# Clitic incorporation and abstract semantic objects in idiomatic constructions[*]

M. Teresa Espinal (UAB)                    [forthcoming in *Linguistics*]

ABSTRACT

This paper analyses inherent clitics of idiomatic constructions as verbal arguments (Jelinek 1984, Baker 1996, Hale 2003) that are translated as free variables (Delfitto 2002): they are anaphoric to a (hidden) non-referential discourse topic or (right) dislocated constituent. Furthermore, since they denote abstract semantic objects (Asher 1993), they are assumed to be semantically incorporated into the verb to which they are affixed at the syntax-semantics interface (Van Geenhoven 1998), either by unification of selecting features or by predicate modification of some role function of the event under consideration (Farkas and Swart 2003, Dayal 2003). Whereas noun incorporation is a salient property of polysynthetic languages (as has been extensively argued in the linguistic literature; Mardirussian 1975, Mithun 1984), in this paper the incorporation of clitics and nouns is argued to be a covert phenomenon of non-polysynthetic languages, specifically in idiomatic constructions (Espinal 2001).

**1.** INTRODUCTION. The goal of this paper is to describe the lexicalization pattern schematized in (1), and to account for the semantic contribution of those pronominal forms which are selected in idiomatic constructions.[1] This pattern, which involves non-strong (or deficient) pronouns (Cardinaletti 1999, Cardinaletti and Starke 1999), is characterized by morphosyntactic cliticization (cf. Kayne 1975, 1989, 1991; Sportiche 1999; among others), and by semantic incorporation (cf. Van Geenhoven 1998, Farkas and Swart 2003, Dayal 2003).

(1)      V + Cl + (XP)

Assuming a syntactic analysis of clitics, and grounded in the analysis of noun incorporation and person-number-case morphology in polysynthetic languages (Mardirussian 1975; Mithun 1984; Jelinek 1984; Baker 1988, 1996; Hale 2003), this paper analyses the set of clitics occurring in the schema in (1) as incorporated classifiers: as morphosyntactic markers (of selectional, case, person-gender-number features) which are verbal arguments of a lexical verbal head, subject –like regular clitics– to clitic climbing in serial verb combinations. These verbal arguments are translated syntactically as free variables, anaphoric to a (hidden) non-referential discourse topic (Delfitto 2002), but since they denote abstract semantic objects (Asher 1993) they are submitted to a process of semantic incorporation (mainly, thematic unification à la Farkas and Swart 2003, or predicate modification à la Dayal 2003). Following this approach, this analysis makes the existence of compositional meanings, associated with incorporated structures, compatible with the possibility of acquiring non-compositional and figurative meanings in the encyclopedia. Thus, reacting against the tradition that idiomatic constructions involve exclusively non-compositional meanings (Katz and Postal 1963; Fraser 1970; Chomsky 1980, among others), this paper will show that the study of idioms is relevant in order to make a distinction between syntactically encoded meaning, which is compositional, and conceptually encoded meaning, which is non-compositional (Marantz 1996, 1997; Mateu and Espinal 2007).

The pattern in (1) –which is characteristic of idioms across different Romance languages such as Catalan, French, Italian, and Spanish, and also some non-Romance languages such as Greek– is illustrated from (2) to (6).[2]

(2)  CATALAN  (Espinal 2004a)[3]
 a. *Ballar-la*
   dance it-3rd.fem.acc.sg                'suffer'
 b. *Enginyar-se-les*
   think CL them-3rd.fem.acc.pl            'manage'
 c. *Ni somiar-ho*
   not dream it-ntr.acc                    'no way!'
 d. *Dir-ne de tots colors* (a algú)
   say CL.gen of all colours to somebody   'offend'
 e. *Caure-hi de quatre potes*
   fall CL.loc of four legs                'deceived'

(3)  FRENCH    (Rey and Chantreau 1979)
 a. *L'echapper belle*
   it-3rd.fem.acc.sg escape nice           'escape from a danger'
 b. *Se les rouler*
   CL them-3rd.fem.acc.pl roll             'do nothing'
 c. *Le payer* (à quelqu'un)
   it-ntr.acc pay to somebody              'pay someone out for sthg'
 d. *En entendre de belles*
   CL.gen understand of nice               'listen to incredible things''
 e. *S'y prendre*
   CL CL.loc take                          'manage'

(4)  ITALIAN    (Turrini et al. 1995)
 a. *Passarla liscia*
   pass+it-3rd.fem.acc.sg smooth           'get away scot-free'
 b. *Spararle grosse*
   shoot+them-3rd.fem.acc.pl big+fem.pl    'tell tall stories'
 c. *Dirlo ai quattro venti*
   say+ it-3rd.ntr.acc at four winds       'tell all and sundry about sthing'
 d. *Dirne delle belle*
   say+ CL.gen some nice                   'talk nonsense'
 e. *Farci caso*
   make+ CL.loc case                       'pay attention'

(5)  SPANISH
 a. *Armarla*
   arm it-3rd.fem.acc.sg                   'make trouble'
 b. *Pirárselas*
   go CL them-3rd.fem.acc.pl               'leave'
 c. *Pasarlo bien*
   pass it-3rd.ntr.acc well                'have a good time'

(6)  GREEK    (Bibis 2002:chapter4; Bibis and Roberge 2004)
 a. *Tin*                    *vrikame*
   her-3rd.fem.acc.sg        found-PAST.1st.pl
   'We've enjoyed ourselves'
 b. *To*                     *(e)vrase*        *(ja ta kala)*
   it-3rd.ntr.acc.sg         boiled-PAST.3rd.sg (for the good)

'(S)he is (really) late'

c.      *O Yanis*  *ta*              *eftise*
         D Yanis  them-3rd.ntr.acc.pl       spat-PAST.3rd.sg
         'Yanis is out of breath'

What is interesting about these paradigms is both the clitic pronominals selected, and the meaning associated with them. With regard to their morphological forms, they include mainly third person feminine accusative singular and plural pronouns; the neuter accusative; and in some languages (like Catalan, French and Italian) even the genitive and the locative forms. Neither first person nor second person pronouns occur in this pattern.[4] According to traditional grammarians, the presence of specific clitics is the overriding property necessary to assess the idiomatic meaning of the expressions illustrating this paradigm (Bally 1951:II,68, Grevisse 1986).[5] From now on I shall refer descriptively to this class of clitics as *inherent clitics* (henceforth i-clitics).[6]

        With regard to their semantics, they are not expletive,[7] since they denote some type or other of abstract semantic object. In the encyclopedia, they are not conceptually or referentially interpreted since complex predicates formed by V + incorporated clitics are connected to some syntactically non-transparent non-compositional meanings obtained at the output of various metaphorical and metonymic conceptual processes that apply non-systematically in idiom comprehension (Gibbs 1995). Let me illustrate this claim by considering the minimal contrasts between (7a-b) and (8a-b).

(7)     a.       A l'època dels exàmens aquests estudiants <u>ho</u> / <u>les</u> ballen – NON-IDIOM.
              at the time of+the exams these students it-ntr.acc / them-3rd.fem.acc.pl dance
              'When students take exams they dance that / those'

        b.       A l'època dels exàmens aquests estudiants *la ballen*     – IDIOMATIC
              at the time of+the exams these students it it-ntr.acc dance
              'When students take exams they suffer'

(8)     a.       M'<u>ho</u> pagaràs                  – NON-IDIOMATIC
              me it-ntr.acc pay
              'You'll pay me for this'

        b.       *Me <u>la</u> / <u>les</u> pagaràs*         – IDIOMATIC
              me it-3rd.fem.acc.sg / them-3rd.fem.acc.pl pay
              'You'll pay for this'

What these data show is that on the one hand, in non-idiomatic expressions the neuter clitic *ho* denotes a meaning that could be explicitly formulated by means of a full propositional form or a non-discrete object (e.g. what you order the students to dance in (7a), or what you owe me in (8a)). The third person feminine accusative plural clitic *les* denotes some particular dances. In each case the semantic content of the clitic is assumed to be dependent on the identification of some relevant antecedent. On the other hand, in idiomatic expressions, the third person feminine accusative singular i-clitic *la* denotes an abstract object: a generic and atemporal situation in (7b), or an imprecise object in (8b). I-clitic *les* in (8b) denotes an indeterminate plural object. In idiomatic expressions, third person clitics have an even more indeterminate and imprecise denotation than the pure neuter pronominal *ho* (Mariner 1975:56, 60), in addition to their being more productive (see note 4). Therefore, they cannot be

appropriately described as being expletive, since even though languages usually have only one or two 'neuter' pronouns (e.g. *ho* lit. it-ntr.acc), they license a large ontology of 'abstract' objects by means of other clitic or weak pronouns (e.g. *la* lit. it-3$^{rd}$.fem.acc.sg, or *les* lit. them-3$^{rd}$.fem.acc.pl).

Moreover, following relevant studies on pronouns, it is important to consider the fact that clitics are defined as being severely deficient pronouns (Cardinaletti and Starke 1999). Syntactically, they impose an adjacency requirement to the V; they cannot occur in coordinate structures; nor can they be modified by adverbs; nor can they occur alone in peripheric positions. Semantically, they are also deficient pronouns, as their antecedent must be prominent in the discourse and can refer to non-human entities. Besides these properties, i-clitics occurring in idiomatic constructions show further more marked deficiencies (Bibis 2002) –the reasons for this higher-level deficiency being the following: i-clitics are

(9)  (i)  reduced to a limited number of lexical items (a maximum of five items in the set of Romance languages being considered, which systematically exclude first and second person pronouns),

(ii)  highly selected with respect to their verbal host,

(iii)  mutually exclusive,

(iv)  cannot be doubled,

(v)  not assigned conceptual θ-roles in semantic terms,

(vi)  non-referential, and

(vii)  interpreted as free variables which are anaphoric to an abstract (pragmatically determined) object (cf. Asher 1993).

Paying attention to these considerations, this paper aims: (i) to describe what the specific syntactic and semantic properties of i-clitics in idiomatic constructions are, as compared to other clitics, (ii) to identify the contribution of i-clitics to the meaning of idiomatic expressions, and (iii) to understand the reason why, within the whole set of clitics available in Romance syntax, such a specific, limited subset is able to occur in (1) and is able to denote abstract semantic objects. Finally, this paper aims (iv) to account for the semantic similarities between i-clitics and object nouns in idiomatic constructions, in the sense that they both trigger abstract object denotation.

In Section 2, I shall introduce the two different types of clitic pronominals that occur in Romance idiomatic constructions: thematic clitics and i-clitics. Sections 3 and 4 contain a full description of the syntactic and semantic properties characteristic of i-clitics in idiomatic constructions that show the pattern in (1). Section 5 will be devoted to present my analysis of i-clitics at the syntax – semantics mapping, with special reference to the notion of semantic incorporation. In Section 6 the syntactic and semantic properties of i-clitics in V + Cl idiomatic constructions will be compared to those characteristic of bare object nouns in idiomatic expressions with the schema V + (D) + N.

**2.** TWO TYPES OF CLITIC PRONOMINALS. Among the clitic pronouns that occur in an enclitic or in a proclitic position (Burzio 1986, Cardinaletti 1999, Mascaró and Rigau 2002) affixed to the verb of an idiomatic expression, a distinction should be made between those weak pronouns which are associated with a conceptual-thematic requirement of the verb, and those which are not, here termed the class of i-clitics.

Let us consider the idioms in (10), which contain the third person masculine accusative clitic *lo* and the third person dative clitic *li* respectively.

(10)    a.    *estar per tancar-<u>lo</u> / estar perquè <u>el</u> tanquin*
            be for lock him-3$^{rd}$.masc.acc.sg / be for him-3$^{rd}$.masc.acc.sg lock+SUBJ
            'be insane'
    b.    *donar-<u>li</u> un camí* (a algú)
            give him/her-3$^{rd}$.dat.sg a way to somebody
            'counsel'

It is important to notice that the underlined clitics are syntactically and semantically associated with argument positions of the verbs *tancar* 'close' and *donar* 'give'. Accordingly, they are interpreted as thematic arguments of the corresponding predicates. Furthermore, these clitics are interpreted as referring to some specific individuals, in accordance with contextual accessibility conditions. As the examples in (11) and (12) show, these clitics allow right and left dislocation, and substitution for other clitic pronouns with different person-gender-number formal features, mainly first and second person pronominal clitics.[8]

(11)    a.    *Està per tancar-<u>la</u>$_i$ , la Maria* $_i$
            be for lock her-3$^{rd}$.fem.acc.sg the Maria
            'Mary is insane'
    b.    Noi$_i$, *estàs perquè <u>et</u>* $_i$ *tanquin*
            guy  be for you lock+SUBJ
            'You, guy, you are insane'

(12)    [pro]$_i$ No sé què fer, *dona'm*$_i$ *un camí*
        not know what do give me a way
        'I don't know what to do, counsel me'

Contrasting with third person accusative masculine and dative pronominal clitics, i-clitics, as those exemplified in (2) - (6), neither saturate any conceptual-thematic argument of the verbal predicate they are affixed to, nor absorb any conceptual-thematic argument of the predicate they occur with. In addition, i-clitics do not allow substitution for first and second person pronouns. Notice the contrast between (7b) and (13).

(13)    #A l'època dels exàmens aquests estudiants <u>em</u> / <u>et</u> ballen
        at the time of+the exams these students me / you dance

An additional observation relates to the fact that some idiomatic expressions allow alternative i-clitics as free variants, but notice that, although this alternation does not contribute to referentiality, it does codify different abstract object denotations. The examples in (14) illustrate the fact that the idiom *enginyar-se-les* (lit. think+CL+them 'manage') allows *les / ho* alternation, and (15) admits *la / ho* substitution. The examples in (16) illustrate Spanish dialectal variation with regard to a specific idiomatic expression. What is interesting is the clitic alternation between neuter *lo,* third person feminine accusative singular *la*, and an empty pronoun (i.e. an invisible abstract clitic pronoun, Keyser andRoeper 1992). The examples in (17) illustrate Catalan free variation between feminine *la / les* pronominal accusative forms. The

examples in (18) illustrate French historical variation between an empty pronoun and third person accusative feminine singular clitic *la*. The examples in (19) show an Italian / French alternation between *la / en*; and, finally, (20) illustrates an Italian / Catalan alternation between optional *ci / ∅*.[9]

(14)  a.    No sé pas com *se <u>les</u> enginya*, però sempre se'n surt (EC)
        not know not how CL them-3rd.fem.acc.pl thought but always CL succeeds
        'I don't know how (s)he manages, but (s)he always succeeds'

     b.    *S'<u>ho</u> ha enginyat* tan bé, que ha convençut tothom (EC)
        CL it-ntr.acc has thought so well that has convinced everybody
        '(S)he managed so well, that (s)he convinced everybody'

(15)  a.    L'any que no plou, els pagesos de secà *<u>la</u> passen prima* (RM)
        the year that not rain the peasants of dry+land it-3rd.fem.acc.sg pass thin
        'Without rain dry land farmers starve'

     b.    *<u>Ho</u> passen prim*
        it-ntr.acc pass thin
        'They starve'

(16)  SPANISH
     a.    *Pasar(se)<u>lo</u> bien* (Castillian Spanish)
        pass CL it-ntr.acc well
        'have a good time'

     b.    *Pasar<u>la</u> bien* (Mexican Spanish)[10]
        pass it-3rd.fem.acc.sg well

     c.    *Pasar bien* (Uruguayian Spanish)
        pass well

(17)  CATALAN
     a.    *Me <u>la</u> pagaràs*      (=8b)
        me it-3rd.fem.acc.sg pay
        'You'll pay for this'

     b.    *Me <u>les</u> pagaràs*
        me it-3rd.fem.acc.pl pay

(18)  FRENCH
     a.    *Échapper belle* (XVIIe s., Rey and Chantreau 1979:350)
        escape nice
        'escape from a danger'

     b.    *<u>L'</u>échapper belle*
        it-3rd.fem.acc.sg escape nice

(19)  ITALIAN
     a.    *Saper<u>la</u> lunga*
        know it-3rd.fem.acc.sg long
        'be smart'

  FRENCH
     b.    *<u>En</u> savoir long*
        CL.gen know long

(20)    ITALIAN
    a.      *Perder(ci) il sonno*
            loose CL.loc. the dream
            'awaken'
    CATALAN
    b.      *Perdre el son*
            loose the dream

These clitic alternations suggest strongly that the linguistic contribution of i-clitics to the meaning of idiomatic expressions, if any at all, is distinct from that of conceptual-thematic clitics. For example, inherent accusative clitics are not interpreted in accordance with the referential values usually associated with the set of person, gender, and number features which, as has been argued independently in the theory of grammar, formally define accusative pronoun morphemes.[11]

These data also illustrate the fact that non-referentiality is not exclusive of third person feminine accusative pronouns. The adverbial weak pronoun *hi,* which shows a deictic use in non-idiomatic expressions (Kayne 2006), lacks any specific spatio-temporal referential property when occurring in the lexicalization pattern schematized in (1). Accordingly, Italian *ci* can alternate with a null form −as illustrated in (20)− and the Catalan idiomatic expressions in (21) −which both select the i-clitic *hi*− do not denote any place to run into nor any event to fall onto; rather, they denote an abstract spatio-temporal location.

(21)    a.      Dius que vols la feina d'aquí una hora? *Ara hi corro      – ara hi corro*
                say that want the work of here an hour now CL.loc run      'be joking'
                'Are you saying you need this finished within an hour? Are you joking?'
        b.      Ho comprens tu? Jo no *hi caic* (IEC)              – *caure-hi* 'make sense'
                it understand you I not CL.loc fall
                'Do you understand it? I doesn't make sense to me'

A further consideration that is interesting to raise at this point is that, assuming the hypothesis that i-clitics (such as Catalan *la / les / ho / en / hi*) do not satisfy specific conceptual-thematic requirements of verbal predicates, some flexibility regarding the combination of these clitics with verbal heads is expected, and in fact this is what we find. The examples in (22) illustrate six different V + Cl idiomatic expressions with the transitive verb *dir* 'say', showing that the contribution of each clitic form to the meaning of the whole idiomatic construction (mainly, the locative clitic *hi*) should be dissociated from the conceptual-thematic selection of the verbal head.

(22)    a.      *Dir-la grossa*
                say it-3$^{rd}$.fem.acc.sg big                          'talk nonsense'
        b.      *Dir-les pels descosits*
                say them-3$^{rd}$.fem.acc.pl for the unsewed    'chat'
        c.      *Dir-s'ho tot*
                say CL it-ntr.acc all                                  'disagree'
        d.      *Dir-ne una de fresca* (a algú)
                say CL.part a of fresh to somebody            'offend'
        e.      *Dir-hi la seva*

say CL.loc the own                          'intervene'

Bearing in mind the generative assumption (Borer 1984, Kayne 1975) that clitics that bear thematic roles are the ones that are linked to an NP / DP maximal projection −which is the real argument of the verb− the clitics illustrated from (14) to (20) are non-thematic arguments, since they are not linked to a full NP / DP argument position. Therefore, the assumption that the meaning of a pronominal clitic comes from the particular reference of the maximal projection to which it is linked does not seem to hold in the case of i-clitics.

In the following sections, I shall describe the set of syntactic and semantic properties characterizing the strong deficiency of i-clitics postulated at the end of Section 1, and I shall introduce the hypothesis that the non-thematic and non-referential status of i-clitics are the result of their not bearing an anaphoric relationship with any argumental and referential DP. I-clitics often show no syntactic relation with any explicit DP, and should a linking process apply between the clitic variable and some other constituent, this XP is never internal to the VP projection, but rather is a discourse (hidden) topic or a (right) dislocated constituent.


**3.** SYNTACTIC PROPERTIES.
**A.** I-clitics are fixed for each idiomatic expression with a fixed case. As shown from (14) to (20), some occasional alternations involving the feminine accusative clitics, the neuter, the partitive, the locative or an empty pronominal form can be observed, but the presence, the absence, or the available alternations among i-clitics are also fixed. In fact, the set of affixes allowed by agglutination into a verb stem is always highly restricted, and depends on the selectional requirements of the verbal head.

In general, with some few exceptions (e.g. *dinyar-la / *dinyar* 'die'), a verb minus its i-clitic is still a well-formed verb.


**B.** I-clitics are mutually exclusive. In addition, accusative and neuter clitics are incompatible with a bare object noun.

This syntactic complementary distribution would appear to support Keyser and Roeper's (1992:89) abstract clitic hypothesis, according to which "each [English] verb is associated with a category-neutral abstract clitic position" which would be filled by i-clitics, bare nouns, etc. However, inherent locative *hi* can co-occur with an object NP (e.g. *afegir-hi salsa* lit. add+CL sauce 'exaggerate') or DP (e.g. (22e)), inherent partitive *ne* can co-occur with a QP (e.g. (22d)), and Italian *ce* can co-occur with third person accusative feminine clitic *la* (e.g. *farcela* lit. fare + ce + la 'manage', *avercela* lit. avere + ce + la 'have something indeed'), thus suggesting either that these two constituents do not compete for the same syntactic position, or that the abstract clitic hypothesis does not provide a thorough explanation.

Nevertheless, it should be noted that an i-clitic can combine with another clitic pronoun that is bound to a thematic argument position, for example, the goal dative object *hi*, as illustrated in (23b).

(23)   a.      *Portar-la votada* (a algú)
                take it-3$^{rd}$.fem.acc.sg voted to somebody        'have it in for someone'
        b.      *La*$_{Inherent}$ *hi*$_{Dative-Goal}$ *porten votada* des que es vol presentar com a degana
                it-3$^{rd}$.fem.acc.sg her take voted since that CL wants present as a dean
                'They've had it in for her since she decided to be the dean'

**C.** Some i-clitics have overtly incorporated into the verb.[12] This diachronic clitic incorporation is shown explicitly in some Valencian Catalan idioms in which the clitic *hi* occurs marginally prefixed to the verb (e.g. the verbal idiom *valer-s'hi* lit. be worth+Cl 'be allowed' −which alternates with *hi valdre's*− and other verbal idioms such as *hi veure's* lit. Cl see Cl 'manage to see', and *n'hi haure* lit. Cl have 'have').[13] Most commonly, however, i-clitics are inflectional affixes for case, person, number and gender which are close to derivational affixes.[14]

**D.** I-clitics cannot be extracted. Notice the following examples from Mendívil (1999:552, ex. (18)):

(24)  SPANISH
      a.     *¿<u>Qué</u> has armado?             - *armar<u>la</u>* 'make trouble'
           what has armed
      b.     *¿<u>Qué</u> has corrido?           - *correr<u>la</u>* 'live it up'
           what has run

**E.** Any attempt to substitute the accusative clitics *la / les* for a nominal in a passive construction provides an ungrammatical output.

(25)    a.     *<u>La bronca</u> va ser armada per la seva mare  - *armar-<u>la</u>* 'make trouble'
           the row PAST be armed by the her mother
    b.     *<u>Les situacions</u> van ésser passades morades - *passar-<u>les</u> morades*
           the situations PAST be passed purple       'be in trouble'

**F.** The antecedent of the clitic cannot occur at the left periphery of the clause as focus, since this would imply having specific foregrounded information as an antecedent −which is not the case, as explicitly illustrated in (26).

(26)    a.     *LA BRONCA, *va armar-<u>la</u>* sa mare     - *armar<u>la</u>* 'make trouble'
           the row       PAST arm it-3$^{rd}$.fem.acc.sg her mother
    b.     *A FER EL QUE DIU ELL, *correm-<u>hi</u>* tots  - *correm-hi* 'let's do it'
           to make the what says he    run CL.loc all

The last four properties strongly suggest, together with A, that the clitic is merged with the verb from the first steps of the derivation, checking immediately its selectional and case features (Masullo 2004), and providing a means for advancing an (accusative, partitive, or locative) argument to a primary internal argument position.

**G.** I-clitics of idiomatic constructions never occur in clitic doubling structures. Therefore, the i-clitic, as an overt incorporated nominal or classifier in polysynthetic languages, is the true syntactic argument of the verb, and any attempt to duplicate the clitic provides an ungrammatical result. See Mendívil (1999:552, ex. (19)):

(27)    SPANISH                                - *arreglárselas* 'manage to'
       *Nos <u>las</u> arreglamos* (*<u>las dificultades</u>) *como pudimos*
       us them-3$^{rd}$.fem.acc.pl manage the difficulties as could
       'We managed as well as we could'

The fact that clitics do not agree with an overt QP / NP is also illustrated in (28), which includes an overt topic specified at the left or at the right peripheral position of the sentence containing the idiom. None of these examples allow a coindexing relationship between the topic and the clitic and, therefore, a syntactically determined coreferential interpretation is not accepted.

(28)  a.    Con <u>tantas dificultades</u> $_i$, *nos <u>las</u> $_j$ arreglamos* como pudimos (*$_{i...i}$)
            with many difficulties us them-3$^{rd}$.fem.acc.pl manage as could
            'With so much trouble, we managed as well as we could'

      b.    *Nos <u>las</u> $_i$ arreglamos* como pudimos, ante <u>las situaciones que se habían creado</u> $_j$ (*$_{i...i}$)
            us them-3$^{rd}$.fem.acc.pl manage as could in face of the situations that CL had created
            'We managed as we could, given the situations that had been created'

**H.** I-clitics cannot be interpreted as syntactically bound to a genuine quantifier. Thus, similar to what occurs in (28a), (29a,b) do not allow the coindexed and bounded reading between the universal or *wh-* quantifier and the inherent accusative clitic, whereas (29c) licenses the bound variable reading of the thematic accusative clitic by the existential quantifier.

(29)  a.    <u>Tota la vida</u> $_i$, que <u>*la*</u> $_j$ *ballem* (*$_{i...i}$)          - IDIOMATIC
            whole the life that it-3$^{rd}$.fem.acc.sg dance
            'We are always suffering'

      b.    *Quina tarda* $_i$! *Com* <u>*la*</u> $_j$ *ballem*! (*$_{i...i}$)          - IDIOMATIC
            what afternoon how it-3$^{rd}$.fem.acc.sg dance
            'What an afternoon! How are we suffering!'

      c.    <u>Algunes dances</u> $_i$, <u>les</u> $_i$ ballarem a final de curs          - NON-IDIOM.
            some dances them-3$^{rd}$.fem.acc.pl by end of course
            'We are going to dance some dances by the end of the course'

In (29a-b) the existential import associated with the V + Cl construction is not due to an existential quantification over specific, individuated entities; rather, it is associated to the fact that, under a neodavidsonian analysis of the meaning of sentences, all sentences describe events of a specific type and all verbs translate as predicates of events (Parsons 1990, Herburger 2000). Accordingly, the meaning associated with (29a) is: there are events (subevents of the whole life) according to which, for all events that are events of suffering, the experiencer of those events is the first person plural.

My analysis of i-clitics in Section 5.2 shows that these cannot be interpreted as variables bound to a genuine quantifier because the postulated incorporation of the clitic leads to a suppression of a theme argument (see note 19, below). It is assumed that, once incorporated, the clitic becomes a qualifier of the predicate that cannot be bound independently.

**I.** Some idiomatic expressions show a structural pattern with an additional XP constituent co-occurring with the clitic pronominal. This constituent is either an obligatory adjunt of the idiomatic expression itself (e.g. *engegar-les <u>sense engaltar</u>* lit. throw-them without face 'say something straight out'); an obligatory quantifier (e.g. *no tenir-les <u>totes</u>* lit. not have-them all 'uncertain, worried'); an obligatory nominal NP /

DP constituent (e.g. *afegir-hi <u>salsa</u>* lit. add-LOC CL sauce 'exaggerate', *ballar-hi <u>el diable</u>* lit. dance-LOC CL the devil 'disagreement'); an obligatory prepositional complement (e.g. *anar-se'n <u>a can Pistraus</u>* lit. go-CL to house Pistraus 'fail'); or the predicate of an adjoined small clause (as exemplified in (30) and (31)).

(30)    a.      *Ballar-la$_i$ grassa$_i$*
                dance it-3$^{rd}$.fem.acc.sg fat.fem.sg        'be rich'
        b.      *Saber-la$_i$ llarga$_i$*
                know it-3$^{rd}$.fem.acc.sg long.fem.sg      'be smart'
        c.      *Mamar-se-les$_i$ dolces$_i$*
                suck CL them-3$^{rd}$.fem.acc.pl sweet.fem.pl 'be trouble free'
        d.      *Cantar-les$_i$ clares$_i$*
                sing them-3$^{rd}$.fem.acc.pl clear.fem.pl     'be outspoken'

(31)    SPANISH
        a.      *Hacerla$_i$ buena$_i$*
                make it-3$^{rd}$.fem.acc.sg good.fem.sg     'get in trouble'
        b.      *Tenerla$_i$ jurada$_i$* (a alguien)
                have it-3$^{rd}$.fem.acc.sg sworn.fem.sg to someone 'have it in for someone'
        c.      *Pasarlas$_i$ canutas$_i$*
                pass them-3$^{rd}$.fem.acc.pl tough.fem.pl    'have a tough time'
        d.      *Prometérselas$_i$ muy felices$_i$*
                promise CL them-3$^{rd}$.fem.acc.pl very happy.fem.pl
                                      'be in for a very pleasant surprise'

It is interesting to notice that, when this XP has an adjectival form, it must agree in gender and number with third person accusative clitics, thus suggesting that the clitic is the external argument (that is, the subject) of the adjective (Stowell 1983), and the internal argument (that is, the object) of the verb.[15] Agreement between the clitic and the adjective may be a consequence of a coindexing relationship between these two constituents. Furthermore, these data suggest that whatever happens to these clitics at the syntax-semantics interface (cf. clitic incorporation in Section 5.2), traces of clitics retain the φ-features (PNG features) of the original clitic. (since there is overt agreement).

      I would like to insist on two further aspects. First, i-clitics can show grammatical agreement with an adjoined adjective (but not with a fully specified argumental nominal, as illustrated in (27)), and second, these adjectives never have a resultative interpretation.[16] Adjectives occurring in the pattern in (1) are always adjuncts.

      It should further be noticed, contrary to what Mendívil (1999:553) indicates, that the adjectival adjuncts in (30) and (31) can occasionally be extracted, as expected from the syntax of non-idiomatic expressions showing this same pattern:[17]

(32)    SPANISH
        a.      ¡Buena la has hecho!                        (see (31a))
                good it-3$^{rd}$.fem.acc.sg has done
                'You've got in trouble'
        b.      Con lo canutas que las ha pasado, no creo que vuelva  (see (31c))
                with the tough that them-3$^{rd}$.fem.acc.pl has passed not think that comes
                back

'(S)he has been through so much trouble, I don't think (s)he will come back'

Summing up, all these properties point to the following conclusions: (i) i-clitics are selected by specific verbs, i.e. each verbal head occurring in the schema in (1) selects particular i-clitics; (ii) i-clitics are mutually exclusive, but inherent (accusative) clitics can occur with a thematic (goal) clitic, and a partitive or a locative i-clitic can occur with an object NP / DP complement; (iii) the clitic affixed to the verb is an object argument of the verb (either a direct object, or an internal or external oblique object; cf. Rice 2003); (iv) i-clitics cannot be extracted, they cannot passivize, they cannot be associated to a focused constituent, nor can they agree with a fully specified argumental DP / NP in a double object structure; (v) i-clitics cannot be interpreted as variables bound to a quantifier phrase because at some point in the derivation they form a bond unit with the verb; and, finally, (vi) i-clitics can formally agree in gender and number with an adjectival modifier in postverbal position because at some point in the syntactic computation of these idioms the clitic VP-argument agrees in gender and number with an adjoined adjectival predicate.

In the next Section I shall consider the most significant semantic properties of the set of case marked i-clitics that occur in the structural schema in (1).

**4.** SEMANTIC PROPERTIES.
**A.** As already pointed out, a characteristic semantic property of V + Cl idiomatic expressions is the fact that i-clitics are not licensed as conceptual arguments of the verbal predicate under a conceptual approach to thematic relationships (Jackendoff 1990).

If we consider once again the idiom *ballar-la* 'suffer', it should be noticed that the clitic cannot be interpreted as the theme of the dancing event. Let us recall the contrast between (7a-b), repeated here for convenience.

(33)    a.    A l'època dels exàmens aquests estudiants <u>ho</u> / <u>les</u> ballen  – NON-IDIOM.
                 at the time of+the exams these students it-ntr.acc / them-$3^{rd}$.fem.acc.pl
                 dance
                 … i aquells d'allà també ho fan
                 and those of there also it-ntr.acc do
                 'When these students take exams they dance that / these dances, and those students over there do it too'
    b.    A l'època dels exàmens aquests estudiants *la ballen*     – IDIOMATIC
                 at the time of+the exams these students it it-$3^{rd}$.fem.acc.sg dance
                 …# i aquells també ho fan
                 and those also it-ntr.acc do
                 'When students take exams they suffer, # and those students do it too'

Notice that the merged clitic in (33b) is able to check or discharge syntactically the selectional D feature, as well as the accusative case feature of the verb, but cannot be interpreted as a conceptual theme. My hypothesis is that this happens because, at the level of LF, the verb together with the i-clitic form a new predicate, and the i-clitic cannot be interpreted as expressing a conceptual theme of the event. Thus, whereas in (33a) *aquests estudiants* 'these students' is interpreted as the agent of the verbal predicate (as proved by the fact that the sentence can continue with an agentive

proform *fer-ho* 'do it'), in (33b) it is interpreted as the experiencer, and the idiomatic sequence cannot continue with the same verbal proform.[18]

**B.** The verbal head in V + Cl idiomatic constructions can either be transitive or intransitive, but the clitic is an intransitivizer morpheme in the sense that it takes as input either a transitive or an intransitive verb and provides as its ouput a one-place predicate with a free variable. Furthermore, i-clitics may lower the valence of the verb by one, and demote the thematic assignment to the subject by one.[19]

Let me consider the intransitivizer effect of i-clitics in three distinct idiomatic expressions: *brandar-la* 'dance, walk around', *ballar-la* 'suffer', and *pintar-hi* 'not to have a say in something'. The accusative feminine singular clitic *la* combines with the transitive verb *brandar* 'take (a weapon)' and saturates its internal argument, giving rise to a verbal predicate with a free variable in argument position. This same clitic combines with an intransitive verb, such as the agentive unergative *ballar* 'dance', and translates it into an intransitive predicate: the non-agentive predicate *ballar-la* 'suffer', with a free variable in cognate object position. The proposal that i-clitic *la* is an affix that triggers semantic intransitivization is proved by the fact that, once the i-clitic is unified with the verb, the subject's thematic argument is demoted (i.e. V + *la* idioms only license experiencer, but not agentive subjects). Similarly, the locative clitic *hi* can combine with both transitive and intransitive verbs and the output is always an intransitive predicate. For example, the transitive agentive *pintar* 'paint' is transferred to the intransitive non-agentive *pintar-hi* 'not to have a say in something', by extending the argument structure of the lexical verb with an oblique object that is not part of its argument structure and simultaneously introducing a thematic demotion to the external argument of the verbal predicate. In other words, the internal argument of the verbal predicate is replaced with the semantic contribution of the affixed locative clitic, and the thematic interpretation of the nominal subject is demoted.[20]

The V + Cl bond, similar to type I and type II noun incorporating languages, is both syntactic and semantic (cf. Mardirussian 1975, Mithun 1984). In syntactic terms, the i-clitic 'closes' the transitivity of the verb by checking its selectional and case features. In semantic terms, the i-clitic is a variable which is unified with the verb that selects it, either by pure argument selection or by verbal modification, thus providing a one-place predicate as output.[21] Consequently, idiomatic expressions including i-clitics are generally used to describe habitual, or ongoing activities (Mithun 1984:850). In Section 5.2 I shall account for this syntactic and semantic bond as being the output of an incorporating process.

**C.** As already indicated, i-clitics occurring in the schema in (1) lack a referential interpretation. Their meaning, therefore, never corresponds to individual entities or to individual locations. Consider (34).

(34)  a.  Avui porta el cotxe nou: es nota que *l'abilla*    - *abillar-la* 'wealthy'
         today drives the car new CL note that it-3<sup>rd</sup>.fem.acc.sg prepare
         '(S)he drives the new car today. (S)he's obviously wealthy'

     b.  Aquest *se les empesca* totes per no fer la feina    - *empescar-se-les*
         this CL them-3<sup>rd</sup>.fem.acc.pl think for not make the work      'manage'

     c.  Ja *hi ha afegit salsa* altre cop              - *afegir-hi salsa* 'exaggerate'
         already CL.loc has added sauce other time
         '(S)he's exaggerating once more'

These examples show that neither the third person accusative clitic pronouns *la / les*, nor the locative clitic *hi* have referential independency. Hence, although formally speaking third person accusative clitics have φ features, i-clitics appear not to have any sort of referential interpretation. Furthermore, there is no argumental DP in the sentential context of the idiomatic VP which provides the syntactic antecedent of the pronoun.[22] Thus, in (34c), as in (21) and (22e), *hi* is not coindexed with any VP internal or external prepositional or adverbial object that refers to any specific antecedent located in time or space.

The account of this non-referentiality follows from the fact that, syntactically, the clitic does not agree with any argumental antecedent that refers to any specific individual. Semantically, i-clitics show a close connection to incorporated nouns in polysynthetic languages, and, like them, i-clitics are non-referential, and non-individuated (Mardirussian 1975, Mithun 1984, 1986).

**D.** The domain in which i-clitics are interpreted is larger than the idiomatic expression containing the i-clitic. The examples in (35) show that the relevant context for interpreting the clitic can be phonetically realized in a previous or subsequent discourse clause, or in a topicalized position. In (35) the underlined constituents provide the most accessible contexts for identifying a pragmatic antecedent for the i-clitic outside syntax (Reinhart 1983).

(35)    a.    <u>En aquesta operació</u> *li hi va la vida* (RM)
            in this operation him/her CL.loc goes the life
            '(S)he is going to risk his/her life in this operation'
    b.    He fet una <u>confitura</u>! *Us en llepareu els dits* (EC)
            have made a jam us CL.gen lick the fingers
            'I've prepared such a jam! It's delicious!'
    c.    Si vol <u>anar-se'n de viatge</u>, ara mateix *ho té pelut*, perquè no disposa de diners
            if wants go CL of trip now right CL.ntr.acc has difficult because not has of money
            'If (s)he wants to travel, right now (s)he is in trouble, for (s)he doesn't have any money'

However, most often the antecedent of the clitic is left unexpressed, and it must be recoverable on the basis of underdetermined contextual and discourse information accessibility.[23] Thus, consider (36). The example in (36a) shows the use of the idioms *clavar-les pels descosits* 'be outspoken', with no linguistically specified antecedent for the clitic, whereas (36b) introduces an *ad sensum* antecedent-anaphor relationship between the meaning of being *xerraire* 'gossipy', and *les*, which apparently denotes 'the things (s)he talks'.

(36)    a.    Aquest és dels que <u>*les clava pels descosits*</u>
            this is of+the that them-3rd.fem.acc.pl nail by+the unsewed
            'He is outspoken'
    b.    És un <u>xerraire</u>, *les clava pels descosits* i no sé com tenim humor d'escoltar-lo (RM)
            is a gossipy them-3rd.fem.acc.pl nails by+the unsewed and not know as have humour of listen him
            'He is gossipy. He is outspoken. I'm amazed you are listening to him'

**E.** I-clitics, in spite of not being conceptual arguments of the verbal predicate they are affixed to, and in spite of not having a referential denotation, are not expletive.

Bibis (2002:106-7) claims that clitic pronouns in idiomatic expressions behave expletively. His thesis is based on the observation that V + Cl in Modern Greek "involve deficient / expletive or marginal clitics". However, although it might be admitted without discussion that i-clitics show strong syntactic deficiency, and that their semantics is different from that corresponding to regular clitics, they cannot be claimed to be expletive, if by means of this term we refer to a syntactic constituent which makes no semantic contribution to the interpretation of the complex string containing this constituent (Espinal 1992). On the one hand, i-clitics are the required part of a lexicalization pattern which is common in various languages, not only in Romance. On the other hand, they are not devoid of a linguistic contribution in syntax. I-clitics have categorial and case features which are syntactically relevant, and denote an ontology of abstract objects which are semantically relevant.

The problem we are now left with is to identify exactly what the semantic contribution of i-clitics is, since their semantics is not determined by the set of morphosyntactic features by means of which they are syntactically defined.

**F.** From a purely descriptive point of view, it happens to be the case that i-clitics are expressions that denote different types of abstract objects, have no spatio-temporal location, no causal efficacy, and their denotation cannot be perceived.

A common claim made in the linguistic literature on pronouns is that neuter [-G] pronouns (such as Catalan *ho,* and Spanish *lo*; cf. Bosque and Moreno 1990) and partitive clitics (such as Catalan, French and Italian *en/ne*; cf. Todolí 2002, Grevisse 1986, and Renzi 1988) in non-idiomatic expressions denote some type or other of abstract objects.[24]

What is not so straightforward, from a semantic point of view, is how to account for the fact that the rest of the i-clitics (third person accusative clitics, as well as the adverbial / locative clitic) may also denote abstract objects. Therefore, an important issue has still to be resolved: why is it the case that inherent *l*-clitics (third person clitics, Déchaine and Wiltschko 2002), and adverbial clitics may also denote various types of abstract objects? Most plausibly, the fact that Romance has no morphological 'neuter' category associated with the nominal system must be related to the fact that the pronoun system is legible for abstract object denotation.[25]

If we take the case of *l*-clitics into consideration, it should be pointed out that traditional grammarians do not seem to agree on what has been named (i) the 'feminization of the neuter' (by means of which either a neuter form is conceived as feminine; cf. Spitzer 1941, Mariner 1975), or (ii) the use of a feminine form with a 'neuter value' (i.e. an abstract meaning; cf. Rey and Chantreau's 1979:350 assumption regarding the idiomatic expression *l'échapper belle* lit. it escape nice-fem 'escape from a danger').[26] This disagreement suggests that what have usually been conceived of as third person feminine accusative pronominal forms (such as, Catalan and French *la, les*; Spanish *la, las*, as exemplified in (2)-(5)) could well be the remainders of previous neuter forms, which morphologically look like feminines, or examples of feminine forms that express the 'neuter', conceived as an abstract semantic category. This is the reason why apparently feminine clitics are semantically abstract, or denote abstract semantic objects.

Let me now consider the meaning of adverbial or locative clitics (e.g., Catalan *hi,* French *y,* and Italian *ci*). In non-idiomatic expressions these clitic pronominals may

have a deictic use, and may substitute for the complement of a preposition different from *de* 'of', thus referring to concrete locative entities accessible from discourse context. However, in idiomatic expressions they introduce some type or other of abstract entity anaphora: proposition, property, or abstract locative / temporal anaphora (some relevant examples are given in (42), below). Whereas, syntactically, inherent *hi* is either a morphosyntactic marker of unaccusativity, a marker of an external object, or an obligatory adjunct, semantically, it lacks its locative referential properties.[27]

According to Asher's (1993:1) definition of abstract objects, these "are things like propositions, properties, states of affairs and facts". A look at V + Cl idiomatic units, however, suggests that the class of abstract semantic objects is even larger.

Semantically speaking, i-clitics are designed to encode variables which connect to various sorts of abstract objects: propositional objects, generic situations, properties, abstract or indeterminate spatio-temporal locations, as well as indeterminate objects, but they never refer to facts (Delfitto 2004). The set of objects denoted by i-clitics, as has been attested in Catalan, is summarized in (37).

(37)    TYPES OF SEMANTIC OBJECTS DENOTED BY INHERENT CLITICS
    a.    *ho*  - propositions, properties, generic situations, and indeterminate objects
    b.    *les*  - indeterminate semantically plural inanimate objects
    c.    *la*  - indeterminate semantically number neutral inanimate objects, properties, generic situations, and abstract spatio-temporal locations
    d.    *en*  - properties and indeterminate objects
    e.    *hi*  - propositions, properties, and abstract spatio-temporal locations

Let me first illustrate the meaning of the i-clitic *ho.* This pronoun is involved in different anaphoric relationships that lead to various interpretations: in (38a) the neuter pronoun is interpreted discoursively in relation to an abstract propositional object, either the antecedent subordinate clause (the coming back event) or the interrogative clause (the thinking state); in (38b) the neuter pronoun is interpreted in relation to the property expressed by the predicate (being a sort of wizard); and in (38c) the neuter pronoun is interpreted as having a discourse relationship to an indeterminate empty topic (denoting a generic situation) which is assumed to be accessible from context. In this sense, i-clitics are used to background known information in discourse, a characteristic semantic property of type III noun incorporating languages (Mithun 1984).[28]

(38)    a.    [Que et penses [que tornarà?]] *Ni somiar-ho*
        that you think that come+back not dream it-ntr.acc
        'Do you think that (s)he's going to come back? No way'
    b.    Aquell home era, *per dir-ho així*, una mena de bruixot
        that man was by say it-ntr.acc thus a sort of wizard
        'That man was, I would say, a sort of wizard'
    c.    [Top e ] *Ho té pelut* per guanyar les eleccions
        it-ntr.acc has hairy to win the elections
        'It's unlikely that (s)he wins the elections'

Similarly, the unsaturated clitic *les* always connects beyond he computational system to some indeterminate (semantically plural) inanimate object, which only marginally

occurs explicitly in discourse topic position, but most often requires to be accessible from available discourse information. This means that the topic semantically saturates the accusative plural clitic pronominal by denoting a plural object which only occasionally is made explicit, as in (39a), but quite often it is to be inferred from the discourse context: maybe 'misfortunes' in (39b) or 'tall stories' in (39c).

(39)    a.    *Se <u>les</u> enginyen* totes, <u>les maneres de no pagar a la Seguretat Social</u>
        CL them-3$^{rd}$.fem.acc.pl think all the manners of not pay to the security social
        'They always find the way to avoid paying the SS'

    b.    [$_{Top}$ e ] Aquest home és dels que <u>les</u> *veu venir*. Sempre encerta el que passarà
        this man is of+the that them-3$^{rd}$.fem.acc.pl sees come always guesses the that will happen
        'This guy knows anyone's guess. He always hits the nail on the head'

    c.    No li facis gaire cas. [$_{Top}$ e ] Té el mal costum de *dir-<u>les</u> pels descosits*
        not him-3$^{rd}$.dat.sg make much case has the bad habit of say them-3$^{rd}$.fem.acc.pl by+the unsewed
        'Take no notice of him/her. (S)he usually talks carelessly'

What is interesting about (38c) and (39b,c) is that the topic corresponds to a salient discourse entity which can be linguistically explicit or pragmatically inferred.

    Another clitic that most often provides resumption of a phonologically unrealized dislocated topic is the third person accusative singular clitic *la*. Again, an empty topic must be assumed in (40) in order to identify, from salient discourse information, the content resumed by the pronoun: maybe 'a punishment' in (40a), the property of being rich in (40b), or an imprecise or generic situation in (40c).

(40)    a.    [$_{Top}$ e ] Sempre *se <u>la</u> carrega*, pobra noia!
        always CL it-3$^{rd}$.fem.acc.sg loads poor girl
        'She always gets into hot water, poor girl!'

    b.    Avui porta el cotxe nou: [$_{Top}$ e ] es nota que *<u>l'</u> abilla* (=34a)
        today drives the car new CL note that it-3$^{rd}$.fem.acc.sg prepare
        '(S)he drives the new car today. (S)he's obviously wealthy'

    c.    [$_{Top}$ e ] Au, fins aviat, i que *te <u>la</u> passis bé*!
        until now and that you it-3$^{rd}$.fem.acc.sg pass well
        'See you soon. Have a good time!'

Clitic *en* resumes a topic, maybe a phonetically empty topic, which, depending on accessible discourse information, can either denote a property ('brico' in (41a)), or an indeterminate object ('misfortunes' in (41b)).

(41)    a.    *No <u>en</u> sé ni un borrall,* <u>de bricolatge</u>
        not CL.gen know nor a bit of brico
        'I haven't a clue about brico'

    b.    [$_{Top}$ e ] Es va quedar viuda molt jove amb tres criatures i *<u>n'</u> ha passat de totes*
        CL PAST rest widow very young with three children and CL.gen has passed of all

'She became a widow very young with three children. She's had a hard time'

Finally, clitic *hi* resumes information regarding propositions ('the speaker making some specific work' in (42a)), properties ('to understand something' in (42b)), or abstract spatio-temporal locations (in (42c).

(42)   a.     Dius que vols la feina d'aquí una hora? [<sub>Top</sub> e ] *Ara hi corro* (= (21a))
          say that want the work of here an hour      now CL.loc run
          'Are you saying you need this finished within an hour? Are you joking?'

       b.     Ho comprens tu? [<sub>Top</sub> e ] Jo *no hi caic* (IEC 1995)
          it-ntr understand you I not CL.loc fall
          'Do you understand it? I don't'

       c.     Em sembla que no juga prou net; [<sub>Top</sub> e ] *hi ha gat amagat* (Raspall and Martí 1984, 1996)
          me appears that not play much clean CL.loc has cat hidden
          'I think (s)he's not playing fair. There is something fishy going on here'

To summarize, i-clitics are not expletive, because they license various sorts of abstract object denotations and the variable introduced by the clitic needs to be interpreted by means of a (hidden) topic or a right dislocated constituent.

**G.** A further semantic property of idiomatic expressions occurring in the pattern in (1) is the fact that these idioms cannot be resumed by factual predicates such as *això va passar* 'it took place', as illustrated in (43a), or by a factual resumptive anaphoric expression *la cosa* 'the thing', as shown in (43b) (Delfitto 2004) .

(43)   a.     *La va dinyar*. #I <u>això va passar</u> en acabar la Guerra Civil
          it-fem.acc.sg PAST die and that PAST happen at end the war civil
          '(S)he pegged out. And this happened after the Civil War'

       b.     *En va fer un gra massa.* #<u>La cosa va succeir</u> després de la manifestació
          Cl.gen PAST make a grain more the thing PAST happened after of the demonstration
          '(S)he went too far. The thing happened after the demonstration'

Similarly, notice the contrast between (44a,b).

(44)   a.     [Aquella dansa]<sub>i</sub> la<sub>i</sub> va ballar. I <u>això va passar</u> al segon bis
          that dance it-fem.acc.sg PAST dance and that PAST happen at+the second encore
          'That dance was danced. And it took place at the second encore'

       b.     Durant la guerra *la va ballar*. #I <u>això va passar</u> també a la seva família
          during the war it-fem.acc.sg PAST dance and that PAST happen too at the {his/her} family
          'During the war (s)he suffered. And it took place too to his/her family'

The factual predicate *això va passar* 'it took place', which requires the external argument to be a factual event, cannot have anaphoric connections with the one-place predicate denoted by V + Cl idiom *ballar-la* 'suffer'. Therefore (44b), in contrast to

(44a), is not a well-formed discourse. "Facts are inherently relational entities possibly involving times and locations" (Delfitto 2004:155), whereas VP idioms containing i-clitics characteristically denote generic and atemporal situations.

Summing up, in this Section I have shown that: (i) i-clitics, semantically speaking, do not discharge a conceptual argument of the verbal predicate; (ii) the predicate corresponding to the V + Cl schema in (1) is licensed as a one-place predicate, often triggering a thematic demotion of the external argument of the verbal predicate; (iii) i-clitics lack a referential interpretation, and their semantics is that of a variable; (iv) the domain in which i-clitics are interpreted is larger than the syntactic domain of the idiomatic expression containing the i-clitic, thereby enabling the required context for interpreting the clitic to be phonetically realized in a discourse environment, either in a topicalized or a right dislocated position; (v) i-clitics are not expletive, since they denote different types of abstract, indeterminate, non-specific objects; and, finally (vi) the denotation corresponding to one-place predicates containing an i-clitic is generally incompatible with a factual interpretation.

In the next section, I turn to the syntax-semantics mapping of i-clitics in idiomatic constructions. In Section 5.1 I shall deal with the variable status of i-clitics at the syntax-semantics interface, and in Section 5.2 I shall analyze the semantics of clitic incorporation as a process of complex predicate formation. I shall develop the idea that the variable and the verbal head form a single predicate at the point of semantic interpretation in such a way that i-clitics are interpreted as internal arguments of their incorporating predicate.

## 5. ON THE SYNTAX – SEMANTICS MAPPING OF I-CLITICS.

**5.1.** The relevant question that still remains to be answered is what exactly is the syntactic status of i-clitics.

According to Déchaine and Wiltschko (2002) the notion of "pronoun" is not a primitive of linguistic theory and lexical pronouns of natural languages can be morphosyntactically assigned to (at least) three pronoun types: pro-DP, pro-φP, and pro-NP. Their hypothesis is that, having distinct morphosyntactic status determines the external syntax of pronominals (that is, their distribution as either an argument, a predicate, or both); their internal syntax (showing a D syntax, a N syntax, or both); their binding properties (as either an R-expression, or a variable); and also their semantics (denoting either a definite expression, or a constant). Furthermore, they put forward the specific hypothesis that French *l*-clitics, and by extension Romance third person clitics, are members of an intermediate functional projection φP that intervenes between D and N, and which encodes φ-features (person, gender and number).[29] These pronominals happen to have neither the syntax of determiners nor that of nouns –a behavior which must be related to their distribution (sometimes that of pro-arguments and sometimes that of pro-predicates), and to the referential defectiveness of Romance articles (sometimes expletive, sometimes not; Vergnaud and Zubizarreta 1992).

Consistent with a φP analysis is the fact that inherent *l*-clitics have interpretable φ-features in non-idiomatic constructions, but un-interpretable φ-features in idiomatic constructions. That is, a pro-φ analysis for *l*-clitics captures the fact that, although they inflect for case, person, gender and number, these features are still not always relevant in semantic terms.[30]

A φP analysis for *l*-clitics is interesting as far as it allows us to capture some significant similarities between the semantics of i-clitics and the semantics of noun

incorporated nominals. If inherent *l*-clitics were DP, they could not be incorporated unless the clitic was expletive –which I have shown not to be the case. Moreover, consistent with a φP analysis is the fact that the binding-theoretic status of *l*-clitics is that of a variable.

Bearing in mind the fact that the set of i-clitics that can occur in the schema presented in (1) not only includes *l*-clitics, but also *ho, ne,* and *hi*, which are marked for case but not for φ-features, it must be concluded that the class of i-clitics project either φP / ClP or NP, but never DP (following Déchaine and Wiltschko's 2002 nominal proform typology).

Furthermore, based on the syntactic analysis put forward for languages belonging to the pronominal argument type (e.g. Navajo, Warlpiri, Mohawk, etc.), where the person-number-case morphology internal to a verb actually represents the arguments of the verb (Jelinek 1984, Baker 1996, Hale 2003, among others), i-clitics of idiomatic constructions are analyzed not as markers of agreement morphology but as the syntactic arguments of the verb.[31] The difference with non-inherent referential clitics is that i-clitics denote abstract semantic objects, and, because of this, they must be semantically incorporated into the V they are affixed to. Therefore, whereas noun incorporation is a salient property of polysynthetic languages (as has been extensively argued in the linguistic literature),[32] this paper postulates that it is also a covert phenomenon found in idiomatic constructions in non-polysynthetic languages.[33]

Let us suppose that i-clitics are selected by, and merged with, the verb in the process of deriving the sequences from (2) to (6) in the computational system. All i-clitics carry grammatical case, which reflects the fact that they have different grammatical functions: they correspond either to internal direct objects, to internal oblique objects, or to external oblique objects of the verb (Rice 2003:59). Once the clitic is syntactically merged with the verb, it checks its selectional and case features (Masullo 2004).

Notice further that, similar to clitic pronouns in polysynthetic languages, i-clitics are never bound by a nominal in an argument position since full nominal antecedents never occupy argument positions. Antecedents of i-clitics are always optional –maybe even hidden– discourse topics or right dislocated constituents, with non-argumental functions. This means that i-clitics are neither syntactically nor semantically bound, since the topic, if present, corresponds to accessible discourse information, and i-clitics are both syntactically and semantically free (Heim and Kratzer 1998). I-clitics are not syntactically bound, because they are not c-commanded within their Minimal Governing Category by their antecedent, and they are not semantically bound, because they do not have a binder within that part of the subtree which translates into a predicate abstract.[34] Consequently, the binding status of i-clitics is that of a free variable.

Following Delfitto's (2002) hypothesis on the semantics of pronominal clitics in clitic left-dislocation constructions, the present study supports the claim that pronominal clitics encode the presence of an unsaturated argument position, and that idiomatic constructions with i-clitics encode functional abstraction over argument positions of the verbal predicate.[35] Semantically speaking, V + Cl idiomatic constructions are interpreted as unsaturated expressions with λ-abstracted variables, but they do not encode λ-abstraction over a thematic argument of the verbal predicate, because, as already indicated, i-clitics are never bound by a nominal in an argument position. V + Cl idiomatic constructions encode λ-abstraction over a range of non-referential topics (or right dislocated constituents) that denote an ontology of abstract

objects. This is what would account for the interpretative difference between thematic clitics and i-clitics, as described in Section 2.

Accordingly, the logical status of i-clitics is that of a variable –a variable which may range over propositions, properties, generic situations, spatio-temporal locations, or indeterminate objects, depending on the internal properties of each clitic (see (37)) and the informational content of the most accessible topic. I-clitics are never syntactically coindexed with a referential antecedent. Therefore, idiomatic constructions with a merged i-clitic contribute semantically to encoding formal objects of the sort λμ (...μ...), where μ stands for a metavariable over a variety of abstract semantic objects. They never contribute individual entities, which is the salient characteristic of regular clitics. This means that i-clitics do not semantically saturate the verbal predicate by themselves. Inherent pronominal clitics must combine with information available from the discourse context (sometimes an explicit syntactic dislocated constituent or topic, and sometimes a hidden topic, which may be salient in discourse) in order to identify a conceptual antecedent for the clitic.

Clitics signal that the antecedent is highly accessible, even though it is a covert antecedent.[36] When an explicit topic is not accessible from contextual information, it is assumed that some piece of contextual information will have to be inferred in order to saturate the clitic-related position. When the topic argument of λ is a referential expression, then the coindexed variable also gets a referential interpretation. When the topic argument of λ is not a referential expression, but an abstract or indeterminate object, then the variable denotes an abstract semantic object. Therefore, the content of the topic is crucial in order to infer the final interpretation associated with the variable. Summing up, i-clitics are pure arguments that are syntactically merged with their selecting verb. As for their binding theoric status, at LF they are translated as free variables. V + Cl constructions are unsaturated expressions in the sense that they encode functional abstraction over argument positions of the verbal predicate. They encode λ-abstracted variables which must combine with a (hidden) topic or a (right) dislocated constituent –the argument of a λ-abstract. This topic must not necessarily be explicit in the linguistic context, and might denote some sort of abstract object: a propositional object, a generic situation, a property, an abstract or indeterminate spatio-temporal location, or an indeterminate object. That is, any contextual assessment of an idiomatic expression containing an i-clitic should start with the topic expression, since the antecedent of the clitic is to be identified in discourse following the instructions encoded by the markers of accessibility that each clitic constrains.

**5.2.** Let us now focus on the following question: how are variables in V + Cl idiomatic constructions semantically interpreted? The answer to this question is that free variables denoting abstract objects are interpreted as internal arguments of an incorporating verb and form a complex predicate with that verb.

Since i-clitics share some characteristic semantic properties with incorporated nominals (genericity and non-referentiality; loss of argumental status; expression of background known information; supplementation by more specific linguistic material external to the complex verb, etc.; cf. Mithun 1984, Baker et al. 2004:143, 151), and since I assume a compositional analysis of the syntactically encoded meaning of idiomatic constructions (Nunberg et al. 1994, Marantz 1996, Mateu and Espinal 2007), i-clitics are expected to be incorporated at LF. A free variable argument will be interpreted by incorporation with the selecting verb. According to this assumption, any fruitful study of the semantics of idiomatic constructions is to be based on a clear distinction between compositional meanings which are syntactically transparent (e.g.,

those determined by argument structure representations) and conceptual meanings which are not syntactically transparent (e.g., those non-compositional meanings obtained via general conceptual metaphoric and metonymic processes; cf. Gibbs 1995).

The variable corresponding to the i-clitic is semantically incorporated into a verb either as a restriction of that verb's internal argument, with a unification of selecting features (Farkas and De Swart 2003, 2004), or as a modification of that verb (Dayal 2003). See (45).[37]

(45)   a.     *Unification*          (adapted from Farkas and De Swart 2003, 2004)
             Replace the relevant selecting features *y* of a verbal predicate with the set of features *z* contributed by a clitic argument of the verb

       b.     $\lambda\mu \; \lambda x \; \lambda e \; [\; \mu\text{-V (e)} \; \& \; \theta \text{ (e)} = x \;]$     (adapted from Dayal 2003)

According to (45a) the clitic, which introduces some selectional features but no discourse referent, is unified with the verbal predicate that selects this set of non-referential features. In contrast, the incorporating formula in (45b) expresses relations between individuals and abstract semantic objects. The metavariable argument, which stands for the set of semantic objects denoted by i-clitics (see (37)), is unified with the verb, and the μ-V bond unit forms a complex predicate in semantic terms. The individual argument x that is assigned whatever semantic role θ (experiencer, patient, etc.) is associated with the external argument of the event.[38]

What is crucial in these formulations is the way in which idiomatic constructions with i-clitics resolve into incorporating complex predicates, and clitic incorporation involves the attribution of abstract semantic objects. The clitic, that introduces some selectional features but no discourse referent, unifies with the verbal predicate that selects this set of non-referential features. After semantic incorporation has occurred, both formal φ-features (specifying person, gender and number) and case features (accusative, genitive, locative, etc.) are semantically cancelled out. In this way i-clitic constructions resolve into incorporating complex predicates.

The advantages of this analysis are inferred from its predictive possibilities. First, it supports both the referential hierarchy and the implicational mapping hypothesis postulated independently by Cyrino et al. (2000:59).

(46)   a.     *Referential hierarchy*

| non-argument | proposition | [- human] | | [+ human] | |
|---|---|---|---|---|---|
| | | | $3^{rd}$ | $2^{nd}$ | $1^{st}$ |
| - specif | | | | + specif | |
| [- ref] ← | | ----------------- | | ---→ | [+ ref] |

       b.     *The Implicational Mapping Hypothesis*
             The more referential, the greater the possibility of a non-null pronoun

Since first and second [+ human] clitic pronouns are argumental and specific, they are the highest in the referential hierarchy and cannot occur in the schema in (1). Third-person [- human, - animate] clitics, and those denoting a proposition, are non-arguments from a conceptual-thematic sense, and even non-referential; therefore, they are the lowest in the referential hierarchy, and are the only ones that can freely occur in the schema in (1). In addition, in accordance with (46b), the less referential pronominal items are those that have more possibilities of becoming phonetically null,

and the more referential items are the first to become phonetically substantive. Abstract-denoting clitics show two tendencies: either they allow substitution for a null variant (see the data in (16), (18) and (20)), or add a distinguishable semantic interpretation (see for example the data in (14), (15) and (17)), depending on the semantic constraints encoded by the different clitics and the contents of the most relevant discourse topic.

Second, this analysis predicts the set of properties listed in Sections 3 and 4:
(i) Since i-clitics are syntactic arguments, they are fixed for each idiomatic expression (3A). If they have accusative case, they are mutually exclusive with other accusative arguments, only non-accusative i-clitics can combine with an object noun (3B); they can be neither extracted nor passivized, since case-marked clitic pronouns are argumental affixes (3D, 3E); they cannot occur in clitic doubling structures, since they are never bound by a nominal in an argument position (3G); and they allow the presence of external predicates (3I).
(ii) Since i-clitics are free variables, they do not have a referential content (4C) and cannot occur in the left periphery as focus (3F); they cannot be interpreted as syntactically bound by a quantifier (3H); and, the syntactic domain in which they are interpreted is larger than the idiomatic VP containing the inherent clitic (4D). Coherent also with the analysis that i-clitics are free variables is the fact that they are not expletive (4E) and denote different types of abstract objects (4F).
(iii) Since i-clitics are incorporated, they sometimes even allow overt incorporation into the selecting verb (3C), and they are not licensed as independent conceptual arguments of the verbal predicate (4A). V + Cl constructions denote one-place predicates which encode formal objects that include a metavariable over a variety of abstract semantic objects (4B) and, therefore, cannot be resumed with factual predicates (4G).

So far I have studied the syntax-semantics mapping of i-clitics. I have shown that the cluster of syntactic and semantic properties described in Sections 2, 3 and 4, especially the abstract semantic denotation characteristic of this class of clitics, follows if i-clitics are analyzed as syntactic unsaturated arguments which are submitted to a semantic process of incorporation by means of which the variable codified by the clitic forms a complex predicate with the verb.

A similar analysis has been suggested in the linguistic literature in order to account for the semantics of object nouns in idiomatic constructions and for the fact that, no matter whether they are bare NPs or DPs, object nouns are interpreted as abstract property- denoting objects, not identifying individual entities. In the next Section I will show that there are close parallels between the contribution of i-clitics and that of object nouns to the meaning of idiomatic constructions: semantically speaking, whereas i-clitics are abstract object denoting expressions, object nouns (even bare singular count nouns) are property denoting expressions.

**6.** THE SEMANTICS OF OBJECT NOUNS IN IDIOMATIC CONSTRUCTIONS. Let me now review some of the main arguments that support the hypothesis that object nouns in idiomatic constructions are property-type arguments semantically incorporated into the verbal predicate.[39]

**A.** Object nouns of V + N idioms allow various degree quantifiers in prenominal position, even if the object noun is a bare singular count noun (such as *boca* 'mouth'). Existential and universal quantifiers are never permitted. Consider (47).

(47)   a.     *fer    boca*
              make mouth                                  'whet one's appetite'
       b.     *fer* més/molta/força/massa *boca*
              make more/much/a lot/too much mouth
              'to whet one's appetite [to a certain degree]'
       c.     *fer una/cada boca
              make a/each mouth

Notice that these data are predicted appropriately from the hypothesis that object nouns are interpreted as properties, rather than as kind-denoting objects (Chierchia 1998). In this class of idioms the head noun denotes a semantic entity which can be modified by an expression denoting degree, and can only be quantified by an expression denoting an existential quantification over degrees (Espinal 2004b). In other words, these bare nouns denote properties (cf. the concept of inner mass predicates, postulated by Bosque and Masullo 1998).

**B.** Object nouns of idiomatic constructions, even if preceded by a definite D, denote non-specific, non-referential objects. Unlike what is expected in regular syntax, bare singulars are not unusual in object position of Romance idiomatic expressions. Object nouns of idiomatic constructions look like mass nouns and are interpreted as abstract objects. Let us consider (48) – (50) (see also note 9, above).

(48)   FRENCH
       a.     *conter fleurette*
              say     flower+DIM           'woo (someone)'
       b.     *rouler carrosse*
              wheel car                    'show off (something)'

(49)   ITALIAN
       a.     *battere cassa*
              beat box                     'ask for money'
       b.     *chiudere bottega*
              close store                  'give up'

(50)   CATALAN
       a.     *fer denteta*
              make tooth+DIM               'show off'
       b.     *passar el rosari*
              pass the rosary              'say one's rosary'

The bare singulars in (48), (49), and (50a), as well as the definite nominal in (50b), do not contribute discourse referents and do not denote any specific flower, car, box, store, tooth, or rosary. They must be interpreted together with the predicate they are an object of. But notice that they do not saturate the predicate by argument instantiation; rather, a complex predicate is composed from the properties corresponding to the verbal predicate and those corresponding to the incorporating nominal.

This strategy strongly suggests that at some level of representation, namely at the syntax – semantics interface, object nouns denote properties and form complex predicate units with the target verb.

**C**. Object nouns of VP idioms do not designate a discourse referent, therefore they are not referential expressions (cf. Simatos 1986, 1997). Object nouns, either singular or plural, with either a fixed D or a zero D, have a non-referential interpretation and, accordingly, allow no discourse coreferentiality.

This claim is tested by the impossibility of substituting the object nominal by a left-dislocated constituent (as in (51b)); by the lack of discourse transparency between the object nominal and an overt pronoun in discourse (as in (51c)); by the fact that object nominals do not allow restrictive relative clauses (as in (51d)); and by the impossibility of allowing a question over the object noun (as in (51e)).[40]

(51)   a.     *passar el rosari*
            pass   the  rosary                'say one's rosary'

        b.     #El rosari $_i$, passa-me'l $_I$      NON-IDIOMATIC READING ONLY
            the rosary pass+CL
            'That rosary, give it to me'

        c.     Va estar *passant el rosari* durant tota la tarda.        #Al final, el va
            PAST be pass the rosary during whole the afternoon at the end it PAST
            perdre
            loose
            '(S)he was saying her rosary during the whole afternoon. #At the end, (s)he lost it'

        d.     #Va *passar el rosari* que li havien regalat
            PAST pass the rosary that CL have given (Non-idiomatic reading only)
            '(S)he gave away the rosary that had been given to her as a present'

        e.     L'àvia va *passar el rosari* tota la tarda. -#Quin?
            the grandmother PAST pass the rosary whole the afternoon which one
            'My grandmother said one's rosary during the whole afternoon'

**D**. Object nouns in VP idioms are weak in the sense put forward by McNally and Van Geenhoven (1998:1): a nominal α is weakly interpreted (or weak, for short) in a context C if and only if α denotes a property in C.

Being weak has at least two further consequences. The first is that object nouns take narrow scope with respect to all other operators (e.g. the negative operator, or modal and intensional predicates) that might occur within the same clause. It happens to be the case that object nouns of idiomatic constructions are always unambiguously assigned narrow scope, and in this sense they differ from indefinite NPs. Therefore, (52) cannot license a wide scope reading for the object noun. This means that the nominal is scopally inert; i.e., it cannot have wide scope relative to any operator or quantifier in whose scope the predicate occurs (Farkas and De Swart 2004:47). Accordingly, the meaning of (52b) is such that it is not the case that there is an event located during the whole afternoon, such that say-word is predicated of Joan on such an event in the afternoon.

(52)   a.     *no obrir boca*
            not say word                     'not to say a word'

        b.     En tota la tarda en Joan *no va obrir boca*
            in whole the afternoon D Joan not PAST open mouth
            'During the whole afternoon Joan did not say a word'

The second consequence is that object nouns of idiomatic constructions lack a *de re* interpretation, even if preceded by a definite determiner. Hence, in (53) it is not possible to understand the object noun as denoting some specific object and, therefore, as having a definite / presuppositional interpretation, according to which either the feature plural or the definite D (or both) contribute a presupposed discourse referent and predicate plurality of it. Rather, the definite $DP_{pl}$ is mapped into a generic inclusive reading (Laca 1990).

(53)   a.     *deixar els llibres*
             leave the books               'stop studying'
       b.     Ha *deixat els llibres* a mig curs
             has left the books in middle course
             '(S)he has stopped studying in the middle of the academic year'

The evidence presented in this section allows us to conclude that in the computational system object nouns, like i-clitics, denote abstract objects. More specifically, object nouns denote one type of abstract object –a property-type expression.

The contrast between object nouns on the one hand (no matter which their lexical status is, as a mass or as a count noun) and i-clitics on the other is that in idiomatic constructions the first can only denote properties, whereas the second, can denote various types of abstract objects. Therefore, the commonality of object nouns and i-clitics is that they instantiate two mechanisms that refer to abstract objects.

**7.** CONCLUSIONS. We are now in a position to provide an answer to the set of questions posed in Section 1, i.e. (i) what are the syntactic and semantic properties of i-clitics in idiomatic constructions? (ii) what is the contribution of each i-clitic to the meaning of an idiomatic expression? (iii) why, among the total number of clitics available in Romance syntax, is a specific limited subset able to occur in (1) and able to semantically denote one type or another of anaphoric relationship to an abstract or an indeterminate object?, and (iv) what are the semantic similarities between i-clitics and object nouns in idiomatic constructions?

The most relevant syntactic and semantic properties characterizing i-clitics in idiomatic constructions have been described in Sections 3 and 4. These properties have been related to one another on the basis of an analysis of i-clitics in idiomatic constructions that is in consonance with an analysis of noun incorporation and person-number-case morphology in polysynthetic languages (Jelinek 1984, Baker 1996, Hale 2003, Mithun 1984). Accordingly, it has been assumed that i-clitics are arguments, whose binding-theoretic status is that of free variables incorporated into the verb stem in the computational system, after unification of selecting features or modification of the event predicate.

Clitic pronouns are formally defined with case and φ-features. *L*-clitics are defined for case, person, gender and number features, whereas the neuter, genitive and locative clitics are marked exclusively for case. I-clitics share identical formal features with non-inherent clitics, but i-clitics are not coreferential with any discourse referent. Therefore, only [- I, - II] (i.e. third person) inanimate non-specific non-referential clitics, as well as neuter accusative, genitive and locative clitics, can occur in the schema in (1) whereas first and second person weak pronouns, as well as referential third person masculine accusative, dative clitics, and reflexives belong to the class of thematic clitic pronominals.

From a semantic perspective, I have shown that idiomatic constructions with i-clitics are grammatical tools designed to encode λ-abstraction over discourse accessible information, an explicit syntactic topic, or a silent discourse topic. V + Cl constructions encode functional abstraction over argument positions of the verbal predicate. They encode λ-abstracted variables which must combine with a (hidden) topic since they allow covert antecedents. When the topic argument of λ is not a referential expression but an abstract or indeterminate object, then the coindexed variable also denotes an abstract semantic object −either a proposition, a property, a generic situation, an indeterminate object, or an abstract spatio-temporal location.

Finally, the semantic similarities between i-clitics and object nominals of idiomatic constructions may be found in the fact that they all denote some kind or other of abstract object that combines with the verb by incorporation, conceived either as unification of selecting features or as modification of the event. I-clitics in V + Cl constructions, as well as bare nouns in V + (D) + N constructions, are incorporated into the verb at an abstract level of meaning representation. The output of this semantic incorporation of the clitic / nominal is a complex predicate which never allows a *de re* interpretation for the clitic / nominal; never allows a wide scope interpretation for the clitic / nominal; and never permits discourse transparency of the incorporated clitic / nominal.

## References
### Dictionaries

Enciclopèdia Catalana (1982). *Diccionari de la llengua catalana*. Barcelona. (EC).

Espinal, M.Teresa (2004a). *Diccionari de sinònims de frases fetes*. Barcelona / València: Universitat Autònoma de Barcelona. Servei de Publicacions / Publicacions de la Universitat de València / Publicacions de l'Abadia de Montserrat.

Institut d'Estudis Catalans (1995). *Diccionari de la llengua catalana*. Barcelona: Ediciona 3 i 4, Edicions 62, Editorial Moll, Enciclopèdia Catalana, Publicacions de l'Abadia de Montserrat. (IEC).

Raspall, Joana; and Martí, Joan (1984, 1996). *Diccionari de locucions i de frases fetes*. Barcelona: Edicions 62. (R-M).

Rey, Alain; and Chantreau, Sophie (1979). *Dictionnaire des expressions et locutions figurées*. Paris: Le Robert.

Turrini, Giovanna; Alberti, Claudia; Santullo, Maria Luisa; and Zanchi, Giampiero (1995). *Dizionario dei modi di dire con esempi d'autore*. Bologna: Zanichelli editore.

### Other references

Abeillé, Anne; Godard, Danièle; and Sag, Ivan A. (1998). Two kinds of composition in French complex predicates. *Syntax and Semantics* 30, 1-41.

Asher, Nicolas (1993). *Reference to abstract objects in discourse*. Dordrecht: Kluwer.

Baker, Mark (1988). *Incorporation: a theory of grammatical function changing*. Chicago: University of Chicago Press.

Baker, Mark (1996). *The polysynthesis parameter*. New York & Oxford: Oxford University Press.

Baker, Mark; Aranovich, Roberto; and Golluscio, Lucía A. (2004). Two types of syntactic noun incorporation: noun incorporation in Mapudungun and its typological implications. *Language* 81, 138-176.

Bally, Charles (1951). *Traité de stylistique française*. Vol. II. Paris: Klincksieck.

Beyssade, Claire; and Dobrovie-Sorin, Carmen (2005). A Syntax-based analysis of predication. *SALT 15*, 44-61. Ithaca, New York: CLC Publications.

Bibis, Nick (2002). *The syntax of clitics in idiomatic and other fixed expressions*. PhD thesis, University of Toronto.

Bibis, Nick; and Roberge, Yves (2004). Marginal clitics. *Lingua* 114, 1015-1034.

Bonet, Eulàlia (1995). Feature structure of Romance clitics. *Natural Language and Linguistic Theory* 13(4), 607-647.

Borer, Hagit (1984). *Parametric syntax*. Dordrecht: Foris.

Bosque Ignacio; and Moreno, Juan Carlos (1990). Las construcciones con *lo* y la denotación del neutro. *Lingüística* 2, 5-50.

Bosque, Ignacio; and Masullo, Pascual (1998). On verbal quantification in Spanish. In *Proceedings of the III symposium on the syntax of Central Romance languages,* Olga Fullana and Francesc Roca (eds.), 9-63. Girona: Universitat de Girona.

Burzio, Luiggi (1986). *Italian syntax. A government and binding approach*. Dordrecht: Foris.

Cardinaletti, Anna (1999). Pronouns in Germanic and Romance languages: an overview. In *Clitics in the languages of Europe*, Henk van Riemsdijk (ed.), 33-82. Berlin: Mouton de Gruyter.

Cardinaletti, Anna; and Starke, Michael (1999). The typology of structural deficiency: on the three grammatical classes. In *Clitics in the languages of Europe*, Henk van Riemsdijk (ed.), 145-233. Berlin: Mouton de Gruyter.

Casares, Julio (1950). *Introducción a la lexicografía moderna*. Madrid: Revista de Filología Española.

Chierchia, Gennaro. (1998). Reference to kinds across languages. *Natural Language Semantics* 6(4), 339-405.

Chomsky, Noam (1980). *Rules and representations*. New York: Columbia University Press.

Chomsky, Noam (1981). *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, Noam (1995). *The minimalist program*. Cambridge, Mass.: The MIT Press.

Chung, Sandra; and Ladusaw William A. (2004). *Restriction and saturation*. Cambridge, Mass., The MIT Press.

Cyrino, Sonia M.; Duarte, M.Eugenia; and Kato, Mary (2000). Visible subjects and invisible clitics in Brazilian Portuguese. In *Brazilian Portuguese and the null subject parameter*, Mary Aizawa Kato, and Esmeralda Vailati Negrao (eds.), 55-73. Madrid & Frankfurt am Main: Iberoamericana & Vervuert Verlag.

Dayal, Veneeta (2003). A semantics for pseudo incorporation. Ms., Rutgers University.

Déchaine, Rose-Marie; and Wiltschko, Martina (2002). Decomposing pronouns. *Linguistic Inquiry* 33(3), 409-442.

Delfitto, Denis (2002). On the semantics of pronominal clitics and some of its consequences. *Catalan Journal of Linguistics* 1, 41-69.

Delfitto, Denis (2004). Reference to facts and the syntax / semantics interface. *Lingue e Linguaggio* 2, 153-170.

Déprez, Viviane (2003). Moving phrases in nominal phases. Paper presented at the *LSRL XXXIII*, Bloomington: Indiana University.

Dobrovie-Sorin, Carmen; Bleam, Tonia; and Espinal, M. Teresa (2006). Bare nouns, number, and types of incorporation. In *Non-defineteness and plurality*, Svetlana Vogeleer and Liliane Tasmowski (eds.), 51-79. Amsterdam: John Benjamins.

Ernout, Alfred (1953). *Morphologie historique du latin*. Paris: Klincksieck.

Espinal, M.Teresa (1992). Expletive negation and logical absorption. *The Linguistic Review* 9(4), 333-358.

Espinal, M.Teresa (2001). Property denoting objects in idiomatic constructions. In *Romance Languages and Linguistic Theory 1999. Selected papers from Going Romance, Leiden 9-11 December 1999*, Yves D'Hulst, Johan Rooryck & Jan Schroten (eds.), 117-141. Amsterdam & Philadelphia: John Benjamins.

Espinal, M.Teresa (2004b). Lexicalization of light verb structures and the semantics of nouns. *Catalan Journal of Linguistics* 3, 15-43.

Espinal, M. Teresa (2005). A conceptual dictionary of Catalan idioms. *International Journal of Lexicography* 18(4), 509-540.

Espinal, M. Teresa; and McNally, Louise (2006). Bare singular nominals and incorporating verbs. Paper presented at the *III NEREUS International Workshop. Definiteness, specificity and anymacy in Ibero-Romance languages*. Alcalá de Henares: Universidad de Alcalá.

Farkas, Donka; and De Swart, Henriëtte (2003). *The semantics of incorporation*. Stanford: Stanford Monographs in Linguistics.

Farkas, Donka; and De Swart, Henriëtte (2004). Incorporation, plurality, and the incorporation of plurals: a dynamic approach. *Catalan Journal of Linguistics* 3, 45-73.

Fraser, Bruce (1970). Idioms within a transformation grammar. *Foundations in Language* 6, 22-42.

Gibbs, Raymond W. (1995). Idiomaticity and human cognition. In *Idioms. Structural and psychological perspectives*, Martin Everaert, Erik-Jan van der Linden, André Schenk and Rob Schreuder (eds.), 97-116. Hilldale: Lawrence Erlbaum Associates.

Grevisse, Maurice (1986). *Le bon usage*. Paris: Duculot.

Hale, Ken (2003). On the significance of Eloise Jelinek's pronominal argument hypothesis. In *Formal approaches to function in grammar. In honor of Eloise Jelinek*, Andrew Carnie, Heidi Harley and MaryAnn Willie (eds.), 11-43. Amsterdam: John Benjamins.

Hall, Robert A. (1965). The 'neuter' in Romance: a pseudo problem. *Word* 21(3), 421-427.

Hall, Robert A. (1968). 'Neuters', mass-nouns, and the ablative in Romance. *Language* 44(3), 480-486.

Harley, Heidi; and Elizabeth Ritter (2002). Person and number in pronouns: a feature-geometric analysis. *Language* 78(3), 482-526.

Heim, Irene; and Kratzer, Angelika (1998). *Semantics in generative grammar*. Oxford: Blackwell.

Herburger, Elena (2000). *What counts. Focus and quantification*. Cambridge, Mass.: The MIT Press.

Herman, Joseph (1975). *Le latin vulgaire*. Paris: Presses Universitaires de France.

Jackendoff, Ray S. (1990). *Semantic structures*. Cambridge, Mass.: The MIT Press.

Jelinek, Eloise (1984), Empty categories, case, and configurationality. *Natural Language and Linguistic Theory* 2, 39-76.

Kallulli, Dalina (2006). Argument demotion as feature suppression. In *Demoting the agent*, Benjamin Lyngfeldt, and Torgrim Solstad (eds.), 143-166. Amsterdam: John Benjamins.

Katz, Jerry; and Postal, Paul (1963). Semantic Interpretation of Idioms and Sentences Containing Them. *Quarterly Progress Report of the MIT Research Laboratoy of Electronics* 70, 275-282.

Kayne, Richard (1975). *French syntax*. Cambridge Mass.: The MIT Press.

Kayne, Richard (1989). Null subjects and clitic climbing. In *The null subject parameter*, Osvaldo Jaeggli and Kenneth J. Safir (eds.), 239-261. Dordrecht: Kluwer.

Kayne, Richard (1991). Romance clitics, verb movement and PRO. *Linguistic Inquiry* 22, 647-686.

Kayne, Richard (2006). Expletives, datives, and the tension between morphology and syntax. Ms., New York University.

Keyser, Samuel J.; and Roeper, Thomas (1992). Re: the abstract clitic hypothesis. *Linguistic Inquiry* 23(1), 89-125.

Laca, Brenda (1990). Generic objects: some more pieces of the puzzle. *Lingua* 81, 25-46.

Marantz, Alec (1996). Cat as a phrasal idiom. Ms., MIT.

Marantz, Alec (1997). No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania Working Papers in Linguistics* 4(2), 201-225.

Mardirussian, Galust (1975). Noun-incorporation in Universal Grammar. *Chicago Linguistic Society* 11, 383-389.

Mariner, Sebastià (1975). *Estudis estructurals del català*. Barcelona: Edicions 62.

Mascaró, Joan; and Rigau, Gemma (2002). Introduction. The grammar of clitics. *Catalan Journal of Linguistics* 1: 9-15.

Masullo, Pascual (1992). *Incorporation and case theory in Spanish: A crosslinguistic perspective*. PhD thesis, University of Washington, Seattle.

Masullo, Pascual (2004). Clitics aren't climbers!. Paper presented at the *LSRL XXXIV*, University of Utah.

Mateu, Jaume (2002). *Argument structure. Relational construal at the syntax-semantics interface*. Doctoral dissertation, UAB, Bellaterra. Downloadable: http://www.tdx.cesca.es/TDX-1021103-173806/#documents

Mateu, Jaume; and Espinal, M. Teresa (2007). Argument structure and compositionality in idiomatic constructions. *The Linguistic Review* 24(1).

McIntyre, Andrew (2004). Event paths, conflation, argument structure and VP shells. *Linguistics* 42(3), 523-571.

McNally, Louise (1992). *An interpretation for the English existential construction*. PhD dissertation, University of California Santa Cruz.

McNally, Louise (1995/2004). Bare plurals in Spanish are interpreted as properties. *Catalan Journal of Linguistics* 3: 115-133.

McNally, Louise; and Van Geenhoven, Veerle (1998). Redefining the weak/strong distinction. Ms., Barcelona & Nijmegen: Universitat Pompeu Fabra & Max Planck Institut für Psycholinguistik.

Mendívil, José Luis (1999). *Las palabras disgregadas. Sintaxis de las expresiones idiomáticas y los predicados complejos*. Zaragoza: Prensas Universitarias de Zaragoza.

Merlan, Francesca (1976). Noun incorporation and discourse referente in modern Nahuatl. *International Journal of American Linguistics* 42(3), 177-191.

Mithun, Marianne (1984). The evolution of noun incorporation. *Language* 60, 847-894.

Mithun, Marianne (1986). On the nature of noun incorporation. *Language* 62, 32-37

Monteil, Pierre (1973). *Éléments de phonétique et de morphologie du latin*. Paris: Nathan.

Navarro, Ía (2005). Valores de *le* en español mexicano. Un caso de incorporación pronominal. Master thesis. Bellaterra: Universitat Autònoma de Barcelona.

Nunberg, Geoff; Sag, Ivan; and Wasow, Thomas (1994). Idioms. *Language* 70, 491-538.

O'Grady, William (1998). The syntax of idioms. *Natural Language and Linguistic Theory* 16, 279-312.

Ojeda, Almerindo (1984). A note on the Spanish neuter. *Linguistic Inquiry* 15, 171-173.

Parsons, Terence (1990). *Events in the semantics of English: a study of subatomic semantics*. Cambridge, Mass.: The MIT Press.

Partee, Barbara (1987). Noun Phrase interpretation and type-shifting principles. In *Studies in discourse representation theory and the theory of generalized quantifiers*, Jeroen Groenendijk, Dick de Jongh, and Martin Stokhof (eds.), 115-143. Dordrecht: Foris.

Picallo, Carme (2002). Abstract agreement and clausal arguments. *Syntax* 5(2), 116-147.

Picallo, Carme (2005). Some notes on grammatical gender and *l*-pronouns. In *Specificity and the evolution/emergence of nominal determination systems in Romance. Selected papers from the II NEREUS International Workshop*. Klaus von Heusinger, Georg Kaiser, and Elisabeth Stark (eds.), 107-121. Konstanz: Universität Konstanz.

Pomino, Natascha; and Elisabeth Stark (2006). Discreteness and the case of the Spanish 'neuter' demonstratives. Paper presented at the *III NEREUS International Workshop. Definiteness, specificity and anymacy in Ibero-Romance languages*. Alcalá de Henares: Universidad de Alcalá.

Reinhart, Tanya (1981). Pragmatics and linguistics: an analysis of sentence topics. *Philosophica* 27.1.

Reinhart, Tanya (1983). *Anaphora and semantic interpretation.* London: Croom Helm.

Reinhart, Tanya (1995). Interface strategies. OTS working paper, Utrecht.

Renzi, Lorenzo (1988). *Grande grammatica italiana di consultazione*. Vol. 1. Bologna: Il Mulino.

Rice, Karen (2003). Doubling by agreement in Slave (Northern Athapaskan). In *Formal approaches to function in grammar. In honor of Eloise Jelinek*, Andrew Carnie, Heidi Harley, and MaryAnn Willie (eds.), 51-78. Amsterdam: John Benjamins.

Roberston, John S. (1980). *The structure of pronoun incorporation in he Mayan verbal complex*. PhD dissertation, Harvard University 1976. New York: Garland.

Roberts, Ian; and Roussou, Anna (1999). A formal approach to 'grammaricalization'. *Linguistics* 37(6), 1011-1041.

Rosen, Sara Thomas (1989). Two types of noun incorporation: a lexical analysis. *Language* 65, 294-317.

Russi, Cinzia (2003). On the argument structure of Italian *volerci*. Paper presented at the *LSRL XXXIII*, Bloomington: Indiana University.

Sadock, Jerrold (1980). Noun incorporation in Greenlandic: a case of syntactic word formation. *Language* 56, 300-319.

Sadock, Jerrold (1986). Some notes on noun incorporation. *Language* 62, 19-31.

Safir, Ken (1995). Abstract incorporation vs. abstract cliticization. *Chicago Linguistics Society* 31, 280-299.

Sihler, Andrew L. (1995). *New comparative grammar of Greek and Latin.* New York & Oxford: Oxford University Press.

Simatos, Isabelle (1986). *Elements pour une theorie des expressions idiomatiques*. PhD thesis, Université de Paris 7.

Simatos, Isabelle (1997). Expression idiomatique, expression figée, prédicat complexe. In *Mot et grammaires*, Bernard Fradin Jean Marie Marandin (eds.), 89-127. Paris: Didier.

Snyder, William (2001). On the nature of syntactic variation: evidence from complex predicates and complex word-formation. *Language* 77(2), 324-342.

Spencer, Andrew (1991). *Morphological theory*. Oxford: Blackwell.

Spencer, Andrew (1995). Incorporation in Chukchi. *Language* 71(3), 439-489.

Spitzer, Leo (1941). Feminización del neutro. *Revista de Filología Hispánica* 3, 339-371.

Sportiche, Dominique (1999). Pronominal clitic dependencies. In *Clitics in the languages of Europe*, Henk van Riemsdijk (ed.), 679-708. Berlin: Mouton de Gruyter.

Stowell, Timothy (1983). Subjects across categories. *The Linguistic Review* 2(3),285-312.

Szmerényi, Oswald J.L. (1996). *Introduction to Indo-European linguistics*. Oxford: Clarendon Press.

Talmy, Leonard (1985). Lexicalization patterns: semantic structure in lexical forms. In *Language Typology and syntactic description* (vol. 3): *grammatical categories and the lexicon*, Timothy Shopen (ed.), 57-149. Cambridge: Cambridge University Press.

Talmy, Leonard (2000). *Toward a cognitive semantics. Vol. II: typology and process in concept structuring*. Cambridge, Mass.: MIT Press.

Todolí, Júlia (2002). Els pronoms. In *Gramàtica del català contemporani*, Joan Solà, Maria Rosa Lloret, Joan Mascaró, and Manel Pérez Saldanya (eds.), vol. 2, chap. 6, 1337-1433. Barcelona: Empúries.

Tortora, Christina M. (1999). The postverbal subject position of Italian unaccusative verbs of inherently directed motion. In *Semantic issues in Romance syntax*, Esthela Treviño and José Lema (eds.), 283-298, Amsterdam: John Benjamins.

Van Geenhoven, Veerle (1998). *Semantic incorporation and indefinite descriptions*. Stanford: CSLI.

Vergnaud, Jean Roger; and Zubizarreta, M. Luisa 1992. The definite determiner and the inalienable constructions in French and in English. *Linguistic Inquiry* 23(4), 595-652.

**Notes**

[1] The term *lexicalization pattern* differs from that postulated by Talmy (1985, 2000). In this paper it refers to a syntactic construction which tends to be lexicalized in the sense that *lexicalization* involves the opposite of grammaticalization (Roberts and Roussou 1999). In this particular sense, Bibis and Roberge (2004) put forward the notion of *marginalization*.

[2] From now on I shall base my analysis on Catalan and Spanish data, although I shall mention data from other languages when relevant. Idioms are introduced in italics, whereas the glosses within commas correspond to their conceptual or encyclopedic meanings.

I am grateful to Adriana Pintori, Paolo Lorusso, Josep Quer, and Dimitra Lazaridou, with whom I checked the Italian and Modern Greek examples.

An idiomatic expression is a multi-word expression which satisfies at least one of the two following conditions: (i) the meaning of the lexicalized expression cannot simply be composed from the literal meaning of its components, and (ii) the morphological or syntactic form of the lexicalized expression is highly fixed. See Nunberg et al. (1994), O'Grady (1998), Espinal (2005) for criteria for idiomaticity.

[3] See Bonet (1995) for a detailed morphological description of pronominal clitics in Romance. I simplify this characterization in the glosses that follow.

[4] Notice that few clitics of the full third person clitic paradigm occur in idiomatic constructions, and that feminine pronouns are preferred over masculine forms (Casares 1950:239-241, Bibis 2002:121). In Catalan, idioms with *la* double those with *les* and triple those with neuter *ho* (Espinal 2004a).

Dutch idioms belonging to this pattern include weak pronouns, such as the neuter singular *het* 'it', the plural form *ze* 'them', and the locative *er* 'there'. Notice that *het* and *ze* are not specified for case, and *ze* is not specified for gender either; *er* "obligatorily appears to the left of prepositions as a suppletive form for pronouns with non-human reference" (Cardinaletti 1999:51).

(i)  a.  *Het*         *benauwd hebben*
        it-3rd.ntr.sg   constricted have           'be short of breath' / 'be afraid'
     b.  *Ze*          *zien vliegen*
        them-3rd.pl   see fly              'be crazy'
     c.  *Er*   *mee zitten*
        there with sit                     'worry about, bother'

I owe this observation as well as the data in (i) to Henriëtte de Swart (p.c.).

[5] Concerning French, Bally (1951:II,68) claims exactly that: "La présence des pronoms personnels et particules *le, la, y, en,* dans un groupe de mots, en l'absence de toute relation de ces pronoms avec un substantif énoncé précédemment est un indice de groupement phraséologique".

Grevisse (1986:1036) notes that "*en* et *y* ont une valeur imprécise dans un grand nombre d'expressions", and also that "dans certaines expressions, *le*, parfois *la* et *les* s'emploient sans antécédent" (p.1025).

See also note 23, below.

[6] The term *inherent clitic*, as well as its synonymous *intrinsic (idiomatic) clitic* (Abeillé et al. 1998), is used to refer to complements, sometimes argument complements of the verb, and other times adjunct complements of the verb, which can only be realized affixally. This term should be distinguished from the notion of *implicit argument*, which refers to a conceptual argument that is not expressed syntactically (Jackendoff 1990:55). Constructions involving i-clitics and constructions having implicit arguments share the property of having uninstantiated thematic arguments, but whereas in the first class the uninstantiated thematic argument is restricted by the contribution of the incorporated clitic, in the second it remains unrestricted (cf. Farkas and De Swart 2003: 67-8).

[7] A syntactic analysis of V + Cl idiomatic expressions should be distinguished from a lexicalist approach (cf. Abeillé et al. 1998), which basically conceives i-clitics as expletive, non-compositional items.

[8] The reflexive clitic, which often occurs in idiomatic expressions, belongs to the class of conceptual-thematic clitics.

[9] It is interesting to compare these alternations with cross-linguistic variations on the D selection in V / P + object idioms (I stands for Italian, C for Catalan, S for Spanish, E for English, OF for Old French, and MF for Modern French).

(i)  a.  *non chiudere occhio* ［I］/ *no tancar l'ull* ［C］
         not close (the) eye
         'not to sleep a wink'
     b.  *caer en manos de* ［S］/ *fall into the hands of* ［E］
     c.  *go away with one's hands empty* ［E］/ *anar-se'n amb les mans buides* ［C］
     d.  *promete les monts et les vaux* ［OF］/ *promettre monts et merveilles* ［MF］ (Rey – Chantreau 1979)
         'to promise the earth'

In Espinal (2001:120) these variations are conceived as suggesting that, when the bare nominal of an idiomatic construction combines with a D, there is no type-changing effect (Chierchia 1998, Partee 1987) from a property denoting type expression to a kind denoting expression, since what is expected is that the same semantic analysis is provided to these V / P + object idioms across languages and across different stages of the same language, quite independently of the presence or absence of an explicit but deficient or expletive D (Vergnaud and Zubizarreta 1992).

[10] In Mexican Spanish some occurrences of accusative and dative clitic alternations can be also attested, but whereas third person accusative feminine clitic *la* denotes an indeterminate object, or an abstract situation, third person dative clitic *le* encodes an abstract path (e.g. *pasarle* lit. pass+him-3rd.fem.dat.sg, 'go from one place to another'). See Navarro (2005) for an analysis of the productive predicate-*le* formation in this language.

[11] Within the generative literature it has been assumed that pronominal categories are endowed with case, and φ content, which provides information on grammatical person, gender and number. Person is expressed with a subset of the possible combinations of the values [±I, ±II] (that is, [+I, -II] corresponds to first person, [-I, +II] to second person, and [-I, -II] to third person). Gender is expressed as [±FEM] and number as [±PLU]. Furthermore, an implicational mapping from gender to number has been postulated in the literature: [αGEN] → [αPLU] (Picallo 2002).
In Catalan, as well as in other Romance languages, the unmarked gender is masculine, and the unmarked number is singular. Neuter *ho,* partitive *en* and adverbial *hi* are [-P, -G, -N], the distinction among these three pronominal forms being defined on the basis of structural case, and on the absence or presence of locative content.

[12] The term used by traditional grammarians is *agglutination*. Both French and Catalan show a widespread phenomenon of agglutination of the pronoun *ne* into the verb (Grevisse 1986:1039, Todolí 2002:1425).

[13] I owe these examples to Júlia Todolí (p.c.).

---

[14] Merlan (1976:178) quotes previous work by Kroeber (1909) in which it is claimed that pronominal incorporation is not the introduction into the verb of independent pronouns, but is rather an inflectional process.

[15] When a neuter pronoun substitutes a third person accusative clitic (e.g. (i a)), there is no gender and number agreement, thus showing that the gender and number features of the feminine accusative forms are what actually constrain gender and number agreement with the adjectival constituent.

(i) CATALAN
    a. *passar-ho prim*
       pass-it.ntr thin                                    'starve'
    b. *passar-la prima*
       pass-it.fem thin.fem

When a postverbal adjective does not show gender and number agreement with the feminine accusative clitic, an A to Adv conversion rule is assumed to apply. This would account for (ii a) in Mexican Spanish, which holds conceptual synonymy with the idiom in (ii b).

(ii) SPANISH
    a. *pasarla*           *bonito*
       pass-it-3$^{rd}$.fem.acc.sg nice.masc        'have a good time'
    b. *pasarla*           *bien* (= (16b))
       pass-it-3$^{rd}$.fem.acc.sg  well

I owe the example in (ii a) to Josep M. Brucart (p.c.).

[16] See Mateu (2002) for an account of the presence vs. absence of resultative constructions in Germanic vs. Romance languages, which is based on the elasticity of the meaning of the verb in the first group, and the existence of contrasting patterns and conflation processes in these two language families (see also Talmy 1985).

[17] I owe these examples to Àlex Alsina (p.c.).

[18] I wish to thank Jaume Mateu for pointing out this test to me (p.c.).

[19] This demotion of the external argument reminds us of much research in the linguistic literature, when a suppression operation has been postulated, either in the lexicon or in the syntax, in order to account for phenomena such as passivization, reflexivization, anticausatives, middles, etc. As pointed out by Kallulli (2006:14), these operations "suppress either an argument position (external or internal), a theta role in the thematic grid of the verb, or some element in the lexical-semantic structure of a predicate (depending on the theory) (Grimshaw 1990, Woolford 1993, Levin - Rappaport-Hovav 1995, Reinhart - Saloni 2004, among others)". For the notion of argument demotion, see also Jackendoff (1990).

It should be pointed out that, although the phenomenon of argument demotion that applies in idioms with i-clitics, shares with anticausatives and passives the effect of intransitivizing, an important distinction should be made. Thus, under the Principles and Parameters analysis of passives (cf. Chomsky 1981), case absorption goes hand in hand with external argument suppression; whereas in the V + Cl idiomatic pattern i-clitics keep their case features while the external argument is deagentivized, and the subject and the i-clitic do not form an argument chain.

The idiomatic schema in (1) should also be distinguished from atransitive particle constructions (McIntyre 2004) since, for example, the inherent locative *hi* can co-occur with an overt accusative DP.

(i) *dir-hi la seva* (= (22e))
    say CL.loc the own                        'intervene'

I assume that argument demotion is a conceptual effect of an operation of semantic incorporation that will be introduced in Section 5.2.

[20] Something similar has recently been claimed by Russi (2003) regarding the Italian *volere / volerci* contrast: "Grammaticalization of *ci* leads to a substantial modification of the subcategorization frame and argument structure of the derived verb *volerci: ci* incorporation changes a monotransitive *volere* into" an unaccusative verb.

[21] Mithun (1984) distinguishes four types of noun incorporation: type I is defined as a process of lexical compounding that lowers the valence of the verb by one; type II is defined as a syntactic process that advances an oblique argument into the case position vacated by the incorporated noun; type III

−characteristic of polysynthetic languages− involves the manipulation of discourse structure by a backgrounding of the incorporated noun and a qualification of the verb; and, finally, type IV is characterized by the fact that a general noun stem is incorporated to narrow the scope of the verb but, additionally, the compound stem is accompanied by a more specific external NP which identifies the argument implied by the incorporated noun.

On some occasions, the intransitive effect of $V_T$ + Cl parallels Mithun's (1984) analysis of noun incorporation, conceived as a process that derives intransitive verbs from transitive ones. On other occasions, the intransitive effect of $V_I$ + Cl may be analyzed as the specification of a cognate object. Nevertheless, some cases should be analyzed as a modification of the incorporating verb.

[22] The examples in (i) illustrate a parallel phenomenon involving the indeterminacy of accusative feminine DPs in contemporary Catalan idioms (Mariner 1975).

(i) a. *fer la seva*
       make the.fem hers                              'go one's own way'
    b. *a la llarga*
       at the.fem long.fem                            'eventually'

[23] Renzi (1988:640) claims that with Italian verbs such as *raccontarne, dirne, saperne, sentirne,* etc. the understood antecedents of the clitic are nouns such as *storie, imprese, esperienze*, and with the verb *avere*, the antecedent of the clitic *ne* is generically interpreted as denoting *situazione, fatto*, etc.

Sometimes, historical information is needed in order to recover the initial deictic antecedent of an inherent clitic pronominal form. Thus, the French idiom *les mettre* 'go out' is well-known to have derived from *mettre les bouts* (Grevisse 1986:1025) or *mettre les bouts de bois* (Rey-Chantreau 1979:118): "Les *bouts de bois* sont métaphoriquement les jambes. *Mettre* équivaut à *prendre ses jambes à son cou. Bout de bois,* dans cette locution, a pour synonymes *les baguettes, les bambous, les bois, les cannes* (tous attestés entre 1903 et 1915)".

[24] On the one hand, Bosque and Moreno (1990:16-22), insist on the claim that neuter pronouns do not denote individual entities but facts, events, qualities, or abstract and unspecified entities, and they make a distinction among Spanish individuative *lo* (when the range of the variable is a set of non-human entities, e.g., *Lo bueno de que venga* lit. it.ntr good of that comes 'what's nice about his coming'), qualitative *lo* (when the range of the variable is a set of properties, e.g. *Me impresionó lo alto del edificio* lit. me impressed it.ntr high of+the building 'I was impressed at the height of the building'), and quantitative *lo* (when the range of the variable is a set of quantities, e.g. *Juan trabaja lo necesario* lit. Juan works it.ntr necessary 'Juan works as much as it is required').

On the other hand, the meaning of Spanish neuter pronouns, which never refer to sex or animation, has been summed up as 'abstract' (Hall 1965:423), and has been equated to that of mass nouns, continuous and non-segmentable entities (Hall 1968, Ojeda 1984).

Similarly, French en can also be said to denote an abstract object, more specifically, a property and an indeterminate object, because "*en* may substitute for a lexical noun in the context of adjectival (weak) quantifiers, as well as modifying adjectives" (Déchaine and Wiltschko 2002:427), and it may also substitute for a PP introduced by the preposition *de* (Grevisse 1986).

[25] As pointed out by Hall (1965:426): "From the historical point of view, therefore, the term 'neuter' has some justification only on the morphological level and only for the non-Ibero-Romance languages; it has none at all on the syntactic or semantic level. (...). The Ibero-Romance pronouns might better be termed 'abstract', since this meaning, rather than absence of reference to sex or animation, is their prime characteristic".

See also Pomino and Stark (2006) for a reinterpretation of Harley and Ritter's (2002) feature geometry, according to which the so called 'neuter' is not a specification of the operation of individuation, but the absence of a [discrete] feature, a case of non-individuation.

[26] Spitzer (1941), after studying various historical Romance data, postulates what he calls 'the feminization of the neuter'. It is well known that the source of pronominal *l*-clitics in Romance is the Latin demonstrative pronoun series *ille* (masculine nominative singular) – *illa* (feminine nominative singular) – *illud* (neuter nominative singular) (see (i)). Diachronically, it has been suggested (Spitzer 1941:341-342) that the *–s* of *illos – illas* (masculine and feminine accusative plural forms) was transferred to *illa* (neuter accusative plural), thus forcing an identity between the feminine accusative plural and neuter accusative plural forms: *illas ossa*. In addition, Spitzer (1941:342) claims that a neuter

nominative plural form, such as *illa ossa*, was conceived as being feminine nominative singular, because of its syncretism with *illa mensa* (feminine nominative singular) (see (ii)).

(i)     LATIN

| | | |
|---|---|---|
| *ille locus* (masc. nom. sg.) | *illa mensa* (fem. nom. sg.) | *illud ossum* (neut. nom. sg.) |
| *illi locus* (masc. nom. pl.) | *illae mensa* (fem. nom. pl.) | *illa ossa* (neut. nom. pl.) |
| *illos locos* (masc. ac. pl.) | *illas mensas* (fem. ac. pl.) | *illa ossa* (neut. ac. pl.) |

(ii)    *illa ossa* (neut. ac. pl.)     →   *illas ossa*

      *illa ossa* (neut. nom. pl.    →   fem. nom. sg.)

A second idea held by comparatist and Latinist grammarians is that the *–a* ending of neuter (nominative and accusative) plurals in both Latin and Greek is a vestige of an ancient collective number (Ernout 1953, Herman 1975, Monteil 1986):

- Italian plural forms distinguish between *ossi* (the bones, conceived separately) and *ossa* (the set of bones, the skeleton), which corresponds to the collective meaning (Spitzer 1941:341).
- Spanish and Catalan feminine forms have also been associated with collective meanings (e.g. feminine *huerta / horta* 'large vegetable garden' vs. masculine *huerto / hort* 'small vegetable garden').
- Also relevant to this discusssion is the existence in languages like Roumanian of nouns that, being feminine in the singular as well as in the plural forms, do not follow the normal feminine pattern and, from a semantic perspective, have an abstract denotation: these nouns "are mass nouns, whose plurals have collective meanings or refer to different types of the objects designated" (Hall 1965:424; see also Spitzer 1941:351).
- Finally, as an anonymous reviewer points out, additional support for the idea mentioned here comes from the fact that in Ancient Greek neuter plural subjects occurred with verbs in the singular (Sihler 1995, Szemerényi 1996).

I am grateful to Pelegrí Sancho (p.c.) for relevant references on this topic.

[27] A non-referential locative pronominal has already been postulated in the linguistic literature for some linguistic constructions different from idioms. Thus, Tortora (1999) concludes that in Borgomanerese (a northern Italian spoken in Borgomanero) the locative *ghi*, which occurs with *arrive* type verbs (those entailing a GOAL, reaching a destination), does not refer to any specific location, but is the overt morphosyntactic instantiation of the GOAL entailed by the verb. Similarly, Russi (2003) concludes that in Italian the non-referential *ci* that occurs in *volerci* is the marker of an implicit resultative (rather than locative) GOAL argument of a two-argument unaccusative predicate. See also Kayne (2006) for an analysis of English *there*, Paduan *ghe*, Italian *ci* and French *y* as deictic modifiers of an abstract nominal associate.

[28] See note 21.

[29] Alternatively, it has been recently postulated that *l*-clitics are members of a Cl(assifier)P, which is to be distinguished from the DP projection. Under this view (Picallo 2005), clitics would instantiate different types of nominal classifiers.

[30] This analysis of clitics is consistent with Déprez's (2003) assumption that grammaticalization is a process that turns φ-features of a lexical item un-interpretable. It is also compatible with Baker's et al. (2004:155) analysis of *wh*-traces in Fiorentino and noun incorporation traces in Mohawk, as far as the deletion that applies to copies after movement (Chomsky 1995) does not delete the PNG features themselves, but does delete the values of those features, thus licensing –as the default case– a non-referential, non-individuated reading with an abstract interpretation.

[31] The idea that i-clitics in idiomatic expressions seem to be the arguments of the verb they are affixed to, similar to the status of clitic pronouns in polysynthetic languages was suggested to me by José Luis Mendívil (p.c.). I wish to thank him for discussion on this hypothesis.

[32] The phenomenon of noun incorporation has received considerable attention in contemporary linguistics from different perspectives: a descriptivist one (Mardirussian 1975, Merlan 1976, Roberston 1980), a typological one (Mithun 1984, 1986), a lexical one (Rosen 1989, Snyder 2001), a morphological one (Spencer 1991, 1995), a syntactic one (Sadock 1980, 1986, Baker 1988, Masullo 1992, Baker et al. 2004), and also a semantic one (Van Geenhoven 1998, Dayal 2003, Farkas and De

Swart 2003, Chung and Ladusaw 2004). These various perspectives reflect the fact that incorporation is not a uniform phenomenon across natural languages.

[33] The claim that derived predicates (after incorporation of non-referential nouns) have the tendency to become idiomatic is found scattered in the literature on noun incorporation (Mardirussian 1975:387, Mithun 1984:853).

[34] Formally, the definitions of syntactic binding and semantic binding could be stated in the following terms (Heim and Kratzer 1998:261-262):
(i)  A node α syntactically binds a node β iff α and β are coindexed, α c-commands β, α is in an A-position, and α does not c-command any other node which also is coindexed with β, c-commands β, and is in an A-position.
(ii) A node α semantically binds a node β iff the sister of α is the largest subtree in a tree γ in which β is semantically free.

[35] Regular Romance clitic constructions are conceived as hidden left-dislocation constructions, and "clitic-constructions [are said to] give rise to unsaturated λ abstracts" (Delfitto 2002:49, ex. (13b)): the clitic binding-theoretic contribution consists of a formal object encoding λx […x…], and λ abstraction must combine with a (hidden) topic which "counts as the argument of a λ abstract" (Delfitto 2002:52). This is illustrated in (i):
(i)  a.  Questo libro, l'ho letto
     b.  [λx (I have read x)] (this book)

[36] Spitzer (1941:352) holds the assumption that familiarity is what allows ellipsis of the nouns to which inherent pronominals are linked. However, as I have been able to show from (38) to (42), the study of V + Cl idiomatic expressions shows that a familiarity or aboutness criterion is neither a sufficient nor a necessary condition for topichood (Reinhart 1981).
"The intuition behind accessibility theory, then, is that the various forms of anaphoric expressions that natural language allows play a role at the interface. It is useful to have so many anaphoric forms, since they can be used to identify different antecedents. The syntax determines only when an anaphoric expression is permitted. When syntax allows more than one (which is usually the case), the choice between the different permitted forms is determined by discourse needs at the interface" (Reinhart 1995:103).

[37] See also Safir's (1995:282) concept of abstract incorporation. The idea that i-clitics with a generic interpretation are incorporated into the V at LF was put forward in Espinal (2001).
A complex predicate, in the sense used here, is not a syntactic combination of a main verb with a secondary predicate, it is the output of an operation of predicate formation by means of which the non-referential interpretation of the Cl forms a semantic compound with the predicate denoted by the verb.

[38] The main problem of this formulation, which follows closely Dayal's (2003) analysis of noun incorporation in Hindi, has to do with the lack of specification concerning the exact relationship held between the verbal predicate and the metavariable. Some alternative formulations, following respectively Dobrovie-Sorin et al. (2006) and Espinal and McNally (2006), are given in (i) and (ii).
(i)  $\lambda\mu \, \lambda x \, \lambda e \, [ \, V \, (e) \wedge \theta_1 \, (e) = x \wedge \theta_2 \, (e) \text{ has } \mu \, ]$
(ii) $\lambda x \, \lambda e \, [ \, V \, (e) \wedge \theta_1 \, (e) = x \wedge \mu \, ((\theta_2 \, (e))) \, ]$
The formula in (i) also expresses relations between individuals x and metavariables μ that stand for abstract semantic objects. What differs from Dayal's rule is that the abstract object denoted by the i-clitic is conceived as an intensional property of an object of the verb. In contrast, the formula in (ii) assumes that there is some role function $\theta_2 \, (\neq \theta_1)$ such that $\theta_2 \, (e)$ is defined and that μ modifies this role function.

[39] The following arguments were introduced in Espinal (2001, 2004b). The proposal that property-denoting arguments are semantically incorporated has been defended in Van Geenhoven (1998) for West Greenlandic, in Dayal (2003) for Hindi, in Farkas and de Swart (2003, 2004) for Hungarian, and also in Espinal (2001) for Catalan idiomatic constructions.
See McNally (1992, 1995/2004) for the idea that the argument of the existential predicate denotes a property, and Beyssade and Dobrovie-Sorin (2005), Dobrovie-Sorin et al. (2006), and Espinal and

McNally (2006) for the idea that bare singulars in argument position denote predicate modifiers of type <e,t>.

[40] See Nunberg et al. (1994:501-503) for a discussion on the existence of coreference relations between pronouns and parts of idiomatic expressions, and for the claim that some idiom chunks are possible antecedents for pronouns. These authors set up a distinction between objects of idiomatic combinations, which can be the antecedents of object pronouns in a Romance language like Italian (e.g. *fare giustizia* 'do justice'), and objects of idiomatic phrases, which cannot be the antecedents of such pronouns, nor can they be the antecedents of resumptive pronouns in left-dislocation (e.g. *mangiare la foglia* 'be caught on to the deception').