Differential object marking and topicality
The case of Balearic Catalan*

Victoria Escandell-Vidal
Departamento de Lengua Española y Lingüística General, UNED

The aim of this paper is to examine Differential Object Marking (DOM) in Balearic Catalan. While definiteness and animacy can explain the distribution of DOM in other varieties of Catalan, in Balearic, the split between marked and non-marked objects is not dependent on inherent or referential properties of the object noun phrases, but determined by topicality. A preposition is consistently used to mark a subset of topical objects, namely those occurring in clitic left- and right-dislocation structures, which correspond to two kinds of hearer-known topics: shifting topics and continuing topics. The preposition does not occur, however, with hanging topics, which introduce discourse-new topical entities. In this way, a correlation can be found between formal properties and well-motivated discourse functions that explains the distribution of DOM in Balearic. Similar patterns can be found in other Romance varieties as well, thus suggesting that topicality is relevant to account for both intra- and interlinguistic variation in DOM.

1. Introduction

Many languages show an internal variation in the way they case-mark direct objects depending on their semantic and pragmatic properties; as a consequence, the set of direct objects is ideally split into two complementary subsets. This phenomenon is known as Differential Object Marking (hereinafter, DOM), following Bossong (1983–1984, 1989, 1991, 1998), among others. The parameters that have been invoked in the literature to account for the distribution of DOM include intrinsic features of the object (animacy), discourse-related properties (referentiality, definiteness, specificity, and topicality), properties of the whole predicate (degrees of transitivity and affectedness) or a combination of more than one of these. It is usual in the literature to assume that the functional motivation for DOM is the need to distinguish objects from subjects: the more subject-like an object is, the greater the likelihood of its being overtly marked. However, this is still an open and much debated issue.
In Catalan, marked objects are introduced by the preposition *a* (roughly, ‘to’, the same preposition used for datives) and DOM is very restricted. As Aissen (2003:451) points out,

the split is between personal pronouns (of all three persons) and all other nominals. (…) Other definite objects, including proper names and definite human common noun phrases, are not preceded by *a* (whether preposed or not).

In the dialect spoken in the Balearic Islands (Majorca, Minorca, Ibiza, and Formentera), on the contrary, DOM occurs in a wider set of environments. This fact has been noticed by some grammarians (Moll 1975, Rosselló 2002) but, as far as I know, it has not been taken into account in the descriptive and typological characterisations of DOM, nor has it received a principled explanation.

The aim of this paper is to analyse the distribution of DOM in Balearic and explore its implications for a general theory of DOM. In Section 2, I will offer a brief review of the literature on DOM and the major dimensions governing the distribution of marked and unmarked objects across languages. Section 3 provides a quick survey of the data of DOM in Catalan as found in the literature. The data of Balearic Catalan, most of them taken from oral sources, will be presented in Section 4. Section 5 is devoted to exploring the extension of DOM in Balearic and the conditions governing it: the role of topicality and the behaviour of different kinds of topics with respect to DOM will be examined in detail in this section. Section 6 will deal with the connections between topicality and DOM in other Romance languages. The main findings and consequences will be gathered in the last section.

2. Differential Object Marking

DOM is a phenomenon attested in a wide variety of typologically unrelated languages all over the world. Languages with DOM mark direct objects in (usually two) different ways depending on certain intrinsic or discourse-related properties of the object. Though the phenomenon itself can be seen as basically the same across languages, the ways in which it is manifested and the particular conditions governing the split between marked and unmarked objects may vary considerably from language to language.

The marking of a particular subset of objects can be done by various means. Finnish, for instance, assigns different morphological cases: some objects are marked as accusative while others receive the genitive or partitive. Spanish and Romanian resort to an adpositional strategy by using a preposition for the marked case: the preposition can be either the same one that is used for other cases —
Spanish uses the dative preposition *a* (‘to’) — or it can be a special form — as in Romanian *pe* (Farkas & von Heusinger 2003; von Heusinger & Onea 2008). DOM can also be manifested as verbal agreement with the object, as in Bantu languages (Morimoto 2002). Finally, the ability of an object to undergo movement to a different position represents a further case of the split, as exemplified in Icelandic object shift (Ritter & Rosen 2001).

As for the conditions that determine the split, there has been much debate in the literature. In the original approach to DOM in Bosson (1983–1984), the author considers the relevant properties to fall into two classes: ‘inherent features which are independent of the context and which correspond roughly to what is generally called animacy’ and ‘referential features which vary as a function of the syntagmatic and pragmatic environment.’ (Bosson 1983–1984: 8). This sort of bi-dimensional system based on inherence and reference is at the basis of the two following scales:

- **Inherential dimension:** [deictic] > [proper] > [human] > [animate] > [discrete]
- **Referential dimension:** [definite] > [identifiable] > [indefinite] > [non-referential]

Aissen (2003) takes a similar position, which offers an approach to DOM couched in terms of Optimality Theory, considering animacy and definiteness as the basic parameters. Her treatment resorts to the notion of Harmonic Alignment as an explanatory tool to account for possible (and impossible) DOM systems across languages. She concentrates on DOM as manifested through morphological case and the use of adpositions. Some languages are sensitive to definiteness only, though they differ with respect to the cut-off point: in Pitjantjatjara, a Pama-Nyungan language of Australia, only personal pronouns and proper names are marked, whereas Hebrew marks all definites. Animacy seems to be a more elusive category, since it is difficult to find languages in which all and only human-referring objects are case-marked: one such language could be Sinhalese, but unfortunately the marking is apparently optional. The other languages in which DOM is determined by animacy have their cut-off point either before the limit of the human category (as Yiddish does, with DOM restricted to some humans) or beyond it (as Ritharngu, also a Pama-Nyungan language, with DOM extending to higher level animals, such as kangaroos and dogs, but not fish or raccoons). In some other languages, DOM is determined by both parameters. Romanian, Spanish, and Persian represent three cases in point, in which DOM is governed by both definiteness and animacy. In Aissen’s approach, only these two dimensions are taken into account; she acknowledges the role of other parameters, such as person, topicality — particularly when case marking is optional — and telicity, but does not take them into account in her analysis.
While maintaining his basic assumptions, in his 1991 paper, Bossong presents a slightly different view on the semantic factors that govern DOM. He identifies not two, but three basic dimensions, which he calls *inherence*, *reference*, and *constituence*. The (new) domain of constituence has to do with the relative dependence/independence of the object NP with respect to the verbal predicate: the more autonomous an object is, the greater the chance of its being marked. Constituence seems to be a rather complex category, since it can be understood in different ways: as dependent or independent existence, as the distinction between affected and effected objects, or as a verb and object being part of the same pragmatic constituent (usually, the rheme). In this way, further pragmatic considerations come into play in determining the split in the domain of objects.

Other researchers have found new factors that can be relevant to determining the split in DOM languages. Ritter & Rosen (2001) argue for the countability of nominals and events as the triggering feature that determines the split in the class of objects. Definiteness/specificity represents countability in the nominal domain: Hebrew and Turkish use a special particle or suffix to introduce definite direct objects; Bantu languages also show object agreement only when the object is specific, definite or animate; and an object shift only occurs in Icelandic when the direct object is definite or specific. There are languages where both the properties of the nominal and those of the event play a role in determining the way in which objects are marked. In Finnish, only “bounded” objects receive the accusative case. The same constraints both on the nominal and on the event hold for Palauan objects to show agreement on the verb and for object shift in Mandarin. These languages give evidence for the existence of systematic correlations between the countability of nominals and the telicity of events, on the one hand, and the way in which objects are marked, on the other.

In a similar vein, Naess (2004) presents a proposal based on the notions of transitivity and on the degree of the affectedness of objects. She begins by arguing that, contrary to Aissen’s claims, prototypical objects are highly individuated. The object is the more affected argument and affectedness is correlated with a high degree of individuation. This correlation is manifested by the fact that, from a cross-linguistic perspective, objects with a low degree of individuation tend to surface in incorporation constructions (which are not to be considered to be transitive). On the assumption that a definite object is more affected than an indefinite one, Naess suggests considering a high degree of affectedness as the relevant factor in DOM systems. In Finnish, for example, highly affected objects take the accusative case, whereas partly affected ones take the partitive.

More recently, Dalrymple and Nikolaeva (2007) have analysed the role of information structure as crucially determining various instances of DOM. They argue that there are many languages in which the split of objects is not governed
by the semantic features of the object or the event, but by the topical status of the object. According to their proposal, the underlying motivation for DOM is not to distinguish subject from object, as claimed in other approaches, but to emphasise the functional similarity ‘between subjects and topical objects, both of which tend to be grammatically marked, in contrast to nontopical objects.’ (Dalrymple & Nikolaeva 2007: 1) DOM can thus be seen as a form of secondary agreement. As they show, this is the case in Chatino, a Zapotecan language, in which a preposition occurs with topical objects only. Similarly, in Ostyak, an Ugric language spoken in western Siberia, topical objects trigger verb agreement. This does not mean that intrinsic or referential features play no role at all: for example, the occurrence of the particle râ in Persian is determined by both information structure and semantic considerations. In fact, other properties such as definiteness, specificity, and animacy can be ultimately related to the notion of topic as well, since prototypical topics are usually definite, specific, and animate.

The previous review of the literature has shown that the dimensions governing the split in the overt marking of direct objects can vary considerably from language to language, though some universal tendencies seem to emerge. Hawkins (2004) has suggested that universal implications in language structure should be derived from scales of processing difficulty (see also Haspelmath 2005, 2008). Current research is oriented to further refining these distinctions and to determining the relative weight of all these dimensions (de Swart 2007; de Swart & de Hoop 2007; de Hoop & Malchukov 2007; Malchukov 2008).

3. DOM in Catalan

According to Milá i Fontanals (1861:462 and ff), Catalan can be divided into two major dialect blocks: Eastern and Western Catalan. The Eastern block includes Central Catalan (the main dialect in terms of the number of speakers, roughly corresponding to the provinces of Barcelona, Tarragona, and Girona), Balearic (the dialect spoken on the islands of Majorca, Minorca, Ibiza, and Formentera), Northern Catalan (spoken in Roussillon, an area corresponding to the French department of Pyrénées-Orientales), and Alguerese (spoken in Alghero, a city on the Italian island of Sardinia). The Western dialect block includes Valencian (spoken in some areas of the provinces of Castelló, València, and Alacante), North-Western Catalan (spoken in the province of Lleida and other limiting areas) and a transitional set of varieties between them (See also Veny 1978; Hualde 1992: xxiii and ff; Wheeler et al 1999: xviii; Montoya Abat 2002: 8 and ff.).

There are currently two standard forms of the language: one in Catalonia, the Balearic Islands, Roussillon, and Alghero, which I will refer to as Standard
Catalan, and a slightly different one in the Community of València (Valencian). Standard Catalan is the result of a normalisation process that took place at the turn of the 20th century and it is based mostly on Central Catalan, with some features from other dialects. This standard is recommended in written texts, higher registers, formal situations, the administration, and the mass media.

The received view on DOM in Catalan is based on Standard Catalan (cf. Comrie 1979, Bossong 1983–1984, Vallduvi 2002, Bel 2002, Aissen 2003). Catalan grammarians point out that the personal direct object never carries the preposition (Moll 1952: 217; Badia 1980, 1994: 207; Bel 2002): ‘In Catalan the direct object, which is almost always represented by a nominal category, does not carry any identifying mark such as a preposition or agreement.’ (Bel 2002: 1097; my translation, V E-V) This is shown in examples (1)–(5):

(1) Standard Catalan (Comrie 1979: 15)
No havien vist l'alcalde.
not have.pst.3pl seen the.mayor
‘They had not seen the mayor.’

(2) Standard Catalan (Bel 2002: 1097)
En Martí renya el seu germà petit.
the Martin squabble.prs.3sg the his brother little
‘Martin scolds his younger brother.’

(3) Standard Catalan (Conversation 19/01/2002, from Mayol 2002: 76)
I després la Roseta ha pujat a veure aquella dona, and afterwards the Roseta have.prs.3sg gone-up to see that woman, la Mariona.
the Mariona
‘And afterwards Roseta went up to see that woman, Mariona.’

(4) Standard Catalan (Vallduvi 1992: 90)
… donc el Joan el veiem ben poc.
… since the John him.obj see.pst.1pl very little
‘… since John, we saw very little of.’

(5) Standard Catalan (Vallduvi 2002: 1224)
El Pere l’estima, la Maria.
the Peter her_love.prs.3sg, the Mary
‘Peter does love Mary.’

Only pronouns have to be introduced by the preposition a (‘to’), as shown in (6).
More specifically, the contexts in which the occurrence of the preposition is compulsory are the following (Badia 1980, 1994; Moll 1952; Bel 2002: 1098):

– When the object is a strong (i.e. independent, non-clitic) personal pronoun (regardless of whether the pronoun is doubled with a clitic or not):

(7) Standard Catalan

a. *Jo t’ajudo tu i tu m’ajudaràs *(a) mi.
   I you-help (to) you and you me-help-fut.1sg *(to) me
   ‘I’ll help you and you’ll help me.’

b. Només va invitarem la Maria i *(a) ell.
   only have.pst.3sg invite the Mary and *(to) he
   ‘S/he only invited Mary and him.’

– When the object pronoun appears in a reciprocal construction:

(8) Standard Catalan

Ens miràvem l’un *(a) l’altre.
   us look.pst.1pl the-one *(to) the-other
   ‘We looked at each other.’

In addition, the preposition introducing the direct object is not obligatory, but is nevertheless accepted (and even largely preferred) in the following cases:

– When the object is a universal quantifier, such as tothom (‘everyone’) and tots (‘all’):

(9) Standard Catalan

a. Hi he salut *tothom.
   there have.prs.1sg greeted (to) everyone
   ‘I greeted everyone.’

b. Els ha aprovat *tots.
   them have.pst.3sg passed (to) all
   ‘He passed all of them.’

– When the object is the relative pronoun el qual (‘whom, which’), in which case, the preposition merges with the pronoun: a el qual > al qual:
The previous examples thus show that Standard Catalan seems to be sensitive to definiteness in the sense of Aissen. Thus, Standard Catalan follows the pattern of other languages in which the split between marked and non-marked objects is governed by the referential properties of the object. If these languages are ordered along a scale in which definiteness, accessibility, and referentiality are combined, the resulting distribution can be represented as in Table 1:

Table 1. Languages sensitive to definiteness (in the sense of Aissen)

<table>
<thead>
<tr>
<th>Pronoun</th>
<th>P.Name</th>
<th>Def.NP</th>
<th>Indef.NP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Catalan</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pitjantjatjara</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hebrew</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Turkish</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Written Japanese</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The distribution of DOM in Catalan, however, seems to be far more complex than this characterisation might suggest. To begin with, some additional syntactic environments favour the occurrence of the preposition. For instance, with coordinated direct objects, one of which requires the preposition, the use of the preposition usually spreads to all other constituents as a result of syntactic parallelism:

(11) Standard Catalan (Solà 1990: § 9.2)

\[
\text{L’advocat va fer riure a tothom menys al jutge}
\]

‘The lawyer made everyone laugh but the judge.’

The use of the preposition a before a definite animate object ‘is tolerated when its absence might suggest that the noun phrase involved was a subject’ (Wheeler et al. 1999:243):

(12) Catalan (Wheeler et al. 1999:243)

a. \[
\text{T’estima com a la seva mare}
\]

‘She loves you like (she loves) her mother’
The preposition should also be accepted, according to Solà (1994: cap. 9), as a means of introducing dislocated human objects, as in (13):

(13) Standard Catalan (Solà, 1994: cap. 9)

a. A Núria, no crec que la pugues convencer.
   to Nuria, not think.prs.1sg that her can.pst.2sg persuade
   ‘Nuria, I don't think you can persuade her’

b. Als funcionaris no els satisfà la proposta.
   to_the civil_servants not them satisfy.prs.3sg the proposal
   ‘Civil servants are not satisfied by the proposal.’

The construction illustrated in (13) involves clitic dislocation: a constituent is detached to a peripheral position and must be resumed by a co-referential clitic in the core clause.²

On the other hand, in spoken or colloquial registers, the preposition a occurs in an even wider set of contexts, as Aissen herself recognises (Cf. Aissen 2003: 451). In fact, Hualde (1992: 86–87; 237–238) describes Catalan as a language with DOM extending to all human definites: ‘For most speakers, in spoken language all human direct objects are marked by the preposition a (or [әnә] in some areas)’:

(14) Spoken Catalan (Hualde 1992: 86–87 and 241)

a. Veuré a la Maria / veuré [әnә] la Maria.³
   see.fut.1sg to the Mary
   ‘I will see Mary.’

b. Les monges no estimen a/ana les nenes.
   the nuns not like.prs.3pl to the girls
   ‘The nuns don’t like the girls.’

There is some variation in the spoken language with respect to non human animals, which may take the preposition:

(15) Spoken Catalan (Hualde 1992: 241)

El pagès estima (an) el seu cavall.
the peasant love.prs.3sg (to) the his horse
‘The peasant loves his horse.’

Indefinites never take the preposition:

(16) Spoken Catalan (Hualde 1992: 241)

Vaig veure molts de policies al carrer.
have.pst.1sg see many of policemen at_the street
‘I saw many policemen in the street.’
It is possible that this system roughly corresponds to Central Catalan (the dialect with the largest number of speakers, spoken in Barcelona, most of Girona, and the northern half of Tarragona). In fact, relying on Hualde's data, Naess (2004: 1188) classifies Catalan as a language with DOM depending on animacy/humanness.

The contrast between Standard and Central Catalan can be summarised as in Table 2. The dark grey areas represent the extension of obligatory DOM; the light areas show optional DOM.

<table>
<thead>
<tr>
<th>Pronoun</th>
<th>Proper name</th>
<th>Definite NP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+animate</td>
<td>−animate</td>
</tr>
<tr>
<td></td>
<td>+human</td>
<td>−human</td>
</tr>
</tbody>
</table>

Given this situation, one may wonder why there is such a difference between the spoken language and the standardised variety. As mentioned before, Standard Catalan is the result of a normalisation process. Prescriptive grammars have been very keen to warn speakers against the use of the prepositional accusative, which is considered a reprehensible transfer from Spanish:

In modern spoken Catalan the preposition a is very commonly used to introduce the personal direct object; this use is a result of Castilian influence, and is uncommon among less educated people and should be carefully avoided in literary language. (Fabra 1918: 112)

Normative grammars often reiterate that a should be avoided before the direct object, which seems to indicate that spoken registers do in fact depart from the received standard. Badia (1994: 209) points out that

...the exceptions to the general rule (‘personal direct objects are never introduced by a preposition’), whether obligatory or merely tolerated, form a true link to a progressive extension of the preposition for cases rejected by tradition and for structural reasons.

However, a quick look at earlier stages of the language shows a different picture. DOM is attested in Old Catalan in more contexts than those accepted by the current standard. According to Meier (1947; cited in Sancho Cremades 1995), the occurrence of the preposition with proper names was not unusual in the old language, from the Middle Ages to the 19th century (Moll 1952: 217; Sancho Cremades 1995: 188–190). The following examples show unequivocal cases of prepositions appearing before proper names:
(17) 13th century Catalan (Jaume I cited in Sancho Cremades 1995: 188)

Faeren rey a don Anrich.

make.pst.3pl king to Sir Henry

‘They made Sir Henry King.’

(18) 14th century Catalan (St Vincent Ferrer cited in Sancho Cremades 1995: 188)

Així se prove si ames a Jesuchrist.

in.this.way cl prove.prs.3sg whether love.prs.2sg to Jesus.Christ

‘In this way it can be proved that you love Jesus Christ’

(19) 15th century Catalan (Curial e Güelfa, cited in Sancho Cremades 1995: 190)

Ja tenia a Jacob de Cleves a terra per occiure'l.

already have.pst.3sg to Jacob of Cleves to ground to kill.him

‘He already had Jacob of Cleves on the ground to kill him.’

The same goes for 19th century writers, who sometimes used the preposition to introduce human definites4, as can be gathered from Moll’s (1952: 217) considerations:

The writers of the Renaixença period [roughly speaking, the 19th century], prior to the careful grammatical restoration, were not very scrupulous about the deletion of the preposition: Estimen a la mare [They love to the mother] (Verdaguer), No hi han trobat al vicari [They didn’t find to the vicar] (Vilanova).

The diachronic data suggest that DOM is not necessarily a transfer from Spanish and that the normative ban against it has given rise to a system in which ‘the formerly far more widespread use of the preposition a with the direct object has been reduced to the absolute minimum (the free pronoun)’. (Bossong 1983–1984: 11; cf. also Bossong 1998: 225).5 If this view is correct, what the standardisation did was to promote an artificial regression of DOM.

4. DOM in Balearic: The basic data

The dialect of the Balearic Islands is one of the most differentiated within the Catalan domain. Though each of the islands of Majorca, Minorca, Ibiza, and Forménitera has its own varieties, Balearic can be considered quite homogeneous, at least at the syntactic level.

There are a number of reasons for which studying Balearic data can be relevant for a general understanding of DOM. To begin with, the population of the islands (particularly, the small ones) has been monolingual and mostly illiterate for centuries, with very limited contact with Spanish until the 1940s. Only the upper class received basic literacy, and teaching was carried out only through Spanish.
until very recently. This provides us with a language that has evolved almost in isolation from both written Catalan and Spanish — a fact that has contributed to maintaining archaic features (Montoya Abat 2002: 11), while favouring the growth of peculiar and innovative solutions (Veny 1978).

A second reason for interest in the analysis of Balearic comes from the ‘visibility’ of the preposition used for DOM in this dialect. As mentioned before, in some other dialects, it can be difficult to determine whether the preposition a is actually present due to schwa merging (See footnote 4). Balearic, on the contrary, systematically resorts to the allomorphic variant an [әn] to avoid the merging of the two unstressed schwa sounds — very much like the homograph English indefinite article. Thus, the preposition will always be heard, either as [ә] before a consonant or stressed vowel, or as [әn] before an unstressed vowel. The fact that Balearic is more transparent in this respect than Central Catalan makes it an optimal testing ground for studying DOM through oral examples.

In this section, I will present the data on the distribution of DOM in Balearic. As in other Catalan dialects, the preposition is always needed with strong pronouns, reciprocal constructions, relative pronouns, and universal quantifiers. The following examples illustrate the basic distribution:

(20) Balearic (COD: Manacor CAR: Minorca)
A valtros no vos deim forasters, us deim
you.1sg not you.obj.1sg call.prs.1sg outsiders, you.obj.1sg call.prs.1sg catalans.
‘You, we don’t call outsiders, we call you Catalans.’

(21) Balearic (Arxiu, Sant Joan, Ibiza)
Ja no el vaig deixar més i ell tampoc a jo.
already not him.obj have.pst.1sg leave more and he neither to I
‘I had not left him ever since and neither had he.’

(22) Balearic (COD Eivissa ERM, Ibiza)
… coneix a qui sigui …
… know.prs.3sg to who be.sbjv.3sg
‘s/he knows whoever’

(23) Balearic (Arxiu, Sant Joan, Ibiza)
Es va calar vora s’al·lota i sa vella va agafar
cl have.pst.3sg put near the.girl and the old.fem have.pst.3sg take
un mantó gros i els va emmantonar a tots dos. 8
a shawl large and them have.pst.3sg shawl to all two
‘He placed himself near the girl and the old lady took a large shawl and covered the two of them with it.’
Proper names and human definites, on the contrary, consistently appear without the preposition:

(24) Balearic (Arxiu, Maó. Minorca)

(a) He anat a agafar es qui ha vist have.prs.1sg gone to catch the[one] who have.prs.3sg seen
    com matava en Fulano o en Sutano. how kill.pst.3sg the Fulano or the Sutano
    ‘I went to catch he who has seen how he killed so-and-so or what’s-his-name.’

(b) I va anar ja a amenaçar es general. and have.pst.3sg go already to menace the general
    ‘So he went to threaten the general’

(25) Balearic (Arxiu, Ferreries, Minorca)

No treien per pagar aquella, aquella gent qui havien not get.pst.3pl for pay that, that people who have.pst.3pl de menester. of need
    ‘They didn’t get enough [money] to pay the people they needed.’

In (24)a, an instance of the human definite (es qui ha vist... ‘the one who has seen...’) and two proper names (en Fulano, en Sutano, ‘so-and-so, what’s-his-name’) occur as the direct objects of agafar (‘catch’) and matava (‘killed’), both appearing without the preposition. The same goes for the objects of amenaçar (‘menace, threaten’) in (24)b, and pagar (‘pay’) in (25), which bear no special marking. Notice that the Balearic examples unequivocally indicate that no preposition is present; otherwise, the an form would have been required: an es qui ha vist..., an en Fulano, an en Sutano, an es general, an aquella gent.

The example in (26) shows the contrast between human definites (es director, ‘the manager’; l’amo, ‘the owner’), which appear without the preposition, and the relative pronoun, which, as expected, is marked:

(26) Balearic (COD Eivissa ERM, Ibiza)

haurà entrat a s’hotel perquè algú de sa família have.fut.3sg entered to the.hotel because someone of the family
    coneix es director o coneix l’amo de s’hotel, know.prs.3sg the manager or know.prs.3sg the.owner of the.hotel, o coneix a qui sigui.
    or know.prs.3sg to whom be.sbjv3.sg
    ‘He must have obtained the job at the hotel because someone in his family knows the manager or knows the owner of the hotel, or knows whoever.’
To this point, Balearic has DOM for pronouns but not for human definites and proper names, thus sticking with Standard Catalan and departing from the widespread tendency of Central Catalan to mark all human definites with the preposition *a* (Cf. (14)). As in Standard Catalan, the preposition is also used in Balearic to introduce clitic-dislocated human objects, as shown in (27) and (28):

(27) Balearic (*Arxiu*, Sant Josep, Ibiza)

An aquella al·lota no la deixaven parlar amb so que ella volia.

to that girl not her let.pst.3pl talk with the who she want.pst.3sg

‘That girl was not allowed to talk to the one she wanted to.’

(28) Balearic (*COD* Eivissa ERM, Ibiza)

...som eivissencs, són es nostros i an aquest

be.prs.1pl Ibizans, be.prs.3pl the ours and to this

l’ ajudam…

him-help.prs.1pl

‘We are Ibizans, they are one of us and we will help this [guy].’

There are, however, significant cases in which Balearic differs considerably from other dialects. To begin with, the occurrence of the preposition with dislocated human objects, which is apparently a feature common to Standard and Central Catalan, seems to be optional in Balearic, as illustrated in the following examples:

(29) Balearic (*Arxiu*, Ferreries, Minorca)

Hi havia molts de puestos

there be.pst.3sg many of places

que es missatges els havia de pagar l’amo

that the servants them have.pst.3sg of pay the.master

‘There were many places where the servants, the master had to pay them.’

(30) Balearic (*Arxiu*, Ciutadella2, Minorca)

... aquell el pagaven de banda

that[one] him pay.pst.3pl of side

‘...this guy, they would pay on the quiet’

A second difference is that in Balearic the preposition appears with right-dislocated human objects as well:

(31) Balearic (Moll 1975: §239)

No la cridis, a la senyora.

not her call.imp.2sg, to the lady

‘Don’t call the lady.’
This is a context in which no preposition shows up in other Catalan dialects (Cf. (5)). If we compare the examples in (24)–(25) of human definites without the preposition, with the ones in (27)–(28) and (31)–(32) of human definites with the preposition, the provisional generalisation seems to be the following: in Balearic there is no marking when human direct objects appear in their canonical position, but the preposition may appear when they are left-dislocated and must appear when right-dislocated. Thus, for human definites the occurrence of the preposition seems to be dependent on the position of the constituent. A detailed discussion of the exact import of such syntactic positions will be provided in Section 5.

The most striking difference between Balearic and the other Catalan dialects, nevertheless, has to do with the use of the preposition *a* to introduce dislocated non-animate definites, as shown in (33)–(34):

(33) Balearic (Moll 1975: § 239)

Colliu-les, *a les peres, que ja són madures.*

‘Pick the pears, for they are ripe.’

(34) Balearic (Rosselló 2002: 1932)

a. *Ja ho sé,* [әn] això.\(^{10}\)

already it know.prs.1sg, to this

‘This I know.’


not her have.prs.1sg, to this_one

‘This one I haven’t got.’


them have.prs.1sg found.fem, to the mistakes

‘The mistakes I have FOUND.’

These examples attest the use of the preposition to introduce definite non-human constituents and, not surprisingly, the two grammarians that register the data in (33)–(34) are of Balearic origin. This use is rejected in Standard Catalan and apparently sounds quite odd to the speakers of other Catalan dialects, judging by the following words by Solà (1994: 168–169) commenting on Moll’s example:

…at least for the majority of speakers of Catalan it not only is unacceptable, but also inconceivable […] This sort of non-animate complement introduced by *a* can indeed be found in some ancient documents (namely, in the *Gazophylacium* of
The occurrence of the prepositions with dislocated non-animates, both preposed and postposed, is nevertheless widespread in all Balearic dialects and well attested in oral corpora and Internet blogs and chats, as shown in the following examples:

(35) Balearic (Majorcan claunmallorca)

\[ ...es turisme passa ara per hores baixes, \]
\[ i an això es comerç ho pateéis \]
\[ ‘The tourism industry is going through tough times and business is suffering from it.’ \]

(36) Balearic (Majorcan)

\[ An aquesta darrera [frase] noltros la diríem així \]
\[ to this last [sentence] we it.obj say.cond.1pl like-that \]
\[ ‘This last sentence, we would say this way.’ \]

(37) Balearic (Arxiu, Eivissa 3, Ibiza)

\[ a. \ La vaig heredar jo an aquella escola. \]
\[ it have.pst.1sg inherit I to that school \]
\[ ‘So I took the reins of that school.’ \]

\[ b. \ Jo la vaig llegir an aquesta [comedia] \]
\[ I it have.pst.1sg read to this [comedy] \]
\[ ‘This comedy I’ve read!’ \]

(38) Balearic (Arxiu, Formentera)

\[ a. \ …perquè tenen com un peu que per girar es formatge \]
\[ …because have.prs.3pl like a base that to turn the cheese \]
\[ va molt bé, i ara no el volen fer an aquell \]
\[ go.prs.3sg very well, and now not it.obj want.prs.3pl make to that \]
\[ peu a ses escudilles. \]
\[ base to the bowls \]
\[ ‘…because they have a sort of base that is very useful for turning the cheese, but nowadays they don’t want to make that base for the bowls anymore.’ \]

\[ b. \ I dins què el posau, an es [formatge]…? \]
\[ and inside what it.obj put.prs.2pl, to the [cheese]…? \]
\[ ‘So where do you put the [cheese] in?’ \]
(39) Balearic (Arxiu, Maó. Minorca)
Bono, son coses que açò també vindrà en s’altra part, well, be.prs.3pl things that this also come.fut.3sg in the other part, de quan em vaig decidir a fer-ho an açò
of when cl.1sg.obj have.pst.1sg decide to do-it to this
‘Well, these are things that will come out in the other part, about when I decided to do this.’

(40) Balearic (Majorcan mado-lucia)
Encara sigui una espècia autòctona la faria fugir even be.sbjv.3sg a species indigenous it.obj made.cond.1sg go-away jo a n’aquesta!
I to this!
‘Even if it is an indigenous species, I would make it disappear!’

Again, the preposition seems to be optional with left-dislocated definites. Compare (35)–(36) with (41):

(41) Balearic (Arxiu, Eivissa 3. Ibiza)
a. Altres que ni siquiera s’han representat, ‘Una de gatera’,
others that not even cl.have.prs.3pl presented, que aquella l’havien de fer a Sant Josep…
that that it.have.pst.3pl of do at Saint Joseph
‘Other [plays] have never been presented, “Una de gatera”, they had to put on stage at Saint Joseph (town)…’

b. Diuen que llavor es sous no arribaven,
say.prs.3pl that then the money not arrive.pst.3pl, això no ho ha de dir sa gent
this not it have.prs.3sg of say the people
‘They say that money didn’t then get [to its destination], this, people shouldn’t say it.’

Thus, the examples in (29)–(41) show a fully regular pattern in Balearic: definite dislocated constituents are marked with the preposition, whether they are human (Cf. (27)–(32)) or not (Cf. (33)–(40)); the preposition is optional with left-dislocated expressions and obligatory with right-dislocated ones. What is remarkable here is that the preposition would be absolutely rejected if the direct objects occurred in their canonical positions, which means that the split between marked and unmarked objects has to do with the position in the sentence, not with definiteness or animacy.

Since Balearic can mark all dislocated definites, the next step is to check whether dislocated indefinites are also case-marked. I have found no examples of dislocated indefinites in the corpora, so I asked my informants on a couple of sentences originally from Vallduvi (2002: 1255–1256). All my informants accept the
Differential object marking and topicality. The case of Balearic structures without the preposition (Cf. (42)). There seems to be strong disagreement among the speakers, however, with respect to the versions introduced by a in (43): about half of my informants tend to reject them, whereas the other half are perfectly happy to produce them.

(42) Catelan Balearic
   a. *Un gelat, *me'l prendria amb molt de gust.
      an ice.cream, me-it take.cond.1sg with much of pleasure
      ‘An ice cream, I would have with great pleasure.’
   b. Me'l prendria amb molt de gust, un gelat.
      me-it take.cond.1sg with much of pleasure, an ice.cream
      ‘An ice cream, I would have with great pleasure.’
   c. Vaig caçar dos conills. Un conill el vaig matar
      have.pst.1sg bag two rabbits one rabbit it.obj have.pst.1sg kill
      a prop de la llodriguera i l’altre...
      near of the warren and the other…
      ‘I bagged two rabbits. I killed one [of them] near the warren and the other…’

(43) Catelan Balearic
   a. %A un gelat, me'l prendria amb molt de gust.
      to an ice-cream, me_it take.cond.1sg with much of pleasure
      ‘An ice-cream, I would have with great pleasure.’
   b. %Me'l prendria amb molt de gust, a un gelat.
      me-it take.cond.1sg with much of pleasure, to an ice-cream
   c. %Vaig caçar dos conills. A un conill el vaig matar
      have.pst.1sg bag two rabbits. To one rabbit it.obj have.pst.1sg kill
      a prop de la llodriguera i a s’altre...
      near of the warren and to the other…

As for clauses, they can also be dislocated, but they are never marked with the preposition. The examples in (44), though originally from Central Catalan, are common to all dialects:

(44) Catalan (Buenafuente 2001:96 and 203, from Mayol 2002:30 and 70)
   a. Ja ho sabem, (*a) que ets a la platja.
      already it.obj know.prs.1pl, (*to) that be.prs.2sg at the beach
      ‘We already KNOW that you are at the beach!’
   b. Doncs que ho diguin a la botiga, (*a) que va
      so that it say.sbjv.3pl at the shop, (*to) that go.prs.3sg
      amb piles…!
      with batteries
      ‘At the shop they should have said that it worked with batteries!’

All rights reserved
The main facts presented in this section with respect to DOM in Balearic can be summarised as follows: Strong pronouns are always marked; pronouns, proper names and definite NPs are obligatorily marked when they are right-dislocated; the marking is very common for left-detached definites, though it seems to be optional; for a number of speakers, indefinites can also be marked when they are dislocated, preferably if animate; and dislocated clauses are never marked.

In Table 3, a provisional descriptive outline of the distribution of DOM in Balearic is provided. The dark shaded areas represent obligatory prepositional marking; the light shaded areas indicate optional DOM:

<table>
<thead>
<tr>
<th>OBJECTS</th>
<th>DEFINITENESS</th>
<th>Pronouns</th>
<th>Proper names</th>
<th>Definites</th>
<th>Indefinites</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-detached</td>
<td>Right detached NPs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Detached</td>
<td>Left detached NPs</td>
<td>Clauses</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If compared to the extension of DOM in Standard and Central Catalan, which can be accounted for on the basis of two dimensions, namely, definiteness and animacy (cf. Table 2), the conditions on the split of objects in Balearic seem to be far more complex and involve other features related to the position of the objects in the clause. In the next section, an analysis will provide a more principled account for the conditions that govern DOM in Balearic.

5. An account of Balearic DOM

The data presented in the previous section call for an analysis in theoretically well-motivated terms in order to understand what the dimensions that determine DOM in Balearic are and how they relate to each other. In this section I will argue that the main factor governing the split in Balearic between marked and non-marked objects is topicality; other distinctions, such as definiteness and animacy, can also play a role, but this is secondary and limited to establishing the cut-off point within the set of topical objects. I will try to show that more fine-grained distinctions related to topicality can explain the apparent optionality of DOM for some topical objects. More specifically I will argue that the grammatical properties of the various constructions involving the detachment of a constituent (prepositional marking and the need for clitic resumption) correlate with the kind of topic involved.
To develop my argument, I will first discuss the way in which word order is related to information structure in Catalan. Then, I will briefly introduce the general notion of ‘topic’. Finally, I will show that topicality and other distinctions associated with it are the explanatory tools needed to offer a coherent picture of DOM in Balearic.

5.1 Word order and information structure in Catalan

In Catalan, unmarked sentences appear with a (S)VO order13 and the nuclear stress obligatorily falls on the rightmost constituent. Catalan prosody is very rigid and information structure contrasts cannot be achieved by shifting prosodic prominence, as in English. In fact, any attempt to modify the position of the nuclear stress while keeping the rest of the constituents in situ gives rise to an ill-formed structure, as shown in (45):

(45) Catalan (Vallduví 2002: 1230)
   a. Vaig ficar els ganivets al calaix.
     have.pst.1sg put the knives to.the drawer
     ‘I put the knives in the drawer.’
   b. *Vaig ficar els ganivets al calaix.
   c. *Vaig ficar els ganivets al calaix.

Whenever the informational content of a sentence has to be packed in a different way, a set of different syntactic patterns should be used. More specifically, the core clause is intrinsically focal, so non-focal phrases are systematically “repelled” from it and must be detached: this amounts to splitting the sentence into two informational units, either in a topic/comment structure or in a focus/background articulation. As a result, Catalan is a language with designated structural topic and focus positions (Vallduví 1994).

Focal constituents should, therefore, bear the nuclear stress. This is achieved by detaching other (potentially focal) constituents to a peripheral, deaccented position. For instance, if one wants to place the focus on els ganivets (‘the knives’), the following patterns are available:

(46) Focus (Standard/Central Catalan):
   a. Els ganivets vaig ficar al calaix.
      the knives have.pst.1sg put to.the drawer
      ‘I put the knives in the drawer.’
   b. Al calaix, vaig ficar hi els ganivets.
      to.the drawer, have.pst.1sg put there the knives
   c. Vaig ficar-hi els ganivets, al calaix.
      have.pst.1sg put-there the knives, to.the drawer
      ‘I put the knives in the drawer.’
In (46)a, the fronted constituent bears the nuclear stress, whereas the rest of the constituents of the sentence have been deaccented, the result being a contrastive focus structure. In (46)b-c the direct object again receives the nuclear stress due to the detachment of the locative phrase \textit{al calaix} (‘in the drawer’); it is ambiguous whether the expression has a contrastive or narrow focus.

Topics must be overtly separated from the informational update in the core clause due to the core clause’s focal nature. This can be done either by left- or right-detaching the topic expression.\textsuperscript{14} The more frequent topcalising construction is clitic dislocation (hereinafter, Cl-D). Consider the examples in (47), with the constituent \textit{els ganivets} treated as the topic:

(47) Topic (Standard/Central Catalan):
\begin{enumerate}
\item \textit{Els ganivets, els viag ficar al calaix}.
\begin{itemize}
\item \textit{els ganivets} = \textit{Els} \\
\item \textit{els} = \textit{i}
\item \textit{viag} = \textit{vaig}
\item \textit{ficiar al calaix} = \textit{ficar al calaix}
\end{itemize}
\begin{itemize}
\item \textit{Els} = \textit{the knives}
\item \textit{ganivets} = \textit{knives}
\item \textit{els} = \textit{them.obj}
\item \textit{viag} = \textit{have.pst.1sg}
\item \textit{ficiar} = \textit{put}
\item \textit{al calaix} = \textit{to the drawer}
\end{itemize}
\begin{itemize}
\item \textit{Els ganivets} = \textit{the knives}
\item \textit{els} = \textit{them}
\item \textit{viag} = \textit{have.pst.1sg}
\item \textit{ficiar al calaix} = \textit{put to the drawer}
\item \textit{Els ganivets, els} = \textit{the knives, the knives}
\item \textit{viag} = \textit{put}
\item \textit{ficiar al calaix} = \textit{put to the drawer}
\end{itemize}
\begin{itemize}
\item \textit{Els ganivets, els viag ficar al calaix} = \textit{the knives, the knives have.pst.1sg put to the drawer}
\item \textit{viag} = \textit{put}
\item \textit{ficiar al calaix} = \textit{put to the drawer}
\end{itemize}
\item \textit{Els, viag ficar al calaix, els ganivets}.
\begin{itemize}
\item \textit{Els} = \textit{them.obj}
\item \textit{viag} = \textit{have.pst.1sg}
\item \textit{ficiar al calaix} = \textit{put to the drawer}
\item \textit{els ganivets} = \textit{knives}
\end{itemize}
\begin{itemize}
\item \textit{Els, viag ficar al calaix, els} = \textit{them.obj, have.pst.1sg, put to the drawer, the knives}
\item \textit{Els ganivets} = \textit{the knives}
\end{itemize}
\item \textit{Els i vaig ficar al calaix, els ganivets}
\begin{itemize}
\item \textit{Els} = \textit{them}
\item \textit{i} = \textit{have.pst.1sg}
\item \textit{ficiar al calaix} = \textit{put to the drawer}
\item \textit{els ganivets} = \textit{knives}
\end{itemize}
\begin{itemize}
\item \textit{Els i vaig ficar al calaix, els} = \textit{them, have.pst.1sg, put to the drawer, the knives}
\item \textit{Els ganivets} = \textit{the knives}
\end{itemize}
\item \textit{Els, vaig ficar al calaix, els ganivets}
\begin{itemize}
\item \textit{Els} = \textit{them}
\item \textit{vaig} = \textit{have.pst.1sg}
\item \textit{ficiar al calaix} = \textit{put to the drawer}
\item \textit{els ganivets} = \textit{knives}
\end{itemize}
\item \textit{Els ganivets, els, vaig ficar al calaix}
\begin{itemize}
\item \textit{Els} = \textit{the knives}
\item \textit{ganivets} = \textit{knives}
\item \textit{els} = \textit{them}
\item \textit{vaig} = \textit{have.pst.1sg}
\item \textit{ficiar al calaix} = \textit{put to the drawer}
\end{itemize}
\item \textit{Els ganivets, els vaig ficar al calaix}
\begin{itemize}
\item \textit{Els} = \textit{the knives}
\item \textit{ganivets} = \textit{knives}
\item \textit{els} = \textit{them}
\item \textit{vaig} = \textit{have.pst.1sg}
\item \textit{ficiar al calaix} = \textit{put to the drawer}
\end{itemize}
\item \textit{Els ganivets, els vaig ficar al calaix, els}
\begin{itemize}
\item \textit{Els} = \textit{the knives}
\item \textit{ganivets} = \textit{knives}
\item \textit{els} = \textit{them}
\item \textit{vaig} = \textit{have.pst.1sg}
\item \textit{ficiar al calaix} = \textit{put to the drawer}
\item \textit{els} = \textit{the knives}
\end{itemize}
\item \textit{Els ganivets, els vaig ficar al calaix, els ganivets}
\begin{itemize}
\item \textit{Els} = \textit{the knives}
\item \textit{ganivets} = \textit{knives}
\item \textit{els} = \textit{them}
\item \textit{vaig} = \textit{have.pst.1sg}
\item \textit{ficiar al calaix} = \textit{put to the drawer}
\item \textit{els ganivets} = \textit{the knives}
\end{itemize}
\item \textit{Els ganivets, els vaig ficar al calaix, els ganivets}
\begin{itemize}
\item \textit{Els} = \textit{the knives}
\item \textit{ganivets} = \textit{knives}
\item \textit{els} = \textit{them}
\item \textit{vaig} = \textit{have.pst.1sg}
\item \textit{ficiar al calaix} = \textit{put to the drawer}
\item \textit{els ganivets} = \textit{the knives}
\end{itemize}
\end{enumerate}

The structure in (47)a is an example of clitic left-dislocation (Cl-LD): the dislocated constituent appears on the left and is resumed by the clitic \textit{els} (‘them’) that reproduces the features of the fronted expression; and the nuclear stress falls on the last constituent of the sentence — in the example under discussion, \textit{al calaix}, ‘to the drawer’ — which is part of the informative focus. The construction in (49)b is clitic right-dislocation (Cl-RD): the dislocated constituent is found at the right-most end, and is typically deaccented and realised as a low plateau; no pause is necessary (Prieto 2002: 11.2.1.4; Astruc 2004, 2005; Astruc & Nolan 2007); and again the clitic shows all relevant agreement features while the nuclear stress falls on the last constituent of the core sentence.

Vallduví (1992, 1994) has suggested an interpretive difference between left- and right-detached constituents: left-dislocated expressions are \textit{links}, i.e., they point to a specific locus for an informational update. Right-dislocated items, on the contrary, are \textit{tails}, that is, they further specify how the new information fits on a given file card or indicates a non-default mode of update (Engdahl & Vallduví 1996). This idea is similar to Lambrecht’s (1994) distinction between ‘topics’ and ‘antitopics’. Antitopics occupy the right periphery, cannot have a topic-creating function, are typically given or easily recoverable and are never contrastive.
5.2 Topicality

The systematic relation between dislocation and topicality in Catalan brings the notion of ‘topic’ to the foreground. Though fundamental to linguistic descriptions, it has remained a very controversial issue. A first source of difficulties is the fact that the term *topic* is used for both the entity referred to by a linguistic constituent and the linguistic expression itself. In addition, there is no consensus about the way in which topics have to be characterised. The traditional approach of the Prague school associated ‘topic’ with ‘old information’ and ‘familiarity’ was identified as the defining property of topics. In more recent times, however, this view seems to have been abandoned. This is not to deny, of course, that most topics refer to familiar, discourse-old or discourse-given entities, but familiarity does not seem to be a necessary condition for topichood, since a topic can also introduce new entities in discourse, such as the expressions *A friend of mine, A boy in my class* usually do (Reinhart 1981; Endriss and Hinterwimmer 2007). As a consequence, current research seems to favour the idea of topichood as *aboutness*: the topic is the entity the proposition is about.

**TOPIC EXPRESSION:** A constituent is a topic expression if the proposition expressed by the clause with which it is associated is pragmatically construed as being about the referent of this constituent. (Lambrecht 1994: 131)

The notion of *aboutness* is, however, a difficult one to elaborate in a theoretical way. Reinhart (1981) then has suggested to characterise topics not in terms of the past (as the relation of a topic to the previous discourse), but of the future, i.e., as the effect of the topic on the ongoing discourse. In this view, what a topic does is to provide a storage address at which the remaining information of the sentence has to be stored. As Krifka (2007: 41) puts it:

The topic constituent identifies the entity or set of entities under which the information expressed in the comment constituent should be stored in the C[ommon] G[round] content.

The characterisation of topics as providing an address to add the information in the comment makes it possible to encompass both the cases of ‘old’ (‘given’, or ‘inferable’) topics and ‘new’ topics under a single label.

Several authors have suggested further distinctions within the class of topics. These have to do with the activation status of the topic entity, which according to Prince (1992) can be measured from the point of view either of (the speaker’s beliefs about) the hearer’s beliefs or of the on-line construction of the discourse model. The two criteria cross-cut, as shown in the following table:
Discourse-new topics are independent from the hearer’s cognitive status, since a discourse-new entity can be either totally new to the hearer or just not activated, but nevertheless known or inferable; any discourse-old entity, on the contrary, must be old to the hearer as well.

Now the question is how topichood relates to syntactic constructions. Frascarelli and Hinterhölzl (2007) identify three types of topics according to the discourse function they perform: (1) shifting topic: a constituent that is “newly introduced, newly changed or newly returned to” (cf. Givón 1983: 8); (2) contrastive topic: an element that induces alternatives which have no impact on the focus value and creates oppositional pairs with respect to other topics (Kuno 1976, Büring 1999); (3) familiar topic: a given, discourse-linked constituent, generally used for topic continuity. They argue that each type has specific phonological and syntactic properties in both Italian and German.

With the above distinctions in mind, I will revise the Balearic data and put forward an explanation for the distribution of DOM.

5.3 Focal vs. topical objects

The previous considerations about word order and information structure point to a rather obvious conclusion: the basic split between marked and non-marked objects in Balearic is determined by information structure. This is indeed the main difference between Balearic and other Catalan dialects. As shown in examples (45)–(47) above, in Standard/Central Catalan the direct object els ganivets is always syntactically realised in the same way, regardless of whether it is a topic or a focus. In Balearic, on the contrary, the preposition is rejected with objects in the focal domain, but it is used to introduce exactly the same expressions when they are presented as topics:

(48) Balearic (Unmarked)
   a. Vaig ficar es ganivets an es calaix.
      have.pst.1sg put the knives in the drawer
      ‘I put the knives in the drawer.’
(49) Balearic (Contrastive Focus)
   a. *Es ganivets vaig ficar an es calaix*
      the knives have.pst.1sg put to the drawer
      ‘I put the knives in the drawer.’

(50) Balearic (Clitic Dislocation)
   a. *An es ganivets, els, vaig ficar an es calaix.*
      to the knives, them.obj have.pst.1sg put to the drawer
      ‘(As for) the knives, I put them in the drawer.’
   b. *Els, vaig ficar an es calaix, an es ganivets.*
      them.obj have.pst.1sg put to the drawer, to the knives
      ‘The knives I put in the drawer.’

Some other relevant contrasts are presented in (51)–(53):

(51) Balearic
    *Coneix* (*a) l’am o / (*an) en Joan / (*an) aquesta comedia.
    know.prs.3sg *(to) the_owner / *(to) the John / *(to) this comedy
    ‘S/he knows the owner / John / this comedy.’

(52) Balearic
       him know.prs.3sg *(to) the_owner / *(to) the John
       ‘S/he does know the owner / John.’
    b. *La coneix, *(an) aquesta comedia.*
       it know.prs.3sg *(to) this comedy
       ‘S/he does know this comedy.’

(53) Balearic
    a. *(A) l’am o / (an) en Joan el coneix.*
       (to) the_owner / (to) the John him know.prs.3sg
       ‘The owner / John s/he knows him.’
    b. *(An) aquesta comedia la coneix.*
       (to) this comedy it know.prs.3sg
       ‘This comedy, s/he knows it.’

Prepositional marking is thus independent of other features, such as definiteness
and animacy.

However, as mentioned above, left-detached objects can also appear without
the introducing preposition. My proposal is that, in order to obtain the right pre-
dictions with respect to the distribution of the prepositional marking, topicality
has to be divided further into different categories.
5.4 Right-dislocated definites

Among detached objects, the less problematic case is that of right-dislocated (RD) constituents: these are obligatorily marked with the preposition and linked to a coreferential clitic. Though RD has been sometimes considered as a type of repair of a performance error, an "afterthought", modern research has shown that its discourse function is not to correct an error, but to organize the discourse. The conditions determining the occurrence of RD constituents in English include entities either implicitly present in the situation or textually evoked (Grosz and Ziv 1996).  

Let us examine the context in which right-dislocation occurs in Balearic. Consider again the example in (38)a, repeated here as (54) for convenience, set in a wider discourse environment:  

(54) Balearic (Arxiu, Formentera)  
A woman is talking about cheese-making and the special container (escudella, ‘bowl’) with a base (peu, ‘base’; lit. ‘foot’) she uses to make the cheese:  
…perquè tenen com un peu que per girar es formatge va …because have.prs.3pl like a base that to turn the cheese go.prs.3sg molt bé, i ara no volen fer an aquell peu very well, and now not it.obj want.prs.3pl make to that base a ses escudelles to the bowls  
‘…because they have a sort of base that is very useful for turning the cheese, but nowadays they don’t want to make that base for the bowls anymore’  

The RD constituent, an aquell peu (‘to that base’) refers to an entity that has been mentioned in the previous utterance and is therefore active. Subsequently, aquell peu becomes the (persistent) topic of the following utterances:  

(55) Balearic (Arxiu, Formentera)  
… no sé perquè, però no, i n’hi han cantidat d’escudelles però, però planes de baix. I entonces per girar es formatge allò és molt difícil, i així amb so peu aquell s’aguanta s’escudella i es va girant damunt s’alta mà.  
‘I don’t know why, but they don’t; and there are a lot of bowls, but they are flat at the bottom. So turning the cheese becomes very difficult; and with that base one was able hold the bowl and turn it with the other hand’  

This suggests that Cl-RD picks a familiar entity and establishes it as a topic. However, this cannot be the whole story. Presenting an active entity as the topic does not require Cl-RD, but merely the use of a marker of high accessibility, such a pronominal. In fact, this is exactly what the pronominal clitic does: it indicates that the entity is highly accessible. The occurrence of the RD constituent must therefore
add further interpretive aspects. As mentioned before, Vallduví (1992: 63) has suggested that RD expressions are 'tails': they provide directions on how the information must be entered under a 'file card', namely by altering the nature of the update process, turning it from a 'retrieve_add' to a 'retrieve_substitute':

The presence of the tail indicates that a nondefault mode of update (UPDATE_REPLACE), involving further ushering, is (in the speaker's eyes) required. (Engdahl & Vallduví 1996: 45)

While I agree with the spirit of the proposal, I do not see the update operation as one of substitution (because the new information does not necessarily erase or replace the previously stored information), but rather one of adding partially conflicting and maybe unexpected information, which is presented as stronger or as more relevant. The idea is the following: an entity, which had been previously introduced in the discourse and associated with some informational content, is now presented in a context that partially corrects or contradicts the previously stored information; the contribution of Cl-RD is thus to mark both that the topic is a highly accessible, activated entity, and that the fact that the informational update represents a partial correction is no obstacle to adding it to the same storage address. In the case under discussion, the base of the bowl, which was highly relevant as a very useful feature of the bowls, is no longer a part of them; the information that the bowls with a particular base are very useful for cheese making must be kept in the corresponding storage address, but the information that the bowls no longer have that base is of a higher order. This is no surprise, since the detachment of the direct object allows a different constituent to receive the nuclear stress and then to be interpreted as the focus of the sentence. The interpretive strategy in this example is not that of replacing, but that of relativising the contradictory information to two different temporal stages (the previous and the current one). This is what licenses Cl-RD: the RD expression points to a contrast between two pieces of information that belong to the same address.

This explanation also applies to the occurrence of RD in interrogatives — an issue that has not been resolved in previous work. An RD constituent provides the storage address where the new piece of information that is presumed to be in conflict with previous knowledge should be added. This explanation also applies to the occurrence of RD in interrogatives — an issue that has not been resolved in previous work. For instance, after having learnt that the bowls with a base are no longer produced, the interviewer (a native speaker of Ibizan) asks the following question:

(56) Balearic (Arxiu, Formentera)

\[ I \ dins \ què \ el \ posau, \ an \ es \ [formatge]…? \]

and inside what OBJ put.PRS.2PL, to the [cheese]…?

‘So where do you put the [cheese] in?’
The CL-RD reactivates the cheese as the topic while at the same time asking for an update that could partially correct his previous idea that bowls with a base are still used to turn the cheese. The same goes for other examples, such as the one in (37), presented now in (57) with its preceding context:

(57) Balearic (Arxiu, Eivissa 3, Ibiza)

The interviewee was a school teacher and is talking about the schools in which she has worked. She is trying to remember the name of the principal of one of the schools, but does not succeed. So she says:

... i aquell tenia una escola allí, i ell, la vaig heredar jo an aquella escola, que era una habitació, que ara és l’Ajuntament...

‘...and that man had a school there, and he, so I took the reins of that school, which was a room, which is now part of the town hall...’

Again, the RD constituent had been introduced to the discourse and reappears in the next sentence. The RD constituent encodes an instruction to take the new focal content as modifying and downgrading the information previously stored under the address aquella escola (‘that school’), namely, that a particular man was the head of that school; this assumption must be made compatible with the fact that the interviewee became the next head of that school. The same explanation goes for (41), an example in which a conflict arises between being an indigenous species (one that probably one would like to preserve) and the suggestion of making it disappear.

It is worth mentioning that CL-RD structures can also be used in out-of-the-blue utterances, i.e., to refer to an entity that has not been previously mentioned in the discourse, but which is, however, easily accessible or inferable on the basis of general knowledge. I think this use fits in well with my proposal. Consider the example in (58):

(58) Catalan (Mayol, 2006)

L’oració Diga trenta-tres mereix una reflexió.

realmente la fan servir tant, els metges?

‘The sentence “Say thirty-three” deserves some thought. Do doctors really use it that much?’

Leaving aside the fact that the dislocated constituent is a subject, not an object, the referent of the RD constituent, as Mayol points out, is easily inferable from the situational context — the sentence being stereotypically used by doctors. According to my suggestion, RD is felicitous here because what the utterance does is precisely to suggest an informational update that conflicts with a standard or default previous assumption.
To sum up, with respect to Cl-RD, the situation is as follows: the construction is used to indicate that a new and surprising piece of information should be stored under the label of a highly accessible or familiar entity. In Balearic, the RD constituent is always introduced by the preposition a and a co-referential clitic should appear in the core sentence.

5.5 Left-dislocated definites

Left-detached (LD) topics usually co-occur with a resumptive clitic in the core sentence. As for their discourse function, my data suggest that there are two different situations in which LD constituents are felicitous. The first one is when the relevant entity has already been mentioned in the previous discourse but it is not the active topic at the moment; the role of LD is precisely to promote it as the topic. Consider example (27) in its wider setting:

(59) Balearic (Arxiu, Sant Josep, Ibiza)
...era una al·lota que s’estimava un jove i es de ca seu no el volien i entonces com que no el deixaven, an aquella al·lota no la deixaven parlar amb so que ella volia, pues hi anava a una hora que ells no hi fosse i se l’enduia i en haver-se-la enduit ja no li podien privar [...] que s’hi casàs...
‘...it was a girl who loved a boy and her parents didn’t like him and then, since they didn’t allow him, that girl they didn’t allow her to talk to the guy she loved, then he would go when they were not at home and take her away and once he had taken her away they could not prevent her from marrying him...’

Two entities are introduced in the first sentence: una al·lota (‘a girl’) y un jove (‘a young man’). Both are potential topics. In the next sentences, the young man is the topical entity and is referred to with a pronominal clitic (es de ca seu no el volien, ‘her parents didn’t like him’; no el deixaven, ‘they didn’t allow him’). In the next sentence, the attention is switched to aquella al·lota (‘that girl’) and the way to indicate this change of topic is the Cl-LD structure: the entity is discourse-old but was not the active topic at the moment.

Cl-LD is also felicitous when a demonstrative pronoun resumes a previously expressed idea as in (35), repeated here as (60):

(60) Balearic (Majorcan claumallorca)
...es turisme passa ara per hores baixes, the tourism-industry go.prs.3sg now through hours low
i an això es comerç ho pateéis.
and to this the trade it.obj suffer.prs.3sg
‘The tourism industry is going through tough times and business suffers from it.’

All rights reserved
The idea that the tourism industry is going through tough times, introduced in the first sentence, becomes the topic of the next one, resumed by means of the LD demonstrative pronoun això (‘this’) and its corresponding co-referential clitic ho (‘it’). Again, the topical entity (in this case, an idea) has been introduced in the discourse and is promoted as the topic.

The same sort of demonstrative resumption can be found in a discourse in which the speaker was talking about the differences between Catalan dialects with respect to the use of the form an in Balearic. After mentioning some sentences, he continues:

(61) Balearic (forums, Majorca)

An aquesta darrera frase noltros la diríem així: ‘Més tirria li tenen an en Mates.’ Amb sa n’de suports perquè en haver-hi dues vocals…

This last sentence we would say this way: ‘Més tirria li tenen an en Mates.’

With the ‘n’ in between. It is a supportive ‘n’, because when two vowels…

All the contexts of Cl-LD examined so far are instances of the same discourse situation: a previously mentioned entity, which was not the active topic at the moment, is promoted as an alternate topic. I will refer to this use as ‘switching topic’.

The second discourse environment where LD expressions appear involves the previous mention not of an entity, but of a larger set of which the entity can be considered to be a member. The LD expression indicates a new topic taken from a contextually active set. Consider the following dialogue:

(62) Balearic

– Has fet (tot) lo que et (ha)via dit?

‘Have you done (all) that I asked you to do?’

– An es plats ja els he rentat

‘The dishes I have done.’

The first speaker asks about a whole set of tasks (for instance, hanging up the clothes and washing up); the second speaker mentions one of them only: he introduces an entity related to one of the tasks as a new topic. The LD expression is interpreted as a ‘contrastive topic’: it is “an aboutness topic that contains a focus” (Krifka 2007: 44), so it induces an interpretation in which an alternative is implicitly established between a constituent and other possible entities of its kind previously introduced in the discourse. This is one of the most salient interpretations for Cl-LD; it has in fact been argued that CL-LD has a contrastive value in Catalan, that is, it indicates a choice of an element from a contextually established set. The topicalised constituent appears as a fronted element, is resumed by a clitic, and has
a rising intonation. Since a contrastive topic is a member of a contextually active set, other alternative topics can be overtly expressed as well, so the second turn in the dialogue in (62) could be continued as in (63):

(63) Balearic

...pero a sa roba no l’he estés encara.
...but to the cloth not it-have.prs.1sg hung yet
‘...but the washing I haven’t hung out yet.’

The LD constituent in a Cl-LD construction can thus serve two different discourse functions: either switching to an alternate topic or introducing a contrastive topic. What these two functions have in common is that the topical entities must be members of an active set. In both cases, the LD topic is chosen and singled out from among other potential entities, so they can be considered as ‘competing topics’, establishing an overt or implicit alternative between an entity and other possible entities of its kind previously introduced to the discourse. The main difference between them has to do with the way in which the competing alternates have been introduced: if it has been done by listing them (una al·lota, un jove, in (59)), the effect is that of switching to a highly related, alternate topic; on the contrary, if a set has been introduced through a defining or general property (lo que et (ha)via dit, ‘what I told you’, in (62)), the selected member contrasts with other potential members of the same set, by triggering the search for alternative members. This difference correlates with certain phonological features as well: switching topics are not phonologically independent from the core clause, whereas contrastive topics are pronounced with a rising intonation and tend to be separated from the core clause by a boundary tone H−.

In all the examples considered so far in this section, the LD expression is introduced by the preposition a(n). However, as mentioned before, there are examples without any preposition. Compare the example in (59) with the one in (64):

(64) Balearic (Arxiu, Ciutadella2, Minorca)
The speaker is talking about the way in which temporary workers’ salaries were established in ancient times:
Sa jugada era que un digués ‘per tal preu segam […]’ i es altres hi havien de venir bé, i aquell era es preferit, aquell cobrava es preu que realment volia cobrar, ja que es altres es (ha)vien de conformar, i aquell el pagaven de banda i ningú es dava compte de res.
‘The procedure was that one said “I’ll work for such a price […]” and the others had to agree and this guy was the favourite, that one would receive the pay that he really wanted to get, since the others had to bite the bullet and this guy they would pay on the quiet and no one would notice anything.’
The topic entity is the worker who suggests a salary in the negotiation, first introduced as un (‘one’), then resumed by aquell (‘this guy’, lit. ‘that (one)’). An intervening topic also appears: es altres, ‘the others’ (es altres hi havien de venir bé; ja que es altres es (ha)vien de conformar). In order to move back to the first topic, the speaker uses an LD structure. Thus, the conditions on LD seem to be exactly the same in both (61) and (66), but only in the first example the preposition is used.

The same goes for LD demonstratives. Compare (60) with (65) (= (41b)). The speaker is talking about the funds raised by NGOs:

(65) Balearic (Arxiu, Eivissa 3, Ibiza)

Diuen que llavor es sous no arribaven,
say.prs.3pl that then the money not arrive.pst.3pl, 
això no ha de dir sa gent.
this not it have.prs.3sg of say the people
‘They say that money didn’t then get [to its destination], this, people shouldn’t say.’

An analysis of the discourse environment reveals that there are no substantial differences between the examples with and without the preposition. In fact, to the best of my knowledge, there is no interpretive difference between the two series of examples (with and without the preposition), nor is there any difference in the discourse environment. I might possibly prefer the examples with the preposition, but the examples without the preposition still sound perfectly natural. This could be considered to be a case of micro-variation — an effect not unusual in the distribution of DOM. However, I think that this apparent optionality can be explained on more principled grounds.

In fact, there is a further construction involving left-detached object topics, the ‘hanging topic’ (HT) or nominativus pendens (‘hanging nominative’). The three main syntactic differences with respect to Cl-D are the following: 1) the detached constituent bears no case-marking; 2) the co-referential, case-marked element in the core of the clause is not necessarily a clitic; and 3) the detached expressions and the co-referential element do not necessarily match in case features. An HT is thus a sort of non-agreeing LD topic. The examples in (66)–(67) illustrate this construction:

(66) Balearic (Arxiu Eivissa 3, Ibiza)

– Quina altra obra vares escriure allí?
‘What other play did you write there?’

– Jo, em pareix que va ser…
I.nom me.obl. seem.prs.3sg that have.pst.3sg be
‘I, it seems to me that it was…’
(67) Balearic (Arxiu, Maó, Minorca)
   a. Noitros ses finestres mos donen a sa planada
      we.nom the windows us.obl give.prs.3pl to the plain
      ‘We, our windows look towards the plain.’
   b. En Gabaldon, som molt amic des fill, també
      the Gabaldon, be.prs.1sg very friend of.the son as_well
      ‘Gabaldon, I am a very close friend of his son as well.’

The co-referential element may be a clitic, as in (66) and (67)a, but it can be a strong pronoun or an epithet as well, or even an implicit element, as in (67)b.\(^{20}\)

The HT construction is used to introduce or shift attention to a new topic (‘shifting topic’); this new topic can be a brand new one, or a previously mentioned topic now revisited (Givón 1983).

Thus, my proposal is that the occurrence of the preposition with left-detached definites is never optional: the examples of LD topics without the preposition are not examples of an anomalous Cl-LD, but regular instances of HT. This means that Cl-LD is consistently constructed with the preposition, while HT is systematically non case-marked. If this proposal is to be accepted, one has to explain why both Cl-LD and HT can appear in the same discourse environments, as shown in the above examples. In the next section, I will try to provide a motivated solution.

5.6 A typology of topics and DOM

The data examined so far provide evidence for three kinds of topics:

1. Clitic-right dislocation indicates the storage address of a hearer-known, active entity, either already mentioned in the previous discourse or easily accessible via world knowledge, to which a new piece of information, presented as contrasting or unexpected, has to be added. The interpretive effect is that of a continuing topic. Continuing topics appear as right-dislocated, deaccented constituents associated with a resumptive clitic in the core clause.
2. Clitic-left dislocation introduces a competing topic, i.e., a new topic taken from a previously identified set. Two subcategories can be identified: if other potential topics have been overtly mentioned before, the construction makes it possible to switch from one to the other as alternate topics (‘switching topic’); if only a general set has been previously introduced — and therefore the relationship between this set and the newly introduced entity has to be inferred — the selection of one member of the set as a topic triggers the interpretive effect of contrast with other members (‘contrastive topic’). An alternative, either overt or implicit, is thus established between an entity and other entities of its kind previously introduced in the discourse. The entity that the Cl-LD
promotes as the topic, though known to the hearer (either by previous mention or by inference), is not the active topic in the immediate preceding discourse. The topicalised constituent appears as a fronted element, is resumed by a clitic, and, when contrastive, can have a rising intonation.

3. The hanging topic construction indicates the introduction of a new topic (‘shifting topic’). Two situations are possible: either the topic is a brand new one, or it is a previously abandoned topic now being recalled. The LD expression is not case-marked and is linked to a co-referential element (not necessarily a clitic) inside the core clause.

My proposal is that the relevant properties of these kinds of topics can be explained, in the spirit of Prince (1992), in terms of the status of the topical entity according to (the speaker’s beliefs about) the hearer’s beliefs, on the one hand, and to the contribution to the discourse model, on the other. I want to suggest that the full range of discourse functions of topicalising structures can be easily accounted for if we add two further dimensions on the side of hearer-based distinctions, namely the source of knowledge (general knowledge vs. previous mention) and the (psychological) activation status of the topical entity (active vs. inactive). The source of knowledge is necessary to include the differences between inferable and mentioned topics. The resulting distinctions are presented in Table 5, where the two main criteria (hearer’s knowledge and discourse novelty) are merged into a single schema to show how hearer- and discourse-based distinctions relate to each other. The shaded area indicates the extension of DOM.

Table 5. Types of topics and DOM in Balearic

<table>
<thead>
<tr>
<th>Hearer-known</th>
<th>Source</th>
<th>General knowledge</th>
<th>Previous mention</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Active</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Continuing topic</strong></td>
<td><strong>Clitic/Right Dislocation</strong></td>
</tr>
<tr>
<td><strong>Inactive</strong></td>
<td></td>
<td><strong>Competing Topic</strong></td>
<td><strong>Clitic Left Dislocation</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Contrastive topic</strong></td>
<td><strong>Set membership</strong></td>
</tr>
</tbody>
</table>

All rights reserved
Now, the basic assumption of my proposal is that Cl-D and HT are each specialised for encoding topicality from a different point of view: Cl-D encodes topicality as a hearer-based notion, i.e. as linked to the hearer’s presumed cognitive status; HT, on the contrary, encodes topicality as a discourse-based function. The two perspectives are independent from each other; however, as the table shows, the two criteria overlap in a certain area, namely that of alternate topics. In fact, both Cl-LD and HT can be a means of switching to a topic that is inactive from the hearer’s perspective and new from a discourse perspective: Cl-LD presents it as chosen from among an array of previously mentioned entities (‘switching topic’); HT introduces it as a topic newly returned to (‘returning topic’). This is precisely the area where the prepositional marking seems optional.

Thus, we find a principled answer to the question of why certain discourse environments are compatible with the two constructions and why prepositional marking seems to be optional in a particular discourse situation. According to my proposal, in the case of alternate topics, the speaker can choose which point of view s/he adopts: either to encode the instruction from the hearer’s perspective, as the change to a known, but inactive topic, or to take the discourse perspective and present it as a new topic.

If this explanation is on the right track, certain consequences can be drawn. First, the occurrence of the preposition is never optional in Balearic, but related to different structures: it will be needed in Cl-LD, not in HT. Second, it will be always possible to predict the discourse environment in which the preposition can appear or not: it will involve situations with alternate topics, which can be seen either as switching topics (and hence constructed as Cl-LD) or as returning topics (presented as HT). Third, it will be possible to predict that the preposition is compulsory to introduce LD constituents expressing a contrastive topic — a discourse function exclusive of Cl-LD. Fourth, we can also predict that brand new topics will occur without the preposition. All these predictions seem in fact to be borne out.

5.7 Detached indefinites

The previous distinctions explain the distribution of DOM with topical definites. In this section I will argue that at least part of the behaviour of indefinites can be explained along the same lines that determine the distribution of DOM with definites; in fact, indefinites provide a further piece of evidence for my analysis.

If DOM in Balearic is sensitive to topicality, it is no surprise that examples with indefinites are scarce and unstable, since, as is well known, indefinites are poor candidates for topic positions. This is true, for obvious reasons, of indefinites that introduce brand new referents without any kind of basis in the previous discourse. However, as shown by Engdahl & Vallduví (1994), Villalba (2000: chs. 2 & 3) and
Vallduví (2002: § 4.5.3.2), this does not mean that all indefinites are systematically excluded from topicalisation. There are two cases in which indefinites can be successfully topicalised. The first is represented by intensional, or non-referential, indefinites, i.e. indefinites that do not introduce new referents, but which merely provide descriptive content that makes it possible to identify not an individual but a class; this is consistent with the assumption that, in addition to individuals, sets can also serve as addresses for storing information (Endriss and Hinterwimmer 2007). The examples in (2)a-b above are instances of topicalised intensional indefinites. The second case is that of discourse-anchored indefinites, i.e., indefinites that selected the new referent from a previously mentioned set, as in (2)c.21

With respect to DOM, as mentioned above, there is no agreement among the informants. The proposal I have presented in the previous sections makes it possible to explain some of the contrasts with regard to the acceptability of the prepositional marking with indefinites. The predictions are that right-dislocated indefinites and Cl-LD contrastive topics are the most favourable environments for the occurrence of the preposition, whereas alternate topics will tend to appear without the preposition. Therefore, the relevant environments to test the extension of DOM with indefinites are RD and contrastive Cl-LD. As for left-detached constituents, when tested without any specific context, speakers usually prefer the indefinite without the preposition; however, when presented with a context that introduces an overt choice among alternatives, they immediately tend to use the prepositional topic. Thus, in the ice cream examples, when asked to choose a dessert, or to choose between coffee and ice cream, they naturally produce or accept examples with the preposition, such as (68):

(68) Balearic

%(A un cafè no, però) a un geladet sí que me'l prendria,
(to a coffee not, but) to an ice.cream, yes that me-it have.COND.1sg
ara, amb molt de gust.
now, with much of pleasure
‘(Not a coffee, but) an ice cream I'd love to have.’

This is exactly the result one should expect when Cl-LD has a contrastive value. Some of my informants overtly explain that whenever there is a choice, the preposition should appear.

At this point, an additional feature, animacy, seems to enter the picture. For some speakers, the preposition can be accepted with Cl-LD animate indefinites (as in (43)c above), but is ruled out with inanimate indefinites (as in (68)). Since animates always appear higher than inanimates on the scale of inherent properties, the implication is that those who accept (68) with an inanimate indefinite will
accept animates as well, as (43)c above — and in fact this is the case — but not necessarily the other way round.

The results are not substantially different with RD: for those who do not accept the preposition, the examples with animate indefinites are better than those with inanimate objects. In addition, given that indefinites represent the end of the definiteness scale, it is not surprising that they show the greatest degree of variation.

Interestingly, the same analysis can be extended to the related partitive constructions. It has been assumed that the partitive preposition is optional for dislocated constituents in Balearic. Compare the following examples:

(69) Balearic
   a. (De) fruita, en hi ha molta, enguany
      (of) fruit, of.it.cl there have.prs.3sg many, this.year
      ‘This year, there is a LOT of fruit.’
   b. (De) cans, en he vist només un
      (of) dogs, of.them.cl have.prs.1sg seen only one
      ‘I saw only ONE dog.’

(70) Balearic
   a. En hi ha molta, enguany, *(de) fruita
      of.it.cl there have.prs.3sg many, this.year, *(of) fruit
      ‘This year, there is a LOT of fruit.’
   b. En he vist només un, *(de) cà
      of.them.cl have.prs.1sg seen only one, *(of) dog
      ‘I saw only ONE dog.’

As the examples reveal, the apparent optionality only affects Cl-LD constituents. My proposal is that when the preposition appears in the sentences in (69), the construction is Cl-LD; when it is absent, we have a HT instead. If so, one would expect that only Cl-LD partitives could have a contrastive interpretation and indeed this seems to be the case:

(71) Balearic
   De pomes, en hi ha hagut moltes;
   of apples, of.them.cl there have.prs.3sg had many;
   de peres, en canvi, poques.
   of pears, in change, few
   ‘Apples, there have been many; pears, on the other hand, only a few.’

If the preposition were dropped, the sentence would sound unnatural in the intended discourse situation. When a new topic is introduced without any sort of contrastive interpretation, the NP without preposition has to be used:
(72) Balearic
a. *Vent, no en fa massa, avui.* 
wind, not of.it.cl make.prs.3sg much, today
‘It isn’t very windy today’
b. *Cabres, en hi ha, encara?* 
goats, of.them.cl there has, still?
‘Are there still any goats?’

(73) Balearic (Arxiu, Ciutadella2, Minorca)
a. *Sa festa de ses toses també era de feina, però ses toses, hi havia, era una cosa molt particular,*
ses toses  hi  havia  molt  d’amor propi. 
the sheep.shearing there have.pst.3sg much of.pride
‘The sheep shearing festival was also a time for work, but the sheep shearing, there was, it was something very particular, in the sheep shearing there was a lot of pride.’
b. *I tot açò hi era, eh? I formatjades hi havia moltes.* 
que en feien moltes. 
that of.them make.pst.3pl many.
‘And there were all these things. And cheese pastries: there were places where they used to make a lot of them.’

A further piece of evidence pointing in the same direction comes from an independent phenomenon related to the use of the preposition with interrogatives. As mentioned above, there is a further allomorphic variant of the preposition *a*, the form *ane* [әnә], which is not permitted in the written standard. As shown in (74)a, *ane* is the (non-standard) form of the preposition for interrogative datives:

(74) Non-standard Catalan (Fabra, 1918: 112)
a. *Ane qui ho enviarem?* 
to who it send.fut.1pl?
‘To whom will we send it?’

Standard Catalan
b. *A qui ho enviarem?* 
to who it send.fut.1pl?
‘To whom will we send it?’

Interrogative accusatives are not marked, as shown in (75); however, in Balearic the form with [әnә] can also introduce [+human] accusative interrogatives, as in (76):
(75) Balearic/Standard/Central Catalan

Qui has vist?

who have.prs.2sg seen?

‘Who have you seen?’

(76) Balearic

[әnә] qui has vist?

to who have.prs.2sg seen?

‘Whom did you see?’

The occurrence of the preposition is quite unexpected. What is relevant for our current discussion is that (76) has a contrastive interpretation, exactly as in Cl-LD. Thus, (75) is a neutral question, whereas (76) is a sort of clarification question, which presupposes that a choice set is active. The context illustrated in (77) would be adequate:

(77) Balearic

Pero tu, [әnә] qui has vist, (an en Joan o a na Maria)?

but you, to who have.prs.2sg seen (to the John or to the Mary)?

‘But, whom did you see (John or Mary) in the end?’

In this case, the prepositional marking has been analysed as a contrastivity marker.

As for the origin of the variant [әnә], I propose that it is the result of double prepositional marking. The idea is the following: first, the preposition a was used to case-mark datives, yielding a qui (‘to whom’); next, this expression was reanalysed as a single word, and the feeling that the first segment was actually a case-marking preposition was lost.22 Then, given the need to overtly case-mark datives, the speakers added the preposition again. This time, however, the phonological environment was different because the new word a-qui began with an unstressed vowel, so that the an variant of the preposition had to be chosen, yielding an a qui: a qui > a-qui > an a-qui. The written form ane thus conceals the internal structure of the expression. Interestingly, however, this is not in fact an isolated case. Bossong (1991: 154) found a similar process in the Rhaeto-Romance dialect of the Engadin Valley (Canton Graubünden, Switzerland). There, the first person pronoun me merged with the preposition a, yielding ame. This new form was reanalysed as non-case-marked, and again received the preposition, with the insertion of an anti-hiatic consonant d: a me > a-me > ad ame.
5.8 The prepositional marking of pronouns

In the previous sections, I have discussed the role of topicality as the governing factor determining the split between marked and unmarked objects. However, there seems to be an exception to this generalisation, that of strong pronouns, which are always marked both when topicalised and when occurring in the domain of focus.

The marking of pronouns has to be related to the functional motivation for DOM. As Bossong points out,

… essentially differential object marking must be considered as a kind of anti-marking device. In typical DOM languages, only those direct objects tend to be marked which share a more or less great amount of semantic features with prototypical subjects while those direct objects showing prototypical object properties tend to be left unmarked. As one proceeds from right to left in the inherence scale, prototypical object properties gradually diminish, and at the same time the probability for objects to be positively marked gradually increases. Positive object marking inside a DOM system marks subject-like objects. […] The general underlying principle can be described in terms of a natural markedness theory whose domain is in the iconic distribution of morphological markedness patterns: disagreement of slot and filler properties favors markedness, harmony favors unmarkedness. (Bossong 1991: 162–163)

In Balearic several additional factors can also contribute to the marking of pronouns. First, the fact that Balearic strong pronouns have only an unmarked case (=nominative).23 Compare (7) above with its Balearic counterpart:

(78) Balearic

Io t’ ajud a tu i tu m’ ajudas a io.

'I’ll help you and you’ll help me’

In addition, pronouns share a most significant property with prototypical topics, namely, that of referring to hearer-known antecedents. By their own nature, pronouns are intrinsically anaphoric, so they necessarily point to highly accessible entities. In this respect, the notion of secondary topic found in Nikolaeva (2001) and Dalrymple and Nikolaeva (2007) can prove useful. At the time of the utterance more than one referent can be under discussion, so that the information provided increases the knowledge of more than one entity. Topics can be ordered with respect to saliency: primary topics are more pragmatically salient then secondary topics. Pronouns are, therefore, perfect candidates for secondary topics, even if they occur in a rhematic or focal position. If so, the fact that they are always marked, both in canonical and in dislocated positions, comes as no surprise.
6. DOM and topicality in Romance

The fact that topicality is a crucial dimension in the distribution of DOM is not uncommon in other Romance languages. The idea that topics are the root of DOM can be found in the proposals by Rohlfs (1971), Pensado (1985), Laca (1987, 2006), and Melis (1995), among others, for the Spanish prepositional accusative. The marking of topical definites is, in fact, what we find in medieval texts, as the following examples show:

(79) Medieval Spanish (Cid 2251 and 275, taken from Laca 2006)
   a. Escarniremos las fijas del Campeador.
      ridicule.fut.1pl the daughters of-the Campeador
      ‘We will ridicule the daughters of the Campeador.’
   b. A las sus fijas en braço las prendia.
      to the his daughters in arm them.obj took.pst.3sg
      ‘His daughters he carried in his arms.’

Sardinian, a language with DOM, illustrates a situation similar to that of medieval Spanish: the preposition can be omitted with rhematic objects, but never with Cl-dislocated constituents. Compare the following examples:

(80) Sardinian (Jones 1995: 44)
   a. Appo invitadu (a) su preideru a su matrimoniu
      have.prs.1sg invited (to) the priest to the marriage
      ‘I invited the priest to the marriage.’
   b. *(A) su preideru, invitadu l’as a su matrimoniu?
      *(to) the priest, invited him-have.prs.2sg to the marriage?
      ‘The priest, did you invite him to the marriage?’

Even in Italian, a language without DOM, the preposition can occur with some dislocated constituents. According to Cardinaletti (2002), in colloquial Central and Southern Italian, a right-dislocated [+human] accusative can be preceded by the preposition a; without the clitic, the sequence would be entirely ungrammatical, as in the following example:

(81) Colloquial Central Italian (Cardinaletti 2002: 33)
   *(L’)abbiamo invitato noi, a Gianni
   *(him)-have.prs.1pl invited we, to John
   ‘We invited John’

It is thus also possible to mark the dislocated item with the preposition in colloquial Italian as well.
Modern Romanian is a language with an extensive use of DOM, though the conditions governing the occurrence of the preposition are a rather complex issue (see von Heusinger & Onea, 2008). It is worth noting, however, that even non-human indefinites can be marked with pe when they are topicalised, very much like in Balearic (cf. (43)):

(82) Romanian (Farkas & von Heusinger, 2003)

(Pe) una din casele vecinului le-am fotografiat.

(to) a of houses.def neighbour.def.gen it-have.prs.1sg photographed

‘One of the neighbour’s houses, I have photographed.’

(83) Romanian (von Heusinger & Onea, 2008)

Un tânăr, mergând pe camp, adună două pietre.

A young_man going on field gather.pst.3sg two stones.

Pe una din ele o luă și o puse în buzunar.

to one of them it take.pst.3sg and it put.pst.3sg in pocket

‘A young man, crossing a field, picked up two stones. One of them he took and put into his pocket.’

The most striking examples, however, come from Porteño Spanish (the variety spoken in Buenos Aires and the River Plate area). Consider the following examples:

(84) Porteño Spanish (poster in a public square, Dumitrescu 1998: 140)

a. A esta plaza la cuidan Aerolíneas Argentinas y usted.

to this square it care.prs.3pl Aerolíneas Argentinas and you

‘This square, Aerolíneas Argentinas and you take care of it’

Porteño Spanish (Masullo 1992: 291)

b. A los libros los envolvió en papel madera.

to the books, them wrap.pst.3sg in paper wood

‘The books, s/he wrapped in wood paper.’

Porteño Spanish (Masullo 1992: 291)

c. Sí, los puso sobre la mesa, a los libros.24

Yes, them put.pst.3sg on the table, to the books

‘Yes, the books, s/he put on the table.’

The examples in (84) show non-human definites introduced by the preposition a. This possibility is totally excluded in European Spanish. Notice that all sentences are examples of clitic-dislocation constructions. Dumitrescu (1997; 1998) indeed shows that if the word order is changed and the dislocated constituent appears as a rhematic direct object, the preposition is systematically dropped:
Differential object marking and topicality. The case of Balearic

(85) Porteño Spanish
a. Aerolíneas Argentinas y usted cuidan (*a) esta plaza.
b. Envolvió (*a) los libros en papel madera.
c. Puso (*a) los libros sobre la mesa.

What this behaviour shows is that the constraints governing the distribution of DOM in Porteño Spanish have remarkable similarities to the pattern that has been described for Balearic in the preceding sections.

A final piece of evidence for the view advocated in this paper comes from an apparently unrelated phenomenon: that of “object dropping” in Brazilian Portuguese and in the Spanish dialects spoken in Quito (Ecuador) and Paraguay. As shown in Schwenter (2006), this phenomenon is sensitive to topicality:

...the greater the topicality of the anaphoric D[irect] O[bject] in B[razilian] P[ortuguese], the more likely its referent was to be encoded overtly; the lesser the topicality, the more likely its referent was to occur encoded as a NULL OBJECT.

(Schwenter 2006: 33)

Significantly, the conditions governing the split between null and overt pronominal objects in these dialects systematically correlates with the distinction governing the DOM split in (Pan-)Spanish:

Table 6. Parallels in object coding (Schwenter 2006: 33)

<table>
<thead>
<tr>
<th></th>
<th>Brazilian Portuguese</th>
<th>Quiteño Spanish</th>
<th>Paraguayan Spanish</th>
<th>Pan-Spanish</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower topicality</td>
<td>Null pronominal (Ø)</td>
<td>Ø + Object NP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less overt coding</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Higher topicality</td>
<td>Overt pronominal (ele/ le)</td>
<td>Accusative a + Object NP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>More overt coding</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In an informal survey on Paraguayan Spanish, I have found that left-detached objects tend to be treated as lower in the topicality scale than right-dislocated constituents. The examples in (86) show the relevant contrast:

(86) Paraguayan Spanish (naturally occurring examples)

a. Esta puerta Ø dejo abierta por si se metió el gato.  
   this door Ø leave.prs.1sg open for if cl get.in.pst.3sg the cat  
   ‘I’m leaving this door open in case the cat got in.’

b. Y con esta pieza le limpio al radiador.  
   and with this piece it.cl clean.prs.1sg to.the heater  
   ‘And with this piece I can clean the heater.’
In (86)a the left-dislocated constituent is an HT, which is not introduced by the preposition, nor is there any overt clitic in the core clause. In (86)b, by contrast, we find both the overt clitic and the prepositional marker for the co-referential expression. This distribution is again very similar to the Balearic data.

7. Conclusion

The data presented in the previous sections have shown that Balearic exhibits DOM in a wider and more complex distribution than is usually assumed for Catalan. The account of DOM in Aissen (2003) — probably the most influential study on the subject — is based on the semantic features of the object DP. The intrinsic properties of the DPs involved, however, cannot fully explain the distribution of DOM in Balearic, since the very same constituent with the very same semantic properties can appear either with or without the preposition, depending on its informational status.

I have argued that the major dimension governing the distribution of DOM in Balearic is topicality. As a language with designated topic and focus position, Balearic is particularly transparent with respect to information structure: topics must occur overtly detached from the informational update. Balearic can encode topicality from two different points of view: from the point of view of the hearer’s knowledge and from the discourse perspective. Hearer-known topical entities are syntactically realized as clitic-dislocation constructions; discourse-new topical entities are presented using the hanging topic structure.

Hearer-known topical expressions are always marked, no matter whether they are active or inactive, previously mentioned or merely accessible, as shown in Table 5 above. This is the case of continuing topics (which reproduce previously mentioned or highly accessible entities) and of competing topics (which select and elaborate on a previously established set). All dislocated definites are consistently marked with the introducing preposition *a* and clitic-doubled with a co-referential weak pronoun in the core clause. As for indefinites, those that meet the independent requirements on topichood can also be marked in the same way, though for a number of speakers an additional requirement of animacy seems to be in force.

On the contrary, discourse-new topics (i.e., topics presented as new from the discourse perspective, regardless of the hearer’s cognitive status) are not marked: they are not introduced by the preposition, nor do they require clitic resumption.

Pronouns seem to qualify as (secondary) topical entities: even if they occur in the focal domain, they are intrinsically anaphoric and encode a very high degree of accessibility from the hearer’s perspective, thus complying with the condition of being hearer-known and providing a salient secondary storage address for the
informational update. Pronouns are thus marked as well: they are introduced by the preposition and clitic-doubled by default.

The data thus support the idea that Balearic DOM is governed by a complex bundle of discourse properties (topicality and hearer’s knowledge), referential status (uniqueness, identifiability), and intrinsic features (animacy). The complexity of this system can be reduced, as suggested by von Heusinger and Kaiser (2003), if we consider that some distinctions are activated only when certain conditions are reached in another dimension. This is exactly what happens in the present case: definiteness turns out to be relevant only when another pragmatic condition is met, namely hearer-old topicality; in a similar way, animacy becomes relevant only when DOM has gone beyond definites along the definiteness scale. The result is, then, a nested hierarchy in which the informational status of the constituent ranges over the referential status (definiteness), which ranges over the intrinsic features (animacy).

Table 7. Kinds of topics and DOM in Balearic

<table>
<thead>
<tr>
<th>TOPOCALITY</th>
<th>DEFINITENESS</th>
<th>Pronouns</th>
<th>Proper names</th>
<th>Definites</th>
<th>Indefinites</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Animate</td>
</tr>
<tr>
<td>Primary topics</td>
<td>Discourse-new topics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Detached)</td>
<td>Hearer-known topics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Secondary topics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(In situ)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The existence of both semantic and pragmatic constraints and the fact that, if operating together, they are hierarchically ordered, are not unusual features in the distribution of DOM. In their thorough examination of a large number of languages, Dalrymple and Nikolaeva (2007) have recently shown that languages with DOM fall into three main categories. The first type corresponds to systems such as Ostyak, in which DOM is dependent on information structure only. Type 2 is assigned to languages where DOM is governed exclusively by semantic features, as in Hungarian and Hebrew. The most widespread type, however, is the third one, in which both information structure and semantic features have a bearing on the distribution of DOM. This type is further divided in two subtypes, depending on the relative strength of informational and semantic features. The second subtype applies to languages such as Hindi or Persian, in which objects with the same semantic features are either marked or unmarked, depending on their information structure role. Balearic represents, thus, a further instance of this general subtype.

The pattern of Balearic, in fact, is not unique; other Romance languages and dialects also show remarkable asymmetries in the construction of topical and
non-topical objects. The data and the account presented here offer new evidence for the idea that topicality is a relevant factor in the emergence and the development of DOM. As Laca (1995) pointed out, the high topicality of the object and its independence within the predicate are the original functions of the prepositional accusative in Spanish. This further supports the idea that the main motivation of DOM is to mark non-prototypical objects:

Prototypical objects form an integrated part of the verbal complex. This is particularly important in the pragmatic domain: verb and direct object prototypically belong to the same pragmatic constituent, usually (although not necessarily) the rHEME (comment) of the utterance. (Bossong 1991: 158)

Not surprisingly, the conditions determining the appearance of DOM are very similar to those that explain the extensions of the definite article in Romance, as pointed out in Company (1991) for Spanish: Subject NPs are usually referential and definite and tend to be information that has already been presented or is mutually known to the participants, particularly if the subject is also human and agentive; object NPs, on the contrary, are expected to be new information. Therefore, subjects were marked first to overtly indicate their referential and topical status.

In addition, topicality seems to be the most relevant parameter for clitic-doubling in Spanish, a phenomenon with a strong relation to DOM. Clitic doubling is a formal way to emphasise “the heightened topicality of the D[irect] O[bject] referent vis-à-vis what is normally the case for DOs.” (Schwenter 2006: 34). Therefore, the systematic correlation between prepositional marking and clitic resumption in Balearic cannot merely be a coincidence.

These considerations suggest that information structure distinctions must be taken into account to obtain a better understanding of the parameters responsible for DOM.

Notes

* This research has been partially supported by the Spanish Ministry of Education through the grant HUM2006-06630 for the project “SPYCE: Semántica procedimental y contenido explícito”. Previous versions of this work were presented at the 3rd NEREUS International Workshop “Definiteness, Specificity and Animacy in Ibero-Romance Languages” (Universidad de Alcalá de Henares, Madrid, October 6–7, 2006), and at the Workshop “Differential Object Marking” (Universität Stuttgart, June 15, 2007). I am very grateful to the audiences for their helpful discussion. I am particularly indebted to Ileana Comorovski, Klaus von Heusinger, Georg Kaiser, Udo Klein, Manuel Leonetti, Edgar Onea, Gemma Rigau, and Begoña Vicente for detailed comments on earlier drafts. Many thanks are due to the anonymous Studies in Language reviewers for very useful suggestions and criticism. I am also very grateful to Aoife Ahern for checking my English. Needless to say, I bear the whole responsibility for the ideas I finally adopted.
1. I am taking for granted that DOM in case-marking and DOM in agreement are instances of the same phenomenon. This issue, however, is not uncontroversial. See Bickel 2008 for discussion.

2. Clitic dislocation has well known syntactic properties, such as recursivity, an ability to occur both in root and in embedded sentences, and sensitivity to islands. The question whether the dislocated constituent reaches its position as the result of movement, or whether it is directly merged there, is a much debated issue that I will not pursue here. For further details on clitic dislocation, see Cinque 1983; Zubizarreta 1998; Vallduvi 1992, 2002; Villalba 2000; López 2003; Anagnostopoulou 2006; Suñer 2006.

3. There are in fact two allomorphic variants of the preposition a [ә]: an [әn] (Cf footnote 6 below) and ana/ane [әnә]. However, as far as I know, the variant [әnә] is never used as the quote and the example suggest. See Section 5.6.

4. When analysing prepositional accusatives in Catalan, a caveat should be born in mind. In the Eastern dialects, the unstressed preposition a is pronounced as a schwa sound [ә]. When it is adjacent to another schwa sound, they merge (Vallverdú Albornà 2002: 144–145). This happens with the final sound of some verbal endings and with initial unstressed e- and a-, pronounced also as [ә]. As a result, when hearing the sequence [әs’timal’perә] it is impossible to know whether the speaker is using the preposition or not, i.e., whether s/he is saying Estima el Pere, (‘S/he loves the Peter’, without the preposition) or Estima al Pere (‘S/he loves to the Peter’ with the preposition). Since the speakers of Catalan have not been taught basic literacy in their own language until very recently, it is no surprise that a number of old (or even modern) texts could render the direct object with or without the preposition in cases of vowel merging. In written texts it is thus difficult to tell apart cases of prepositional accusative from cases of mistaken spelling; and vice versa, and some examples of real DOM might have been obscured by wrong spellings. The only safe approach is to consider expressions in which vowel merging does not take place.

5. The extreme case of DOM in the area is represented by Valencian, where DOM shows up with the same pattern of distribution that has in Spanish (Sancho Cremades 1995: 199):

(i) No veig {a Cast / als xiquets} des de fa sis anys
not see.prs.1sg {to Cast / to-the children} from six years
‘I haven’t seen Cast/the children for six years’

This case will not be considered here.

6. The nasal infix occurs when the preposition precedes determiners and pronouns that begin with an unstressed vowel, such as masculine definite articles (es, el, en ‘the’; see footnote 8) and demonstrative determiners and pronouns (aquest/a, ‘this’; aquell/a, ‘that’; això, ‘this’). Given the non-standard nature of this form, there is no accepted way to spell it. The usual spelling in ancient Catalan texts, from Llull to Maragall, was an:

(i) 13th century (Ramon Llull: Astronomia)
.. e anaxi de les autres natures semblants an aquestes
and likewise of the other natures similar to these
‘…and the same for other natures similar to these’
(ii) Early 20th century (Joan Maragall: La ciutat del perdó)

An aquest i an aquell els salvaren.

to this and to that them save.pst.3pl

'They saved this one and that one.'

The traditional spelling an is the form found in the transcriptions of the oral corpora I used and it is the more frequent form in the Balearic Islands. Some less acceptable spellings can also be found, such as anàquest, an aquest. The form with hyphens (a-n-aquest) is usual in the Catalan dialects of Southern France.

7. My Balearic examples come from various sources. Most of them come from two oral corpora: the Arxiu audiovisual dels dialectes catalans de les illes (Arxiu), and Corpus Oral Dialectal (COD) del català contemporani. The interviews taken from the Arxiu... correspond to different small towns on the islands of Minorca, Ibiza, and Formentera; they are labelled according to the name of the town. The interviews in COD come from different small towns on all the islands and are labelled according to the name of the town and the initials of the interviewee. Other examples have been taken from written texts, Internet blogs and chats, and spontaneous speech. Finally, others have been constructed by myself as a native speaker of Ibizan to illustrate different points; in this latter case, both the form of the examples and their interpretation have been tested with a sample of native speakers (14 individuals: 9 females and 5 males, between 20 and 90 years of age, with different social and cultural backgrounds). In all the examples, I have kept the original spelling.

8. An anonymous reviewer points out that the occurrence of the preposition in this case might be due to the animacy of the referent (tots dos, 'both of them' = the boy and the girl). Notice, however, that the preposition would also show up if the referent were inanimate, as in the following dialogue:

(i) Balearic

- Has pogut trobar es llibres?
  have.prs.2sg can find the books?
  'Did you finally get the books?'
- Els he comprat (a) tots dos.
  them have.prs.1sg bought (to) all two
  '(Yes) I bought both of them.'

9. Three different definite article forms are found in Balearic. The form es/sa (<ipse, ipsa) is the usual one (cf (24)b). The form el/la (<ille, illa) is common to the rest of Catalan dialects; in the Balearic Islands it is used only with some unique definites, such as el rei ('the king'), el bisbe ('the bishop'), l’ajuntament ('the town hall'), la catedral ('the cathedral'); also la casa ('the household') and l’amo ('the owner, the master') bear this form (cf. (26)). The form of the definite article, en/na (<dominus, -a) is used to introduce proper names, as in example (24)a.

10. In these examples, the mark appears in its phonetic transcription because the author refused to write the non-standard form.

11. The examples (37)b and (38)b were uttered by the interviewer, a native speaker of Ibizan.

12. See Section 5.7 for further details on indefinites.
13. This is the more frequent surface word order in unmarked sentences. As one of the reviewers pointed out, Vallduví (1993) has argued that Catalan is a VOS language, claiming that all preverbal subjects are in fact left-dislocated constituents as a result of information packaging needs. The reason why it should be so will become clear in the text below.

14. Therefore, Vallduví (1993, 2002) argues that Catalan appears as an SVO language as a result of the left-dislocation of the subject. There is no clitic for them in Catalan.


16. The effect of substitution in Vallduví’s examples, such as

\[(i) \text{'L’amo l’ODIA, el bròquil. the boss HATES broccoli.}\]

is due to the contrastive focus on the predicate, not to the occurrence of the RD constituent.


18. This account is consistent with the condition established by Ward and Prince (1991:171) for English topicalisation, which seems to be at work for Balearic as well:

**Discourse Condition on Preposing in Topicalization:** The entity represented by the preposed constituent must be related, via a salient partially ordered set relation, to one or more entities already evoked in the discourse model.

19. According to Vallduví (2002), hanging topics are sentence fragments, and their intonational contour is similar to that of an interrogative. For a discussion of other grammatical properties of hanging topics, see, among others, Cinque 1983; Anagnostopoulou 2006; Suñer 2006; De Cat 2007. For Catalan, see Villalba 2000 (ch. 2) and Vallduví 2002.

20. In this case the implicit argument of the relational noun fill (‘son’) represents the resumptive element inside the clause.

21. As an anonymous reviewer notes, this difference correlates with the fact that a future tense is used in the ice cream example, whereas a past tense is used in the rabbit example.

22. This can be favoured for the strong tendency of Balearic to use an epenthetic schwa sound [a] in initial positions for emphasis:

\[(i) \text{[a] que dius? emph what say.prs.2sg? ‘What are you saying?’}\]

\[(ii) \text{[a] on(t) ho has posat? emph where it have.prs.2sg put? ‘Where did you put it?’}\]

This initial sound is not related to DOM.
23. The other Catalan dialects distinguish between the first person singular *jo* (nominative) and *mi* (oblique).

24. For the role of right-dislocation in Spanish, see Sedano 2006.

25. For a detailed discussion of these aspects, see, among many others, Brugé and Brugger 1996; Dumitrescu, 1997, 1998; Torrego 1999; Leonetti, 2008.

References


Cinque, Guglielmo. [1983] 1997. ‘Topic’ constructions in some European languages and con-
nextedness. *Materials on left dislocation*, Elena Anagnostopoulou, Henk van Riemsdijk &
Company, Concepción. 1991. La extensión del artículo en el español medieval. *Romance Philol-
ogy* 44: 40–424.
Comrie, Bernard. 1979. Definite and animate direct objects: A natural class. *Lingüística Silesiana*
3: 13–21.
versity of Chicago Press.
Dalrymple, Mary & Irina Nikolaeva. 2007. Topicality and nonsubject marking: Agreement,
casemarking, and grammatical function. Ms, Oxford University.
Dumitrescu, Domnita. 1997. El parámetro discursivo en la expresión del objeto directo lexical:
español madrileño. *Actas del XII Congreso de la Asociación Internacional de Hispanistas,
Endriss, Cornelia & Stefan Hinterwimmer. 2007. Direct and indirect aboutness topics. *The No-
tions of Information Structure*, Caroline Féry, Gisbert Fanseelow & Manfred Krifka (eds.),
83–96 [Interdisciplinary Studies on Information Structure 6].
Engdhal, Elisabeth & Enric Vallduví. 1996: Information packaging and grammar architecture: A
constraint-based approach. *Integrating information structure into constraint-based and
dam: ILLC.
Escandell-Vidal, Victoria. 2007a. Topics from Ibiza: Differential object marking and clitic dis-
location. *Definiteness, specificity and animacy in Ibero-Romance languages*, Georg Kaiser &
Manuel Leonetti (eds.), 23–43. Konstanz: Universität Konstanz [Arbeitspapiere Fachbereich
Sprachwissenschaft 122].
Escandell-Vidal, Victoria. 2007b. Acusatiu preposicional i dislocació amb clitic. *Caplletra* 42:
185–218.
Farkas, Donka F. & Klaus von Heusinger. 2003. Stability of reference and object marking in Ro-
manian. Workshop on direct reference and specificity, ESSLII, Vienna.
Frascarelli, Mara & Roland Hinterhölzl. 2006. Types of topics in German and Italian. *On inform-
ation structures, meaning and form*, Kerstin Schwabe & Susanne Winkler (eds.), 87–116
Amsterdam: John Benjamins [Linguitstik Aktuell-Linguistics Today 100].
sterdam: John Benjamins [Typological Studies in Language 3].
Grosz Barbara & Yael Ziv. 1998. Centering, global focus, and right dislocation. Centering in
University Press.
Haspelmath, Martin. 2005. Universals of differential case marking. *Explaining Syntactic Univer-
sals*, a three week course taught at the LSA Institute at MIT (LSA 206), July 17-August 5,
2005.
The limits of syntactic variation, Theresa Biberauer (ed.), Amsterdam: John Benjamins [Lin-
guitstik Aktuell-Linguistics Today 132].
von Heusinger, Klaus. 2008. Differential object marking and the lexical semantics of verbs in Spanish, Probus 20: 1–33
Differential object marking and topicality. The case of Balearic


All rights reserved


**Data sources (Web addresses valid as of March 9, 2008)**


*COD*: *Corpus Oral Dialectal del català contemporani*, Universitat de Barcelona (research project HUM 2004/01504, Maria Rosa Lloret Romañach et al; http://www.ub.edu/lincat/).


mado-llucia: http://mado-llucia.balearweb.net/post/19537

Author’s address:

Victoria Escandell-Vidal
Departamento de Lengua Española y Lingüística General
UNED
Senda del Rey, 7
28040 Madrid (Spain)

vicky@flog.uned.es