001

0. Introduction

Romance clitics, and more precisely, clitic pronouns, have been a topic in generative linguistics since the seventies. For linguists like Perlmutter (1971) and Kayne (1975), their most striking property was their linear order, which is different from the order of the phrases that realize the same grammatical functions. The direct object, e.g., follows the verb when it is noun phrase, but precedes it when it is a clitic pronoun. Given the basic assumption of the generative theory of that period, accounts were proposed in terms of derivations from canonical abstract representations of the sentences containing them. In more recent work, the focus has shifted from how clitics how can be derived to how they are to be represented. New proposals on the treatment of Romance clitics have been made in the framework of optimality theory (OT, Grimshaw 1997) and of feature and exponence theories, also known under the label of Distributed Morphology (Monachesi 1999, 2000, Everett 2000, Luís & Sadler 2001).

Among these new approaches, the more radically innovative is the one known as Realizational or Distributed Morphology (DM). OT, in fact continues to consider clitics as morpho-syntactic entities, whereas DM denies them that status. Furthermore, Grimshaw's OT treatment of Italian clitics is limited in scope, since it does not comprise the whole inventory. [1] I will argue that the classical view, according to which Romance clitics are visible elements of c-structure is adequate and that there is no need to revise the classical LFG framework in order to be able to account for their grammatical properties. But under this assumption, two sets of descriptive facts still need to be accounted for, the restrictions and idiosyncrasies of clitic clusters, and the high ambiguity and syncretism of clitics.

The paper is organized as follows: In section 1, I will give a descriptive overview of those properties of Romance clitics that are a challenge to

grammatical analysis. In section 2, I will discuss the DM approach to the analysis of Romance clitics. In section 3, I will show how notorious difficulties of representing Romance clitics can be resolved in the classical LFG framework.

1. The strange properties of Romance clitics

Romance clitics have a number of properties that are a challenge to grammatical analysis. They show mismatches between argument structure and constituent structure (1.1); they have a high degree of ambiguities, categorial as well as functional (1.2), and they show, in the case of Italian and Spanish, irregular phonological variation (1.3).

1.1 Mismatches between argument structure and constituent structure

Clitic clusters have an order that may differ from the order of the corresponding phrases. The default order of DIRECT and INDIRECT OBJECT (OBJECT II), when they are phrases, is "DIRECT OBJECT before INDIRECT OBJECT" (1). When the clitics are third person, the order is the same (2a), but it is opposite when the indirect object is first or second person (2b).

(1) La fée donne une bague au chevalier.
the fairy gives a ring-DO to-the knight-IO
'The fairy gives the knight a ring'.

(2) a. La fée le lui donne.
the fairy it-CLITIC-PERS3-DO him-CL-PERS3-IO gives
'The fairy gives it to him'.

b. La fée me la donne.
the fairy me-CL-PERS3-IO it-CLITIC-PERS3-DO gives
'The fairy gives it to me'.

There are asymmetries in distribution: the first and second person DIRECT OBJECT-clitics cannot form clusters with INDIRECT OBJECT-clitics (3), whereas the third person DIRECT OBJECT-clitics can (4):

(3) a. Elle *me lui présente.
she me-CL-DO-PERS1 him-CL-IO-PERS3 introduces

---

[1] Grimshaw 1997 considers neither *ne*, the partitive and "genitive" clitic, nor *ci* and *vi* in their spatial interpretation.

'She introduces me to him' .

b. Elle *lui me présente.
she him-CL-IO-PERS3 me-CL-DO-PERS1 introduces
'She introduces me to him.

(4) Elle le lui présente.
she him-CL-DO-PERS3 her-CL-IO-PERS3 introduces
'She introduces him to her'.

Furthermore, the French first and second person DIRECT OBJECT and INDIRECT OBJECT clitics cannot be placed at the right-hand side of the verb unless they are followed by another clitic (5), whereas there is no such restriction for the third person clitics (6):

(5) a. *Aidez-me !
help me-CL-DO-PERS1
'Help me'!

b. Donnez-m'en !
give me-CL-IO-PERS1 from-it-CL-PARTITIVE
'Give me some'!

(6) Dis-le!
say it-CL-DO-PERS3
'Say it'!

In spoken French, one finds so-called portmanteau clusters, i.e. the merger of two grammatical functions into one form. This holds for SUBJECT and DIRECT (7) and for INDIRECT OBJECT and DIRECT OBJECT (8); the same phenomenon appears in Old French texts (9)-(10):

(7) [ʒezevy] *je les ai vus*
underlying lexical representation: /ʒez e vyz/
I-CL-SUBJ+CL-DO have seen
'I saw them'

(8) [ʒɥidire] *je le lui dirai*
underlying lexical representation: /ʒə yi dire/
I-CL-SUBJ him-CL-DO+IO will-say
'I will tell him this'

(9)

Prenez	Take
mon avoir, que vos la veez,	my belongings which you see there,
en cele male qui la pent …	in that bag which is hanging there …
Se ge muir portés la lou roi,	If I die, bring it to the king,
si dites que ge *li* envoi.	and say that I am sending *it to him*.

(*La Male Honte*, cit. Foulet 1958:148)

(10) Di mei <…> u sunt voz dras ? — Tell me <…> where are your clothes?
Dame, ceo ne dirai jeo pas — Lady, this I will say by no means.
Kar si *jes* eüsse perduz — Because if I lost them
E de ceo feusse aparceüz, — and would be discovered,
Bisclavret sereie a tuz jurs. — I would be a werewolf forever.

(Marie de France, *Bisclavret*, 71ff; Rychner 1983:63)[2]

1.2 Ambiguities

From a descriptive point of view, the system of French clitics comprises twenty forms. Only eight of them (11) are unambiguous. The other twelve (12) are ambiguous in several respects:

(11) je, tu, ils, ce, la, lui, leur, les
(12) il, elle, ça, nous, vous, elles, on, me, te. se, y, en

The ambiguity is categorial for *elle, ça, nous, vous* and *elles*, i.e., these forms are clitics as well as noun phrases. This becomes visible most clearly in clitic doubling, as in (13) and (14):

(13) Elle, elle a compris.
she she-CL-SUBJ has understood
'She understood'.

(14) Ça, ça va pas.
that that-CL-SUBJ goes not
'That doesn't work'.

The ambiguity is a semantic one for the subject clitic *il*, which is a pronoun or an expletive (15).[3]

(15) a. Il chante. — pronoun
'He is singing'.
b. Il pleut. — expletive
'It is raining'.

Several clitics are ambiguous with respect to the grammatical function that they represent. Some (*me, te, se, nous, vous*) are DIRECT OBJECT as well as INDIRECT OBJECT; cf. (16):

(16) a. Elle me verra.
she me-CL-DO will-see

[2] Similar portmanteau clitics occur in Portuguese; cf. Luís & Sadler (2001:2).

[3] Other clitics may be expletive in idioms: *le* in *l'emporter* 'to gain the upper hand', *en* in *en vouloir à quelqu'un* 'to be angry with', *se* in *s'en aller* 'to leave', *y* in *ne pas y aller de main morte* 'to exaggerate'. I do not discuss the representation of idioms in this paper.

'She will see me'.

b. Elle me parlera.
she me-CL-IO will-speak
'She will speak to me'.

Y is an Oblique (17) or an Adjunct (18):

(17) J'y ai pensé.
I Y-CL-OBLIQUE-INANIMATE have thought
'I have thought of it'.

(18) Je l'y ai vu.
I le-CL-DO Y-CLITIC-SPATIAL-ADJUNCT have seen
'I saw him there'.

En is an Oblique (19) or a Partitive Modificator of the DIRECT OBJECT (20):

(19) Je leur en ai parlé.
I them-CL-IO of_it-CL-OBL have spoken
'I spoke to them about it'

(20) Elle en aime un autre.
she of_them-CL-PARTITIVE an other
'She loves someone else'

And *le* is not only a DIRECT OBJECT (21) but also a Complement of the copula (22):

(21) Je le verrai.
I him-CL-DO will-see
'I will see him'.

(22) Je le serai.
I it-CL-XCOMP will-be
'I will be'.

And the clitics of the first and second person are personal (23) as well as reflexive pronouns (24):

(23) Je te défendrai.
I you-CL-PERSONAL will-defend-PERS1
'I will defend you'.

(24) Tu te défendras.
You you-CL-REFLEXIVE will-defend-PERS2
'You will defend yourself'.

1.3 Irregular phonological variation

Some clitics show irregular phonological variation, i.e. alternations of phonological shape which are not due to postlexical phonology. In French,

the vowel of the feminine singular object clitic *la*, just as the homophonous determiner, is elided before vowel (25), but there is no rule of *a*-elision (26):

(25)	Il l'aime 'He loves her'.	*Il la aime	Elision
(26)	Mon énergie 'My energy'.	*m'énergie	No elision

In Italian, the situation is similar (27):

(27)	a. Lui l'ama 'He loves her'.	?lui la ama	Elision
	b. Alta Italia 'Upper Italy'.	*alt'Italia	No elision

In Italian, *mi*, *ti*, *si* and *ci*, when they precede a direct object-clitic, lower /i/ to /e/ (28), a phenomenon which does not appear anywhere else in that language:

(28) a. *Mi* vedi?
'Can you see me'?

b. *Me* lo dici?
'Will you say it to me'?

And, finally, dissimulation is compulsory when two clitics are combined which have the same or almost the same phonological shape: in Italian, *si+si* appears as *ci si* (29), and in Spanish, *le+lo* is replaced with *se lo* (30):

(29) a. Si capisce.
SI-CL-SUBJ understands-SG
'One understands'.

b. Si capiscono.
SI-CL-RECIPROCAL understand-PL
'They understand each other'

c. Ci si capisce. *Si si capisce.
SI-CL-SUBJ SI-CL-RECIPROCAL understand-SG
'One understands each other'

(30) a. Le respondré mañana.
him-CL-IO will-answer-PERS1 tomorrow
'I will answer him tomorrow'

b. Lo digo.
it-CL-DO say-PERS1
'I say it'

c. Se lo digo. - *Le lo digo.

him-CL-IO it-CL-DO say-PERS1
'I say it to him'

2. Treating clitics as exponents

In the presence of such an amount of intricate descriptive facts, DM appears to be an ideal approach. This idea may seem to be paradoxical in a lexicalist framework such as LFG, given that in DM there is no such thing as a lexicon in the classical sense (Harley & Noyer 1999:3). But one could plausibly think of a syncretistic format, where only certain grammatical domains are represented in terms of mapping from feature bundles to phonological forms. One could consider that the Romance clitic systems are one of these domains. Syntax, then, does not have to account for the ambiguities and variations of the Romance clitic systems; it just has to provide feature configurations for clitics, identify their hosts, map them onto phonological strings, and put them in the correct linear order with respect to their host. In a sense, if such a treatment is chosen, it can actually be said that "there are no clitics" (Everett 2000).

Moreover, the treatment of Romance clitics as phonological exponents is no longer a mere program. Monachesi (1999) has shown how the Italian clitic system can elegantly be represented in HPSG on a DM base. All the problems mentioned above seem to be solved, included the treatment of clitic clusters. By treating them as complex objects at the level of features, and as unanalyzable strings at the level of surface expression, Monachesi (1999:76) not only captures their idiosyncratic restrictions, but also reduces their ambiguities to a large extent. In fact, the ambiguities which clitics show in isolation disappear in clitic clusters. Thus, those clitics that alternatively represent DIRECT and INDIRECT OBJECTS (*mi*, *ti*, *si* etc.), are unambiguous in clitic clusters, and *lo* is always a DIRECT OBJECT, never an XCOMP, when it appears in a cluster.[4]

[4] Another analysis of Romance, more precisely Portuguese, clitics, based on DM principles, is contained in Luís & Sadler 2001. The Portuguese clitics show an intricate behavior regarding their precedence relationship with and the identity of their hosts. If, unlike these authors, one treats morphological segments as constituents, then precedence relations are expressed by c-structure rules. This is doubtlessly feasible.

The lexicalist the representation of clitic clusters which I will propose in section 3 differs from Monachesi's regarding the architecture of the grammar, but it seems to be logically equivalent if one looks at the relationship between the features and their realizations ("morphological segments" in my terminology). In fact, Monachesi represents, e.g. It. *gli*, the masculine third person dative clitic pronoun, as (31):[5]

(31) [STEM … CLTS <dat, pers3, sg, masc>] →
[AFFIX [PHON <*gli*>]]

And she explains this notation with the following words: "It states that if there is a STEM with a CLTS list with one element which is a third person singular, masculine, dative NP, it must be realized as a clitic whose phonological form is *gli*." (Monachesi 1999:73).

A traditional LFG lexical entry would be (32):

(32) *gli*, CL
(↑ PRED)=PRO, (↑ GEN)=MAS, (↑ NUM)=SG, (↑ PERS)=3, (↑ PCASE)=A

This notation states that there is a morphological segment *gli*, which is a clitic and encodes the features pronoun, masculine, singular, third person, and dative case. The only difference, then, seems to consist in how constituency is treated in LFG and HPSG: in LFG, *gli* is some kind of constituent (labeled CL), whereas in Monachesi's HPSG analysis, it is not a constituent of any kind, but just a phonological string.

Now the existence of elegant and descriptively satisfying analyses like Monachesi's do not necessarily prove the adequateness of the underlying theoretical claim. And, in fact, there are a number of objections to the DM approach, which I will address now, going from the general to the specific.

The most general objection is that there is no need to import DM into LFG, since the stratified representations of LFG make it possible to map feature structures onto word-forms which have no morphological structure, i.e., are not analyzable in c-structure. DM may well represent an important progress with respect to theories where form and function must be isomorphic, in such a way that unanalyzable word-forms must be derived

[5] I have omitted those elements of the representation which are not relevant to my argumentation and made some abbreviations more readable.

from underlying analyzable forms, as, e.g. *went* from *go+ed*. But LFG, just like DM, directly relates word-forms and feature configurations, no matter whether the word-forms have an internal constituency or not. In other words, the problem that is solved by DM, is also solved in LFG.

A more specific objection is that current presentations of the DM-model are formulated in such a way that it is not obvious what the phonology looks like. Taken at face value, DM excludes all kinds of Lexical Phonology, since there is no lexicon. One could of course explain that the items, which are spelled out by the realization rules of DM, are precisely the "lexical representations" to which postlexical phonology applies. But if affixes are not "visible" in any part of the grammar, one may lose important generalizations. As an example, take Italian thematic vowels, such as /i/ in [sen'ti:te]*sentite* 'you hear', [sen'ti:vo]*sentivo* 'I heard', [sen'ti:to]*sentito* 'heard (participle)'. The syllables in which they occur are candidates for stress assignment, in spite of their "light" linear structure. If the phonology does not know that they are thematic vowels, this factor of the stress assignment system is reduced to a huge list of arbitrary spell-outs.

If this may still be accepted in merely structural model of phonology, the situation becomes dramatic if one passes from the structural to the procedural point of view. In the FUL-model of speech perception (Preuss et al. in press), acoustic properties are transformed into phonological features and the phonological feature structures are mapped onto abstract lexical representations. The lexical representations of complex, analyzable words are strings of morphological segments, e.g. stems and affixes, and these segments are listed in the lexicon[6]. It seems hard, in fact, to see how phonological features could be mapped directly onto morpho-syntactic features.

Furthermore, there are certain descriptive facts that are hard to treat in terms of DM. One of them is the derivation of manner adverbs from adjectives in Italian. The general structure of these adverbs is (33):

(33) the feminine singular form of an adjective + *mente*

[6] Actually, Preuss et al. (in press) use the term "morpheme", but in their use that term does not imply the notion of Saussurian sign.

The Italian adjective has two distinct inflectional classes; cf. (34):

(34) The *o*-Class

	Sing.	Pl.
Masc.	***o***	***i***
Fem.	***a***	***e***

The *e*-Class

Sing.	Pl.
e	***i***

Hence, the feminine singular form of *o*-Class adjectives ends with *-a*, (e.g. *rapida*) and the singular form of *e*-Class adjectives ends with *-e* (e.g. *cortese*). Accordingly, one gets manner adverbs like the following:

(35) a. rapidamente 'fast'
b. cortesemente 'politely'

Now, how are forms like these derived in DM? The DM-type morphological representations of *rapida* and *cortese* are, respectively:

(36) rapid + [GEN = FEM, NUM = SG]
cortes + [NUM = SG]

Should *mente* be affixed to these representations, in such a way that one gets representations like the ones in (37)?

(37) rapid + [GEN =FEM, NUM = SG]+mente
cortes + [NUM = SG]+mente

Under this treatment the representations in (37) must then be mapped to phonology. But since there must also be a mapping relationship for the adjectives, there would be an unnecessary doubling.

Or should the adjectives first be realized by mapping them onto phonology, like in (38)?

(38) rapid + [GEN = FEM, NUM = SG] /'rapida/
cortes + [NUM = SG] /kor'teze/

The derivational suffix must then be adjoined to the spell-outs at the right-hand side of the double arrow. But this would create monsters, made up of a phonological representation and a morphological segment.

As a consequence, it will certainly be more adequate to regard the suffixes of Italian adjective inflection as visible morphological segments. And if

derivational affixes are visible in the lexicon, why should other morphological segments not be visible as well?

There also is an argument specific to the data treated in the present paper. As has been pointed out above (3), certain clitic clusters which are licensed by argument structure, are ungrammatical. In a DM treatment, one has to rule out those feature configurations for which there are no phonological exponents. Monachesi (1999:80) postulates a "Clitic Realization Principle", which states that

> "All verbal hosts which have a nonempty CLTS list must satisfy one of the realizational constraints".

But these constraints are nothing else than lists of feature configurations paired with phonological forms. If, on the other hand, clitic clusters are stored in the lexicon as morphological segments to which feature configurations are associated, then the non-existence of some segments does not come as a surprise. Idiosyncratic restrictions on distribution are irregularities, comparable to those of defective inflectional paradigms. And were else than in the lexicon should irregularities be stored?

To conclude, it is not obvious that DM should be an advance with respect to LFG. Furthermore, it seems to have some undesirable consequences. Even if the question of how the syntax-morphology and the morphology-phonology interfaces are best modeled within the LFG framework is still open, DM does not seem to impose itself for the topic at hand, namely the treatment of Romance clitics. I will therefore maintain the hypothesis that clitics are morphological segments to which feature structures are associated, and I will try to show how some of the crucial descriptive facts can be accounted for without changing LFG.

3. Using the LFG framework for representing Romance clitics

I will now give a sketch of the c-structure analysis of Italian and French clitics (3.1) and make a proposition regarding the relationship between morphological case and grammatical function (3.2). I will then address three of the challenges which Romance clitics present for a treatment in terms of "visible" lexical items. Two of them concern functional ambigui-

ties, namely ambiguities of grammatical function (3.3) and ambiguity between reflexivity and non-reflexivity (3.4). The third concerns the representation of clitic clusters (3.5). I will close with an outlook at the problems inherent in Italian impersonal *si* and French first plural *on* (3.6).

3.1 The c-structure analysis of Italian and French clitics

Since Grimshaw (1982), French clitics are treated as postlexical affixes to the verb. Even though she does not use this term, the idea is clearly expressed by the c-structure rule she proposes (Grimshaw 1982:90), and which I reproduce here as (39) in a simplified version:

(39) V' → (CL) (AUX) V

The question is, of course, whether this is a correct analysis. Clitics obviously resemble affixes. This holds especially for those cases in which the clitic is placed to the right-hand side of its host. Luís & Sadler (2001:4) show that, in Portuguese, the enclitic forms show morphological interaction with the verb. This also holds for Italian: If the verb form is an infinitive, the clitic is attached to its right-hand side, and the infinitive suffix /re/ loses its final /e/; cf. (40):

(40) fare 'to do' - farlo 'to do it'

This is a morphological alternation that also occurs in Italian word formation. In fact, adverb formation with *–mente*, which inactivates the inflectional features of the adjectival basis, deletes the corresponding morphological segment if the stem ends with /l/ or /r/; cf. (41):

(41) tale 'such a' talmente 'so'
regolare 'regular' regolarmente 'regularly'

But the fact that e-deletion is also phonologically conditioned, raises suspicion. Moreover, e-deletion also is required in phrases; cf. (42):

(42) dottore 'doctor' dottor Rossi 'Dr. Rossi'
potere 'to be able' per poter farlo 'in order to be able to do that'
bene 'well' ben tre volte 'three times'[7]

[7] *Three* pronounced with a stress, expression that three times is more than enough.

It might well be that, at least regarding Italian e-deletion, the interaction of the clitic with the verb takes places in phonology; this is an open question. As far as French is regarded, the verb form never seems to be affected by the presence of a clitic. I will conclude that there is no reason to deny that clitics are inserted postlexically. Their similarity with affixes can probably best be expressed, rather than at the c-structure level, by the representation of their functional properties.

I will thus follow Grimshaw's (1982) general idea[8]. The rule for Italian will be (43), if the sentence is declarative and the verb (auxiliary or full verb) is finite:

(43) V' → (CL1) (CL2) V

Notice that in this treatment, CL is thought to include clitic clusters as well as single clitics. CL1 is the negation particle *non*, which needs a verb as a host. The category CL2 functionally comprises objects, obliques and adjuncts. Unlike French, Italian has no subject clitics. The distinction, at the c-structure level, between two kinds of clitics is necessary because the negation particle always is at the left-hand side of the verb.[9]

For non-negated imperative sentences in the second person and for non-finite forms of the verb, the order is inverted:

(44) V' → (CL1) V (CL2)

For French, the situation is more complex, inasmuch as there is one more type of clitic forms: the subject clitic (CL1), which always precedes the negative particle *ne* (CL2), and which differs in distribution from the other

---

[8] For the representation of clitic clusters, Grimshaw (1982) makes the rule repeat the category CL, assigning each CL a position:

V' → (CL)1 (CL)2 (CL)3 (AUX) V

A different proposal for clitic clusters will proposed here. When, further on, I will use the notations like CL1, CL2, CL3, this will be meant to express not different positions, but different categories of clitic forms.

[9] Monachesi (2000) makes an analogous distinction for Romanian. According to her analysis, there are two kinds of clitics, those, which have word status, and those, which are affixes, and she derives the ordering of clitics from that difference of status. My own distinction is only motivated by facts of distribution.

pronominal clitics (CL3). The rule for French declarative sentences is (45), regardless of whether the verb is finite or not:

(45) V' → (CL1) (CL2) (CL3) V

In imperative sentences, there is no subject clitic, and the other pronominal clitics are placed at the right-hand side of the verb:

(46) V' → (CL2) V (CL3)

In French interrogative sentences[10] the clitic subject follows the verb; the rule is thus (47):

(47) V' → (CL2) (CL3) V (CL1)

These are the facts that justify the assumption of three clitic categories in French, as opposed to only one in Italian.

3.2 Morphological case and grammatical function

In Grimshaw 1982 French clitics are lexically assigned a case attribute with the values "accusative" and "dative"; Monachesi 1999 uses the same case labels for Italian, adding "locative" for *ci*, and one could use "genitive" for French *en* and Italian *ne*, and "nominative" for the French subject clitics. Grimshaw then maps cases onto grammatical function in the c-structure annotations.

It is not obvious, however, why one should have case features in the analysis of languages like French and Italian. It is true that, historically, the clitic forms are continuations of Latin case forms. But this is not a reason to postulate case as a feature in modern Romance. Since no other category than clitics have case, it is simpler to directly assign grammatical functions to clitics in the lexicon and to write entries such as (48) and (49) for French:

[10] In French, interrogation can also be realized by declarative sentences with an interrogative intonation or an interrogative pronoun.

For Italian, which has no subject clitics, (48) has no equivalent, and the category of (46) would just be CL.

(48) *je*, CL1
(↑ PRED) = 'PRO'
(↑ PERS) = 1
(↑ NUM) = SG
(↑ SUBJ)

(49) *la*, CL3
(↑ PRED) = 'PRO'
(↑ PERS) = 3
(↑ NUM) = SG
(↑ GEN) = FEM
(↑ OBJ)

This representation, in a model of grammar that has grammatical functions, eliminates a useless step of computation.

Moreover, there are two types of clitic doubling. The clitic may realize some grammatical function, and the constituent with which it is coreferential, be a TOPIC. The clitic may also, instead of realizing a grammatical function, be just an agreement marker, as in Italian (50) and French (51).[11]

(50) Me la dai una mela?
me-CL-IO it-CL-OBJ_AGR an apple-OBj
'Do you give me an apple'?

(51) Il est sympa ce garçon.
he-CL-SUBJ_AGR is nice this boy-SUBJ
'That's a nice boy'.

This use of clitics as agreement markers can be represented more easily if the grammatical function of the clitic is directly encoded in the lexicon. In order to do this, all that it takes is to revise entries like (48) and (49) above in such a way that instead of grammatical functions, they encode agreement with grammatical functions; cf. (52), the revised formulation of (49):

(52) *la*, CL3
(↑ OBJ PRED) = 'PRO'
(↑ OBJ PERS) = 3
(↑ OBJ NUM) = SG
(↑ OBJ GEN) = FEM

[11] The criterion for the latter structure is intonation: when a clitic is used as an agreement marker, the typical TOPIC intonation of the related noun phrase is lacking. The clitics, which appear in this construction, seem to be restricted. Typically they are the object and, in French, the subject clitic. Whether other clitics are used as mere agreement markers is subject to further investigation.

This entry accounts for both, the use of the clitic as an agreement marker and as the realization of a grammatical function. It creates an object that is merged with the object of the verb if there is one, and alternatively becomes the object if there is no other object.[12]

3.3 Ambiguities of grammatical function

Some clitics show ambiguities with regard to the grammatical function they realize which can only be handled by lexical listing. This is the case, e.g., for French *y*. This clitic may be an oblique with PCASE A (53), a local OBLIQUE (54), and a local adjunct (55):

(53) J'y pense.
I it-CL-IO think
'I think of it.'

(54) J'y vais.
I there-CL-OBL go
'I go there.'

(55) J'y dors.
I there-CL-ADJUNCT sleep
'I sleep there.'

Three entries are needed for (53)-(55):

(56) *y*, CL3
(↑ OBL PRED) = 'PRO'
(↑ OBL PCASE) = A
(↑ OBL ANIMATE) = —

(57) *y*, CL3
(↑ OBL PRED) = 'PRO'
(↑ OBL LOC) = +

(58) *y*, CL3
(↑ ADJ PRED) = 'PRO'

[12] Halpern & Fontana (1994) make a distinction between two kinds of clitics, which they name X° and X^{max} clitics. Roughly speaking, X° clitics support inflection and agreement, whereas X^{max} clitics represent arguments. This distinction is not adequate for Romance, where clitic doubling is functionally ambiguous. Diachronically, clitics develop from argumental phrases. After phrases have turned into clitics, they may keep their status as arguments or lose it. The evolution of Romance clitics seems to have reached a stage where both kinds of functional status, argument and mere agreement marker, are synchronically present.

(↑ ADJ LOC) = +

But for other clitics, generalizations are possible, and they can be expressed via lexical rules. The first and second person clitics *me, te, nous, vous* are both, DIRECT OBJECT and INDIRECT OBJECT. The same holds for the reflexive *se*. This can be expressed by encoding only one of the functions and derive the other. The choice is arbitrary in principle, but the lexical rule can be formulated more simply if the encoded function is the INDIRECT OBJECT. The rule then reads as follows:

(59)

x, CL3	*x*, CL3
(↑ OBL PRED) = 'PRO'	(↑ OBJ PRED) = 'PRO'
(↑ OBL PERS) = {1,2}	(↑ OBJ PERS) = {1,2}
(↑ OBL PCASE) = A	

3.4 The ambiguity between reflexivity and non-reflexivity

For reflexive constructions, French and Italian, like other languages, e.g. German, have a formally unambiguous form only for the third person; in reflexive constructions in the first and second person, the personal pronoun is used.[13] But while, in German, the retrieval of first and second person reflexivity can be left to semantics, French and Italian need to make reflexivity explicit at the f-structure level for all persons for the sake of auxiliary selection in the compound tenses. In fact, the representation of sentences like French (60a) and Italian (61a) must make a feature "reflexive" available, analogously to the b-sentences:

(60) a. Tu t'es trompé.
you-CL-SUBJ-PERS2-SG you-CL-OBJ-PERS2-SG are misled
'You made a mistake.'

b. Il s'est trompé.
he-CL-SUBJ-PERS3-SG-MAS him-CL-REFL-OBJ-PERS3-SG is misled
'He made a mistake.'

(61) a. Mi sono fatto la barba.
me-CL-OBL-PERS1-SG am-PERS1-SG made the beard
'I shaved.'

b. Si è fatto la barba.
him-CL-REFL-OBL-PERS3-SG is-PERS3-SG made the beard

[13] This is probably due to the fact that semantic reflexivity can be derived from the correferentiality of the subject and the direct or indirect object if the subject is in the first or second person.

'He shaved.'

Therefore the personal pronouns of the first and second person must also be represented as reflexive pronouns by a feature, which I will write as "REFLEXIVE = +". One might think of expressing this ambiguity via a lexical rule, comparable to the one I proposed above for the DIRECT OBJECT-INDIRECT OBJECT ambiguity. But since the REFLEXIVE feature is needed for personal pronouns only in sentences with compound tenses, the ambiguity can be represented more adequately as a conditional constraint, i.e. as an inference drawn at f-structure. It can be formulated as (59):

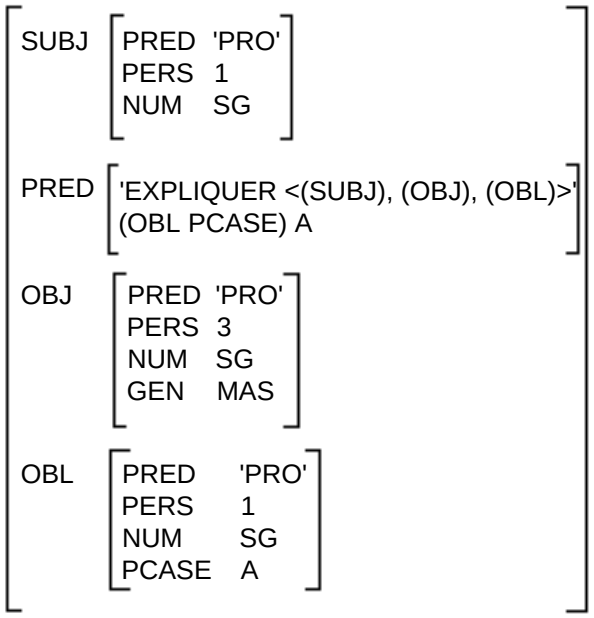
(62) ***Reflexive inference***

SUBJ PERS = {1 ,2} & {OBJ, OBL} AGR = SUBJ AGR {OBJ, OBL} REFL = +

For illustration, consider sentence (63) and its f-structure (64), such as it is built up from the annotated c-structure tree:

(63) Je me l'explique.
I me-CL-OBL-PERS1 it-CL-OBJ-PERS3 explain
'I can see why this is so.'

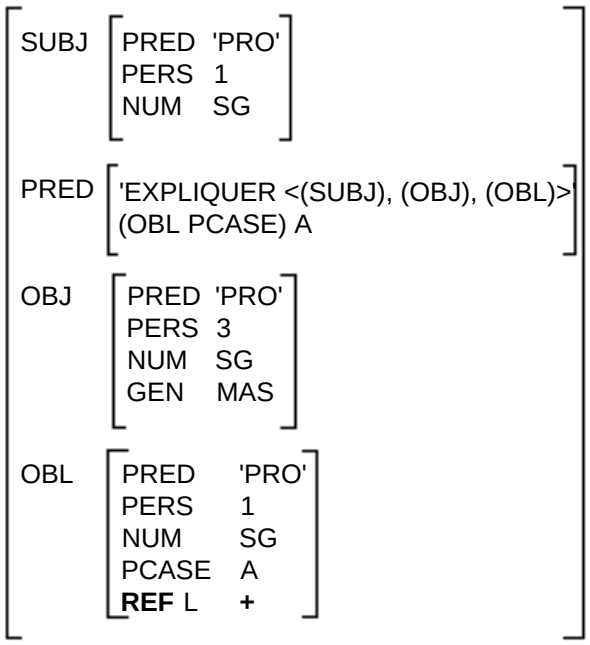
(64)



F-structure (64) meets the conditions for applying Reflexive Inference. When Reflexive Inference is applied, the resulting f-structure is (65). It

differs from (64) insofar as the feature "REFL = +" is included in the functional description of the OBLIQUE.[14]

(65)



3.5 The representation of clitic clusters

As has been said above, a large amount of the difficulties of the French and Italian clitic systems can be avoided by representing clitic clusters as forms that are analyzed only at the f-structure level. The inventory of the items belonging to category CL2 in Italian and CL3 in French will therefore comprise, along with the simple clitics, the clitic clusters. We will thus account for the rigidity of linear order, for the restrictions on co-occurrence, for the merger of two arguments into one clitic and for phonological variation in the simplest way. The notation introduced above for representing simple clitics turns out to be most useful to keep the two sets of

[14] There is nothing, however, in (60), which triggers the application of Reflexive Inference. In a more explicit formulation it should be expressed that Reflexive Inference takes place only of the value of tense is the name of a compound tense.

features apart. As an example, consider the lexical representation of *me la* (66); merged arguments can be represented in exactly the same way (67):

(66) *me la*, CL3
(↑ OBL PRED) = 'PRO'
(↑ OBL PCASE) = A
(↑ OBL NUM) = SG
(↑ OBL PERS) = 1
(↑ OBJ PRED) = 'PRO'
(↑ OBL NUM) = SG
(↑ OBL GEN) = FEM
(↑ OBL PERS) = 3

(67) /yi/, CL3
(↑ OBJ PRED) = 'PRO'
(↑ OBJ NUM) = SG
(↑ OBJ GEN) = MAS
(↑ OBJ PERS) = 3
(↑ OBL PRED) = 'PRO'
(↑ OBL PCASE) = A
(↑ OBL NUM) = SG
(↑ OBL PERS) = 3

3.6 Italian impersonal si *and French first plural* on

In Monachesi (1999) impersonal *si* is considered a subject. If this is a correct analysis, which I will assume for the following discussion, then sentences like (68 b), are a problem for HPSG. Monachesi (1999:98) resolves it by a lexical rule, which removes the subject "from the subject list of the verb" and adds it "as a member of the CLTS list."

(68) a. Si mangia spesso le fragole.
SI eats often the strawberries
'One often eats strawberries.'

b. *La gente si mangia spesso le fragole
the people SI eats often the strawberries
'The people one eats often strawberries.'[15]

In an LFG analysis, this problem does not even appear: sentences like (68 b) are taken care of by the uniqueness constraint; whereas HPSG has to postulate an ad hoc rule.

But impersonal *si* presents still another problem: in sentences with an XCOMP, the XCOMP is in the plural, whereas the verb is in the singular:

[15] Monachesi's glossing and translation. Instead of *one*, *we* could also be used as a translation of impersonal *si*.

(69) Si è contenti.
SI is-SG happy-PL-MAS
'One is happy.'

Monachesi (1999:99ff) proposes a solution based upon a sophisticated theory of agreement, proposed by Pollard and Sag (1994) for French sentences like (70) and further developed by Kathol (1999).

(70) Vous êtes belle.
you-PL are-PL beautiful-SG
'You are beautiful.'

This theory uses indices in such a way that agreement constraints can be stated with respect to an index rather than between two inflected forms. In LFG, one might think of treating the problem in the following way (which I could not elaborate here): The agreement between the verb and its subject is standard. The XCOMP, however, has a (covert) subject of its own, which is only referentially identical with the subject of the sentence. This covert subject must be plural when impersonal *si* is the subject of the verb, and it can be both genders:

(71) Si è contente.
SI is-SG happy-PL-FEM
'One is happy.' (said of women)

An analogous solution is desirable for the French subject clitic *on*, which, in the spoken language, is semantically first person plural and agrees as such with its XCOMP (72) or Past Participle (73), but requires the third person singular for the verb:

(72) On est sympathiques.
ON is-SG nice-PL
'We are nice.'

(73) On est partis.
ON is-SG left-PL
'We left.'

The elaboration of a more expressive theory of agreement for LFG is a task for further research.

4. Clitic pre-emption

The phenomena that Sadler (1997:11) subsumes under the term of clitic pre-emption, concern an aspect of the grammar of clitics which is only

rarely addressed in the literature. Regarding the languages treated in the present paper, the facts are the following.

In French, if an object or an oblique is realized by a pronoun, this pronoun must be a clitic; cf. (74):

(74) a.Il t'aime.
he CL-OBJ-PERS2-SG loves

b. *Il aime toi
he loves NP-OBJ-PERS2-SG

The non-clitic pronoun is possible only in clitic-doubling:

(75) Il t'aime, toi.
he CL-OBJ-PERS2-SG loves NP-OBJ-PERS2-SG
‚he loves you' (main stress on *you*)

Sadler (1997) observes similar facts for Welsh and explains them as the consequence of a general blocking principle. A similar solution may be formulated for French.

But regarding Italian, the situation is slightly different: The cliticity of the pronoun, rather than being necessary, is only the unmarked choice (75a). The non-clitic pronoun is not blocked, it is only marked; cf. (75b):

(75) a. Mi piace.
me-CL pleases
‚I like it'

b. A me piace.
to me-NP pleases
‚I like it' (stress on *I*)

The blocking and markedness effects that clitics may produce deserve to be carefully investigated, also with respect to other languages. But it is very likely that the underlying principles must refer to hierarchies in the lexicon. This is one more reason for a lexicalist treatment of clitics.

5. Conclusion

I hope to have shown that an adequate representation of French and Italian clitics is possible without changing the general LFG framework. It is true that there remain some points which need further reflection, and, which is more important, that the data are only from two languages. But the fact that such grammatical objects as clitics, which are on the border-

line between syntax and morphology speaks in favor of the robustness of the framework. Given its week theoretical commitments, LFG makes it possible to develop theories without changing the overall framework, and it certainly is a wise research strategy to take advantage of such robustness not to modify the general framework unless it is necessary.

References

Bresnan, Joan. 2001. *Lexical-Functional Syntax*. Oxford: Blackwell.

Everett, Daniel L. 2000. "Why There Are no Clitics. On the Storage, Insertion, and Form of Φ-Features". in: Peter Coopmans, Martin Everaert & Jane Grimshaw (eds.). *Lexical Specification and Insertion*. Amsterdam/Philadelphia: John Benjamins. pp. 91-114.

Foulet, Lucien. 1958. *Petite syntaxe de l'ancien français*. 3me éd. revue. Paris: Champion.

Grimshaw, Jane. 1982. "On the lexical representation of Romance reflexive clitics." In: J. Bresnan (ed.) *The mental representation of grammatical relations*. Cambridge, Massachusetts: MIT Press, pp. 87-148.

Grimshaw, Jane. 1997. "The best clitic: constraint conflict in morphosyntax." In: Liliane Haegeman (ed.) *Elements of grammar*. Dordrecht: Kluwer, pp. 169-196.

Halpern, Aaron and Josep M. Fontana. 1994. “X° and X^{max} Clitics”. In: Erin Duncan et al. *The Proceedings of the Twelfth West Coast Conference on Formal Linguistics.* Stanford, CA: CSLI.

Harley, Heidi and Rolf Noyer. 1999. “State-of-the-Art Article: Distributed Morphology”. *GLOT* 4.4. pp. 3-9.

Kathol, Andreas. 1999. "Agreement and the syntax-morphology interface in HPSG." In: R. Levine and G. Green (eds.). *Readings in HPSG.* Cambridge: Cambridge University Press.

Kayne, Richard. 1975. *The Transformational Cycle.* The MIT Press.

Luís, Ana and Louisa Sadler. 2001. *Morphology and Markedness: Porttuguese Object clitics*. Handout of a paper given at the Department of Linguistics of the University of Konstanz on Tuesday, September 11, 2001.

Monachesi, Paola. 1999. *A lexical approach to Italian cliticization*. Stanford, California: CSLI Publications.

_____. 2000. “Clitic placement in the Romanian verbal complex”. In: Birgit Gerlach and Janet Grijzenhout (eds.). *Clitics in Phonology, Morphology and Syntax*. Amsterdam / Philadelphia: John Benjamins.

Perlmutter, David. 1971. *Deep and Surface Structure Constraints in Syntax*. New York: Holt, Rinehart and Winston.

Pollard, C. and Sag, Ivan. 1994. *Head-Driven Phrase Structure Grammar*. Chicago - Stanford: University of Chicago Press and CSLI Publications.

Preuss, Letizia, Marslen-Wilson, William and Lahiri, Aditi. in press. "Abstractness in stem representation: Phonology-Morphology interaction in processing". *Journal of Memory and Language*.

Rychner, Jean (ed.). 1983. *Les Lais de Marie de France*. Paris: Champion.