Variable coding and object alignment in Spanish:
A corpus-based approach

José M. García-Miguel
Universidade de Vigo

Abstract

This article discusses three variable coding properties of Spanish objects: flagging (*a*-marking vs. ø-marking), indexing (clitic doubling vs. no doubling), and clitic case form (accusative *lo* vs. dative *le*). These properties are essential for the formal identification of grammatical relations. They are triggered by similar parameters that partly overlap and partly show distinct distributions, yet they also challenge the boundaries between direct objects [DO] and indirect objects [IO] and raise the question whether the typological alignment of Spanish (di)transitive clauses is indirective or secundative. The study draws on quantitative and qualitative corpus data on formal, semantic, and discourse properties of core participants in Spanish clauses, relating these properties to the distribution of variable coding. It is concluded that *a*-marking, clitic doubling, and *leísmo* are less frequently employed than unmarked objects, no doubling, and accusative case for clitics, that Spanish DO and IO must be taken as extreme points of a more general Object syntactic function, and that, in general, all variable object coding follows an indirective alignment type. Consequently, animate and topical objects are considered as formally and functionally marked atypical objects both in monotransitive and ditransitive clauses.

Keywords: object, grammatical relations, Differential Object Marking, case, argument indexing, ditransitive alignment, Spanish language

Author's address :
Universidade de Vigo, Facultade de Filoloxía e Tradución, Praza das Cantigas s/n, Campus Universitario, E-36310 Vigo, Spain, **e-mail**: gallego@uvigo.es

1. Introduction

1.1 Grammatical relations: Setting the scene

In this paper, I will address the issue of the coding properties of central participants in Spanish clauses. In this language, the identification of a Subject grammatical relation is relatively clear, but the "object" zone is more problematic. Traditionally, two grammatical relations – Direct Object [DO] and Indirect Object [IO] – have been distinguished, but the boundary between these is not clear. My aim is to revise some of the basic criteria that allow us to identify syntactic functions,[1] to examine how those criteria must be applied to Spanish, and to draw conclusions not only about Spanish grammar but also about the more general nature of grammatical relations. More specifically, this paper will examine the nature of Object grammatical relations by considering not only the distribution of their coding devices over monotransitive and ditransitive clauses, but also the text frequency of their main semantic and syntactic realizations. Drawing on corpus-based data, this study will shed light on three phenomena of variable coding of the object in Spanish (*a*/ø, clitic/ø, *lo*/*le*), as well as on core participants in Spanish and on the nature of grammatical relations. More generally, the data to be presented below are relevant to discussions of objecthood and markedness.

In Spanish, the coding property that defines the Subject [Subj] syntactic function is agreement with the verb in person and number, both in transitive and intransitive clauses. A Subject can be instantiated by a noun phrase preceding the verb,[2] as in (1), or following the verb. As Spanish is a so-called "pro-drop" language, the Subject is instantiated in many cases simply by the person and number verb index. The second participant in transitive clauses, the Object [Obj], is usually instantiated by a noun phrase, usually in post-verbal position. In some circumstances to be described below, this noun phrase may be preceded by the preposition *a* (see (1b)).

(1) a.   *Juan encontró   sus llaves.*
          Juan find.PFV.3SG his keys

          'Juan found his keys.'

     b.   *Juan encontró   a   sus amigas.*
          Juan meet.PFV.3SG   to his friends.F

          'Juan met his friends.'

Objects may also be indexed by a pronominal clitic (*lo, la, me,…*) variable for person, number, and gender. Such a clitic may occur alone as expression of the object, as in (2a), or it may co-occur in the same clause with a correferential noun phrase or with a coreferential independent personal pronoun (*él, ella, …*), as in (2b), in what is called "object duplication" or "object clitic doubling" and it represents an instance of object agreement (García-Miguel 1991) or, better, argument indexing (Haspelmath 2013).

---

[1] In this paper, I will use the terms "grammatical relation" and "syntactic function" interchangeably.

[2] Noun phrases in this paper have a noun as head, not a pronoun.

(2) a. *L-a-s   encontró*.
      3-ACC.F-PL meet/find.PFV.3SG
      'S/he met them.'/ 'S/he found them.

  b. *L-a   encontró   a ella/María*.
      3-ACC.F[SG] meet.PFV.3SG to her /Mary
      'S/he met her/Mary.'

I will consider this possibility of pronominal indexation as the defining formal property of objects in Spanish. Pronominal clitics come in two series (cases) in the third person: accusative case (*lo, la, los, las*) and dative case (*le, les*). Case may be used as a formal criterion to differentiate the two syntactic functions Direct Object [DO] and Indirect Object [IO], but there are many problems with such a differentiation, some of which will be dealt with in the following pages.

Both Subject and Objects can be indexed within the verb group (lexical verb, auxiliaries, and adverbal clitics), and this property allows them to be considered core or central participants (García-Miguel 1995: 41–46). Complements such as those in (3), on the other hand, are non-core oblique arguments,[3] and the corresponding clauses are considered intransitive. The intransitive verbs in (3) govern complements with the preposition *a* (functioning as a directional marker) or with other prepositions.

 (3) a. *Juan fue   a Leipzig*.
      Juan go.PFV.3SG to Leipzig
      'Juan went to Leipzig.'

  b. *Juan pensaba   en su familia*.
      Juan think.IPFV.3SG in his family
      'Juan was thinking about his family.'

With this short characterization of Spanish syntactic functions, which essentially follows reference grammars (Alarcos Llorach 1994: chap. 21–24; RAE & AALE 2009: chap. 33–36), I am assuming that in order to say that a given grammatical relation exists in a given language, the claim must be justified both language-internally and cross-linguistically (Comrie 1989: 66; Andrews 1985: 71–77). The possibility of a universal definition of grammatical relations has been challenged within the functional-typological linguistics tradition and grammatical relations are considered not only language-specific but also construction-specific (Dryer 1997; Croft 2001; Bickel 2011). In what follows, I will assume that constructions are the basic units of grammar and that syntactic functions must be characterized in relation to the constructions in which they appear. Elements belonging to different constructions in the same language should be said to share the same syntactic function to the extent that they share formal encoding mechanisms (order, indexing, case, etc.). For cross-linguistic comparison, I will use the labels S, A, P, T, and R as comparative concepts in characterizing grammatical relations (Haspelmath 2011). S is, in any language, the sole argument in the major monoactant (intransitive) construction. A and P are, in any language, the arguments of the major

---

[3] Note that oblique elements may either be arguments or adjuncts, but the labels "core" or "central" are reserved here for Subject and Objects.

biactant (monotransitive) construction. And A, T, and R are, in any language, the arguments of the major triactant (ditransitive) construction. Subsets of arguments that receive the same treatment by a specific construction in a language (e.g. that receive the same case marking) may be called "alignment types" (see Bickel 2011: 203): the subset {S, A} (as opposed to {P}) defines the nominative–accusative alignment type, and the subset {S, P} (as opposed to {A}) defines the ergative–absolutive alignment type.

In comparative terms, the Spanish Subject can be defined as the subset {S, A} and the alignment as nominative–accusative. Clauses (1) and (2) are examples of a transitive construction <A + P>. As for the clauses in (3), they are said to represent the intransitive pattern <S + X>, and not the transitive pattern <A + P>. Thus, I follow the usage of A and P by Comrie (1989) and Haspelmath (2011), and not that of Bickel and Nichols, who generalize A and P [O in their terminology] to all two-argument clauses (Bickel and Nichols 2009; Bickel 2011; Nichols 2011: 460–466).

We may also ask why the [Subj–V–Obj] construction, and not any other biactant construction, has been chosen to represent the transitive construction. There are two main criteria for identifying a particular construction as the major biactant construction in a language (Witzlack-Makarevich 2010:109–111). One criterion, proposed by Comrie (1989: 111) and Lazard (2002: 152) is qualitative: it is the construction used with verbs expressing prototypical actions, such as "kill" and "break". The second criterion is quantitative: it is the construction having greater productivity, higher token frequency, and/or higher type frequency (this last criterion is preferred by Witzlack-Makarevich). In Spanish, as in many other languages, both criteria converge on the same pattern: [Subj–V–Obj] is the biactant pattern with higher type and token frequency and is also the pattern of *matar* 'kill'¸ *romper* 'break', and many other common effective action verbs.

1.2 The problem: Variable coding of Spanish objects

The definition of grammatical relations in Spanish is complicated by the phenomena of variable marking, also known as differential object marking (DOM) (Bossong 1998) or actance variation (Lazard 1984): some objects, even in [V–Obj] combinations involving the same lexical verb and the same lexical noun, may show variable coding properties. The three main variable marking phenomena that we will consider here concern the use of the preposition *a* or ø, the use of object clitic doubling or not, and the dative vs. accusative case of object clitics.

(i) Some, but not all, Objects may be marked by the *preposition a* or ø, as in (4) (see the examples in (1)); in some contexts the choice between *a* and ø entails a semantic difference. In general terms, the preposition *a* is used if the referent is both human and specific.

(4) *Encontré   (a) un amigo.*
     met.PFV.1SG (to) a friend

     'I met a friend.'

(ii) Some, but not all, Objects may be clitic-*doubled*, i.e. they may be instantiated by a pronominal clitic and by a co-nominal. In general, the presence of a full nominal excludes the possibility of a pronominal clitic, but depending on the dialect and

the grammatical context an object clitic and a full nominal may coexist in the same clause, as in (5).

(5) *(**Lo**)   encontré   a Pepe.*
     3SG.M.ACC met.PFV.1SG to Pepe
     'I met Pepe.'

(iii) Some Objects alternate between the *accusative* and the *dative* pronominal clitic, without any clear difference in meaning. This phenomenon is traditionally known as "leísmo" and is usually interpreted as an anomalous use of the dative clitic where the accusative is expected.[4]

(6) *Lo / **le**   encontré.*
     3SG.M.**ACC** / 3SG.**DAT** met.PFV.1SG
     'I met him.'

In general, variation in object marking can be described in terms of *split* alternations and *fluid alternations*. Applied to case marking, split alternations display "alternation[s] of lexical case associated with different verbal lexemes", whereas in *fluid alternations* "the same verb takes alternative case frames depending on transitivity parameters" (Malchukov and de Swart 2009: 341). In this sense, the Spanish alternations exemplified in (4)–(6) are clearly of the fluid type. A stricter definition of fluid alternation, however, requires that "the *same* noun phrase *in the same linguistic context* can alternatively take both case markers" (de Hoop & Malchukov 2007: 1638; original emphasis). This stricter definition holds of examples (4)–(6), but not of all Spanish objects: for example, personal pronouns hardly ever show alternation: they (almost) always require *a*-marking and are (almost) always clitic-doubled. It appears, then, that variation in object marking in Spanish is of a mixed type, in that it represents mostly fluid alternations and partly split alternations. With regard to Spanish *a*-marking, for instance, Klein and de Swart (2011: 8), hold that it is characterized by split alternation based on animacy and definiteness, and by fluid alternation based on specificity. Actually, it is difficult to find grammatical contexts in Spanish where one of the alternating object coding forms is strictly prohibited. Rather, what we observe is a different frequency of each alternating form depending on the inherent properties of the object and, to a lesser extent, depending also on the verb and other properties of the clause as a whole.

One way of examining the coding choice in a fluid alternation is to search for the conceptual or functional features that determine the choice. A different approach involves observing the contexts of use and the frequency distribution of the alternating forms. The analysis carried out in the following pages – informed by this second approach – looks at the statistical distribution of variable marking and investigates the factors underlying this distribution. My overview of object variation in Spanish, then, aims to complement, and not contradict, other more detailed studies on the prepositional object (Laca 1987, 2006; Pensado 1995; Torrego Salcedo 1999; Delbecque 2001;

---

[4] The inverse phenomenon also exists: the usage of the accusative clitic ("*loísmo*" and "*laísmo*") where the dative is expected.

Leonetti 2008, to name just a few) and on object clitics (García 1975; García and Otheguy 1977; Fernández Ordóñez 1993, 1999; Klein-Andreu 2000; Flores 2002; Belloro 2007, among many others). In considering the frequency distribution of the features of core grammatical relations and variable marking, I will try to follow a functional and constructional approach, identifying semantic and discourse motivations for syntactic structures. This corpus-based quantitative approach is related to the belief that grammar emerges from discourse through the repetition of forms (Bybee 2006) and that frequency is a crucial factor in the shaping of linguistic structure.

　　Although variable object coding does not, in principle, depend on verb class and is possible with all transitive verbs, a prominent problem for Spanish grammar is the existence of a minor verb class whose object is always flagged by the preposition *a* and indexed by a dative clitic. This class includes mainly experiential predicates, such as *gustar* 'to like' (Vázquez Rozas 2006), which systematically take a dative Object Experiencer, usually in initial preverbal position if it is a full noun phrase, and a Subject Stimulus, usually in postverbal position:

(7)　*A María le　　　gusta-n　las películas del　　Oeste*.
　　to Mary　3SG.DAT　like-3PL　the movies　　of.DEF　Western
　　'María likes Westerns.'

It appears, then, that we could recognize a major verb class of transitive verbs like *encontrar* 'find, meet' and *matar* 'kill', allowing alternating object coding, and a minor *gustar*-type verb class. The object of the first class would be a Direct Object [DO], and the object of *gustar*-type verbs would be an Indirect Object [IO]. In comparative terms, only the major class would correspond to the <A + P> pattern. But there are some problems with this approach. First, the verb classes are not so clear cut: in addition to *gustar*-type verbs governing an IO, there are other intermediate verbs that prefer *a*-marking or dative case in most but not all contexts, and yet others that do so more rarely. What we seem to find, then, is in fact a continuum. Second, as we have seen, the three DOM phenomena mentioned above have different distributions, and in some contexts there can be fluidity on one dimension and rigidity on another. Thus, the presence of object alternation itself cannot be used as a general test for DO. Third, even if there were disjoint verb classes, in each DOM dimension one of the alternating forms (namely, preposition *a*, clitic doubling, and dative case) also represents the coding of IO. Now, according to the isomorphism hypothesis, "recurrent identity of form between different grammatical categories will always reflect some perceived similarity in communicative function" (Haiman 1985: 19), so the formal and semantic similarities between some DO and IO must in any case be explained. For these reasons, and given that the boundaries between canonical transitive clauses and two-participant constructions of the *gustar*-type are not clear-cut, most parts of the analysis in Sections 3 and 4 below will concern all constructions with two core participants Subject and Object (either DO or IO).

　　As we are jointly treating all biactant verbs taking an object and all object forms, one might ask which of the alternating forms is more representative of the major biactant construction [S–V–Obj]. The semantic qualitative criteria are not fully clear: Lazard argues that, in general, differential object indexing and marking correlate with the degree of individuation of the object, and "the major biactant construction is clearly the construction with marked object" (Lazard 2002: 157); for Spanish, this means

constructions with *a* and clitic doubling. Comrie, on the other hand, claims that "the most natural kind of transitive construction is one where the A is high in animacy and definiteness, and the P is lower in animacy and definiteness" (Comrie 1989: 128); this points to constructions without *a* and without clitic doubling. The status of case (*le* vs *lo*) with regard to individuation and transitivity is less clear. Indeed, we need additional criteria to make this decision as to which object form best characterizes the transitive construction, and in the following sections I will provide quantitative corpus-based data that seem to support Comrie's view. The decision also affects the problem of ditransitive alignment, which will be formulated in the next section.

2. Object alignment: Direct Object or Primary Object?

In Spanish, coding variation can be observed in the second actant P of the major biactant verb class, which comprises verbs like *matar* 'kill', *romper* 'break', *encontrar* 'find', and *ver* 'see'. A noteworthy consequence of this variable marking is that a number of P participants share some coding with R,[5] the recipient of ditransitive constructions canonically instantiated by verbs meaning 'give' and other transfer verbs.

In typological studies, the usual approach to alignment types − accusative, ergative, etc. − has been extended from descriptions of transitive constructions to the description of ditransitive constructions (Dryer 1986; Haspelmath 2005a; Malchukov, Haspelmath and Comrie 2010). Using T and R as the labels for, respectively, the "Theme" or transferred object and the "Recipient" of prototypical ditransitive constructions, three basic types can be distinguished: (i) *indirective alignment* (T=P≠R), opposing a direct object {P, T} to an indirect object {R}, as in German; (ii) *secundative alignment* (R=P≠T), opposing a primary object {P, R} to a secondary object {T}, as in West Greenlandic; and (iii) *neutral alignment* (P=T=R), with two objects in ditransitive constructions, as in English:
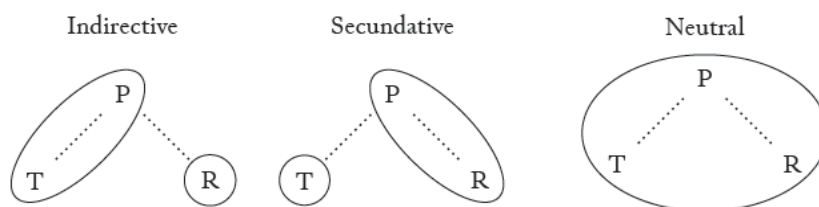

Fig. 1: Ditransitive alignment types (Malchukov et al. 2010: 5)

In Spanish, the R argument (typically Recipient) of prototypical ditransitive clauses takes the preposition *a* and is indexed by a dative clitic; the T argument ("Theme": the transferred object with verbs of transfer) is typically realized as a ø-marked noun phrase and may be indexed by and accusative clitic (although doubling is rare):

---

[5] This is a relevant difference with respect to DOM in other languages, such as Finnish, where the object alternates between accusative and partitive cases, and R is marked by a different case, namely adlative.

(8) *María le   dio   un regalo a   Juan.*
     Maria 3SG.DAT give.PST.3SG a   present to John
     'Mary gave a present to John.'

In other words, broadly speaking, because of DOM, R shares grammatical properties with some Ps, whereas T shares grammatical properties with some other Ps. Such facts have not been an obstacle to recognizing the syntactic functions Direct Object [DO] – for the set {P, T} – and Indirect Object [IO] – mainly for {R} – in the grammars of Spanish (RAE & AALE 2009: chap. 34–35). Accordingly, typological studies classify Spanish as an indirective language type, even if they recognize that DOM is a potential source of problems (see, for example, Haspelmath 2005b; Malchukov et aL 2010: 7). Although he also adheres to this view, Comrie (2012: 19) considers the possibility that Spanish represents a fourth ditransitive alignment type, "namely split-P alignment, in which P sometimes (when low in animacy/definiteness) aligns with T, sometimes (when high in animacy/definiteness) with R". He also notes that sentences similar to (8) "present an initial instance of the relevance of primacy […] This is in effect a variety of indexing of the R on the verb.[…] We can thus say that in Spanish, R has primacy over P and T with respect to indexing on the verb by means of a clitic pronoun" (Comrie 2012: 20).

   As an alternative to the view that Spanish follows the indirective alignment pattern, some linguists have suggested the possibility that perhaps IO could be considered the "true" object in Spanish (see Roegiest 1990: 248), with Spanish being or becoming a special kind of primary object language (Company 2001). On that view, the primary object would be coded by *a*-marking, clitic doubling, and/or dative case, both in monotransitive and ditransitive clauses. This would also likely imply that ø-marked, not doubled, and/or accusative case objects are secondary objects. Note, however, that passivization clearly follows indirective alignment: in general, Spanish allows P-passivization and T-passivization, but R-passivization is not possible. If coding did follow, even partially, the secundative type, this would represent a new typological problem: "Passivization can follow a secundative pattern even if coding is neutral and a neutral pattern even if the coding is indirective. What is unattested is a language with secundative coding but strictly indirective passivization. Thus, R-passivization is generally preferred over T-passivization" (Malchukov et al. 2010: 30) Does Spanish falsify the claim that secundative alignment is incompatible with indirective passivization? I don't think so.

   Company (2001) argues that there are some features and ongoing changes in Spanish that point to it being a special kind of primary-object language. Her argumentation is partly based on the lexical and pragmatic meanings of arguments, and partly on diachronical changes involving the grammatical forms common to IOs and a number of DOs, namely preposition *a,* clitic doubling, and dative case. According to Company, the set of gramatical changes used as evidence comprises the following seven facts:

   (1) 'Anomalous' Dat-marking usurping Acc–DO; (2) Generalized Dat–DO; (3) Consistent marking of Dat with the preposition *a*; (4) Dat–IO duplication; (5) Depronominalization of Dat clitics; (6) Progressive invasion of Dats as the causee of causative constructions; (7) Frequent order V–Dat–Acc. (Company 2001:12)

These seven "apparently unconnected changes" yield "a global shift by which Dats stand as main objects in diverse grammatical areas, displacing in many cases the Acc from the DO position" (Company 2001: 30). These changes are directly related to the coding phenomena addressed in this paper: the use of preposition *a* as an object marker, participant indexation ("object duplication"), and the dative–accusative case alternation.

Before we continue with variable object marking, let us consider Company's points (1) and (5). By "Anomalous Dat-marking usurping Acc–DO" she refers to the fact that in some dialects of Spanish the gender and number marking of R (F.PL in (9)) is attached to the accusative clitic, not the dative. In her opinion, this new cliticization "behaves as a lexicalized, single, basically unanalyzable form: *selos, selas, seles*" (Company 2001: 15).

(9) *Si ellas   me quieren comprar el   caballo,*
     if they.F me want   buy   the horse

     yo se   l-a-s   vende-ré.
     I   them(DAT) 3-ACC.F-PL sell-FUT

     'If they want to buy the horse from me, I will sell it to them.'

The "depronominalization of Dat clitics" is exemplified by the lack of number agreement in examples like (10):

(10) *Pónga=le   las carpetas azules a   los sillon-es*.
      put=3SG.DAT the doilies   blues to the armchair-PL
      'Put the blue doilies on the armchairs.'

It seems that these two phenomena show contradictory behavior: on the one hand, anomalous Dat-marking in (9) exhibits a tendency to encode agreement in gender and number with datives, even in contexts where previously there was no formal possibility to do so (attaching the plural marker to the accusative instead of the invariable *se*); on the other hand, (10) shows a tendency not to mark plural agreement when this is formally possible. Of course, these two contradictory tendencies have a clear functional motivation: the singular *le* in (10) anticipates a plural postverbal object (*sillones*) in the same clause; the pluralized *selas* is used to recover a plural referent (*ellas*) mentioned in a previous subordinate clause. These phenomena are important for the analysis of clitics, which are grammaticalized referent tracking devices, but do not, I believe, affect the alignment of ditransitive clauses.

Company, however, also presents a number of tendencies which are significant: (i) that "animate, individuated, active masculine patients, barely affected by the action of the verb or conceptualized as respectable or important, are preferred to be marked with an innovative Dat-DO *le*" (2001: 17), (ii) that "the use of '*a*' marking for human Accs makes them closer to Dats, eroding the differences between the two objects" (2001: 19), and (iii) that "the steady increase that Dat doubling has had in the evolution of Spanish reveals that the marking of Dat-object agreement has become almost obligatory" (2001: 23). The full set of changes (1)–(7) quoted above is taken as a symptom of two "complementary tendencies: (a) a tendency of Dat case to **displace the Acc** case in DO function, and (b) a tendency to modify, by different means, the Dat case to **reinforce Dat-marking** or, in other words, to strengthen the role of Dats as objects

and not as obliques" (Company 2001:12–13; original emphasis). This final point is important in that part of Company's argumentation aims to show that Dats in Spanish are not obliques. Although I concur, the problem here is in the definition of "core" and "oblique". Above, when commenting on examples (1)–(2), I used a formal criterion: Subject and Objects [both DO and IO] are core or central participants because they share the possibility of being indexed within the verb group; all other arguments and adjuncts, then, are "oblique". From a semantic and discourse perspective, Subject and Objects, as core participants, are also the more prominent participants (García-Miguel 1995: 41–46; Vázquez Rozas 1995), or, in Goldberg's terms, the constructionally profiled arguments (Goldberg 1995: 48–49). It is more difficult to give a cross-linguistically valid definition of core and oblique arguments (but see Nichols 1983; Andrews 1985: 81; Thompson 1997). The main criteria for coreness are expression through noun phrases (vs. adpositional phrases) and/or indexing within the verb group. More broadly, the two participants of the major biactant construction can be considered core arguments, and other arguments formally similar to these two are also core arguments. In ditransitive constructions, primary objects of secundative alignment types, direct objects of indirective alignments, and both objects of neutral alignments would be, by definition, core participants. T in secundative alignments and R in indirective alignments may be considered core or oblique, depending on the specific encoding used in particular languages. In many languages, R takes the same encoding as benefactives, goals, possessors, and other non-core participants.[6] In Spanish, Rs, that is IOs, share many formal and semantic properties with the second participant of the major biactant constructions. IOs, then, are core participants in Spanish, and other prepositional complements of triactant clauses are not (see also García-Miguel 1999).

The controversial elements in Company's thesis are not related to the status of IOs as core participants, but related to the alignment of ditransitive clauses in Spanish, that is (i) whether arguments marked by *a* and cross-indexed by the dative clitic *le* are becoming the "true" (direct/primary) Object, (ii) whether "*leismo*, Dat for Acc, may be understood as a certain kind of Dat promotion" (Company 2001: 17), and (iii) whether "the grammatical behavior in this area [*a*-marking] appears almost like a PO language: there is only one object case marking, which falls on the patient of monotransitives and on the recipient ditransitives" (2001: 21). In what follows, we will examine the extent to which it is true that "Dat-case displaces the Acc case in DO function" (2001: 12), or in more neutral terms, to what extent ditransitive R-arguments share coding (and behavior) properties with monotransitive P-arguments and whether there is a preferred or more representative alignment of Spanish (di)transitive clauses. Answers will be provided in Section 5.

3. The corpus, the database, and some variable properties of core participants

The data used for this study are from the ADESSE database in the state it was on May 25, 2013.[7] The database contains syntactic and semantic analyses of the almost 160,000 clauses that make up the texts of the ARTHUS corpus ("Archivo de Textos Hispánicos

---

[6] According to Newman (1996), besides being a (primary) object, a recipient can be integrated into a "give" clause, mostly marked as a dative, a goal, a locative, a benefactive, or a possessor
[7] The ADESSE database, which can be browsed at http://adesse.uvigo.es/, is a revised and semantically extended version of BDS. The original BDS database is partly accessible at http://www.bds.usc.es/.

de la Universidad de Santiago"). This is a set of 34 texts in Spanish published between 1981 and 1991, with a total size of about 1.5 million words. The archive comprises mainly essays (18%), theater (15%), and narrative literary texts (38%), together with some journalistic texts (12%), and a sample of spoken language (19%) from Madrid, Seville, and Buenos Aires. The origin of the texts is mainly Peninsular Spanish (78%), with the remainder being from Latin America.[8] Syntactic coding of this corpus has resulted in a syntactic database of contemporary Spanish (BDS), which thus contains syntactic analyses of the almost 160,000 clauses that make up the ARTHUS texts. This database was integrated in the ADESSE database and extended with additional semantic features .

Each clause of the corpus was coded for several syntactic and semantic features of the clause (verb form, verb sense, verb semantic class, polarity, modality, voice) and for the main syntactic and semantic features of its inherent arguments: syntactic function (Subj, DO, IO, Oblique Complement, Locative, Manner, Oblique Agent, Attribute), argument indexing (subject agreement and object clitics, if any), syntactic category (Noun Phrase, Pronoun, Infinitive, Finite Clause, ...), preposition (which one, if any), animacy, definiteness, number, semantic (micro-)role, lexical head, order (position with respect to the verb). Most argument features that are encoded in the database are fairly accessible to observation, and clearly linked to semantic and discourse properties. The following variable features are considered relevant for this study and will be analyzed in the following pages:

–   *Animacy*, i.e. animate vs. inanimate, which is an index of inherent agency potential (Silverstein 1976)
–   *Instantiation*, i.e. either "full" instantiation by a syntactic category (NP, Pro, or clause) or argument-indexing only, or both, as in (1) and (2) for Object. Instantiation may correlate with referent accessibility in discourse (Givón 1983; Ariel 1990; Chafe 1994): more accessible referents are retrieved with lighter forms, such as object clitics)
–   *Definiteness*, i.e. definite vs. non definite (indefinite or zero determiner), which is an index of referentiality or individuation, but it is also an index of relative accessibility.
–   *Order* of "full"constituents (NP, Pro, or clause) in relation to the verb, i.e. preverbal or postverbal. Initial position in the clause is a clear index of thematicity (Halliday 2004). Examples of SVO order are given in (1), whereas (11) is an example of OVS order.

(11)   *A los viejos nos acompaña   nuestra historia.* (SON: 220)
       to the elder 1PL accompanies our   history
       'We the elders are accompanied by our history.'

Neither of the values associated with each feature (e.g. the values "animate" and "inanimate" of the feature "Animacy") is obligatory with any grammatical relation. Rather, the preferred value for each feature is determined by the semantic and discourse

---

[8] The complete list of references for the texts comprising the corpus can be found in the Appendix, or at http://adesse.uvigo.es/data/corpus.php.

functions of syntactic structures. While a grammatical relation may, in principle, be characterized by any combination of feature values, not any combination is in principle possible, and this is reflected in frequency. Let us first consider the set of two core participant clauses, that is, clauses with a Subject and an Object. The distribution of feature values in the corpus is set out in Table 1. As the boundaries between canonical transitive clauses (with DO) and IO two-participant constructions of the *gustar*-type exemplified in (7) are not clear (see Section 1.2), this table includes all object types (DO, more accusative-like, or IO, more dative-like). As a matter of fact, the inclusion of IO raises the percentages of animate, definite, preverbal, and non-lexical objects.

Table 1: Properties of participants in Subj–V–Obj [DO/IO] (+ X)
         [N = 77,212 clauses] (ADESSE)

	Subj	**Obj [DO/IO]**
Animacy		
Animate	80.50%	27.14%
Inanimate	19.50%	72.86%
Instantiations		
Indexing only	63.80%	25.90%
"Full" instantiation (NP, Pro, or clause)	36.20%	74.10%
Definiteness (if NP or Pro)		
Definite	90.00%	66.33%
Indefinite	10.00%	33.67%
Order (if NP, Pro, or clause)		
Preverbal	73.67%	3.84%
Postverbal	26.33%	96.16%

Table 1 shows that subjects are animate, definite, preverbal (i.e. thematic or topical), or "reduced" to verbal indexing more frequently than objects. These Spanish corpus data reflect a universal tendency related to the main asymmetries contrasting participants in transitive clauses (see (12)): "subjects tend to be definite, animate, and topic (thematic); while direct objects tend to be indefinite, inanimate, and rhematic" (Comrie 1979: 19). Note, however, that, although objects are not indefinite to a higher degree than definite ones, they do – at least in this corpus of Spanish – present a lower rate of definiteness than subjects.

(12)   A   P
                Human   Non-Human
                Definite   Less definite
                Highly accessible   Less accessible
                Theme   (Part of) Rheme
                [Agent]   [Patient]

     Subjects of intransitives (S) always show significantly lower rates in animacy (agency potential), agreement-only instantiation (accessibility), and preverbal position (thematicity/topicality) than subjects of transitive clauses (see Table 2). But in every

item, the percentages are still higher than those offered by objects in Table 1, and closer to that of transitive subjects.

Table 2: Subject in Subj–V (+X) [N: 66,691 clauses] (ADESSE)

	S (Subj)
Animate (vs. inanimate)	65.66%
Indexing only (vs. "full" instantiation)	51.02 %
Definite (vs. non definite)	87.90 %
Preverbal (vs. postverbal)	60.10 %

These Spanish data confirm that the subject of intransitive constructions neutralizes the polarization between A and P in terms of animacy (agency potential) and preverbal position (thematicity/topicality), and that S is intermediate between the polar participants of transitive constructions. This is also a universal tendency that provides the basis for either accusative or ergative alignment types.[9]

In three-participant clauses (ditransitive constructions), R shows a frequency distribution of feature values very similar to that of subject – obviously with the exception of constituent order, as lexical Rs are postverbal. T tends to show the opposite values, with a frequency of the feature values "animate" and "preverbal" (indicative of thematicity) which is significantly lower in ditransitive T-arguments than in the object of two-participant clauses.

Table 3: Participants in the ditransitive construction [N = 8,445 clauses] (ADESSE)

	A (Subj)	T (DO)	R (IO)
Animate (vs. inanimate)	84.18 %	2.25%	90.24 %
Indexing only (vs. "full" instantiation)	65.95 %	10.65%	74.14 %
Definite (vs. non definite)	90.06 %	53.57%	89.02 %
Preverbal (vs. postverbal)	74.50 %	2.40 %	9.50 %

These findings from Spanish, then, confirm the general principles that "the most frequent and therefore most expected ditransitive associations are animate/definite R and inanimate/indefinite T" (Haspelmath 2007: 83), with the added nuance that T is not predominantly indefinite, but only less frequently definite than A and R. The findings also confirm a high accessibility of R (feature value "Indexing only"), which in most cases is reduced to a clitic. The reduction to a verbal index is even higher for R than for A.

In sum, the distribution in the corpus of a set of basic features shows clear indications of inherent agency potential (feature value "animate") and inherent topicality/thematicity (feature value "preverbal") and accessibility (feature value "indexing only") of central participants. There is a polarization between A and P in transitive clauses, with much higher values for A in every feature; at the same time, the intransitive S-argument shows intermediate results, although closer to that of the A.

---

[9] The tendency to introduce new referents by means of lexical items in S or P slots, but not in A, is known as "preferred argument structure" and said to provide the discourse basis of ergativity (Du Bois 1987 and Du Bois 2003; Du Bois et al. 2003). I view the intermediate status of S rather as a good basis for either ergative or accusative systems.

Likewise, there is a polarization between R and T in ditransitive clauses, with much higher values for R, and with the transitive P-argument showing intermediate values, though closer to those of T.

(13) a. A > S > P
      b. R > P > T

This is the functional background against which variable coding of grammatical relations must be understood. In general, we expect from (13a) that S aligns sometimes with A and sometimes with P, and we expect from (13b) that P aligns sometimes with R and sometimes with T.

4. Object variation in two-participant clauses

In this section, I will present an overview of the distribution of the three phenomena of actance variation exemplified in (4)–(6), namely variable nominal marking (use of preposition *a* or ø), variable person indexing (object duplication by a pronominal clitic vs. no clitic indexing of the object), and variable case of the clitic (dative vs. accusative). All two-participant active clauses with Subject and Object (and optionally additional oblique elements) will be taken into account, regardless of whether the object has traditionally been considered as a Direct Object (as in *romperlo*) or an Indirect Object (as in *gustarle*).

4.1 Preposition *a* and object doubling in two-participant clauses

I will deal with *a*-marking and person indexing (clitic doubling or object agreement) together because they display many common properties, they are triggered by similar factors, and have similar effects (Leonetti 2008). The factors governing the use of *a*-marking and clitic doubling involve animacy and definiteness, the lexical meaning of the verb, and discourse-pragmatic features. The most relevant factors have to do with animacy and definiteness. Since the annotation of our database includes information on syntactic category (personal pronoun, NP, relative pronoun, clause), on definiteness (definite, indefinite), and on animacy (animate, inanimate), these features can be combined and positioned on a scale which, broadly speaking, follows the animacy and referential hierarchy (Silverstein 1976; Lazard 1984; Bossong 1998; Aissen 2003). As Table 4 shows, the descending order of frequencies of the preposition *a* and of doubling correlates with the descending degree of animacy/definiteness:

Table 4: Preposition *a* and doubling with non-subject in 2-participant clauses; full (not just clitic) Object [DO/IO]

	Total	% *a*	% doubling
Personal pronoun	779	99.7%	99.6%
NP animate definite	3830	90.4%	15.1%
NP animate indefinite	1374	47.1%	6.4%
Relative Pro animate[10]	482	42.7%	12.2%
NP inanimate definite	20999	2.6%	2.8%
Relative Pro inanimate	4923	0.5%	1.0%
NP non animate indefinite	14252	0.4%	0.3%
Clause	10747	0.0%	0.2%
TOTAL	57386	10.0%	3.8%

Both phenomena show a clear decreasing tendency along the scale, but the cut-off points are different in each case: the use of *a* is correlated primarily with animacy and definiteness, while doubling is triggered basically by personal independent pronouns. Neither phenomenon is totally obligatory at the higher levels of the scale and neither is fully excluded from the lower levels, except the preposition *a* with complement clauses. The set of relevant factors related to the hierarchy of animacy and definiteness is common to both phenomena:

(14)   Pronoun > NP
       NP > Clause
       Animate > Inanimate
       Definite > Indefinite

Given that object duplication is more restrictive than the use of *a*, one may be tempted to think that the doubling uses constitute a subset of the uses of *a*. This is known in some formal circles as *Kayne's generalization*: "An object NP may be doubled by a clitic only if the NP is preceded by a preposition" (reproduced from Leonetti 2008: 34). However, this generalization does not hold for preverbal inanimate objects, which may be doubled and are usually not *a*-marked.

(15)   *El   dinero   lo   puso   Arturo, naturalmente*. (AYE: 067)
        the money   3SG.ACC.M put.PFV.3SG Arthur   naturally

       'The money was contributed by Arthur, of course.' (lit. 'The money, Arthur contributed it, of course.')

The generalization is closer to being true of animate objects, where there is a clear statistical association between *a*-marking and doubling, but there is no obligatory implication in either direction. For instance, there are several examples in the corpus of preverbal animate NPs which are doubled, but not *a*-marked.

---

[10] Relative pronouns are a special case: relative *que* does not allow *a*-marking, whereas *quien* and (*el*) *cual* require it. The figures for relative pronouns in fact correspond to the use of those forms for animate and inanimate objects.

(16)   *Mis amigos de la universidad los   adoro todavía.* [SEV: 255]
       my friends of the university 3PL.ACC.M adore yet

       'I still adore my university friends.'

Table 5: *a*-marking and doubling of NPs objects with animate referents

	Doubled	**Not doubled**	**Total**
a-marked	642	3470	4112
ø-marked	25	1067	1092
Total	667	4537	5204

$\chi^2 = 137.073$, p-value < 0.001

As was pointed out above, the use of *a* is the norm with definite and animate objects. As for indefinite animate objects, the use of *a* seems to be optional, but it is clearly related to a specific or referential interpretation of the NP. Together with some other parameters, the set of factors conditioning the use of *a* are particular aspects of the individuation of the object (Kliffer 1984), a parameter of cardinal transitivity (Hopper and Thompson 1980: 253). *A*-marking signals that the object is more individuated and, in this sense, that the clause is more transitive.

In addition to the animacy hierarchy, there are a number of semantic and pragmatic factors governing the use of *a*-marking and clitic doubling. Kliffer (1984) notes that the use of *a*-marking seems to depend on verb kinesis. Von Heusinger (2008) shows that *a*-marking first occurred with action verbs like *matar* 'kill' and *herir* 'hurt', which select human objects, and then extended to other verbs like *ver* 'see', *hallar* 'find', *tomar* 'take', and *poner* 'put'. This diachronic path is argued to be related to (degree of) affectedness of the object (von Heusinger and Kaiser 2011). Our data on contemporary Spanish do not show significant differences between these verbs. Nor do we find clear evidence that *a*-marking is associated with agentivity or telicity, as is advocated by Torrego Salcedo (1999: 1784–1790). Laca (2006) observes that there are verbs that have always favored *a*-marking, even with inanimate objects: *llamar* 'call, name', relational verbs like *sustituir* 'substitute', *preceder* 'precede', and *seguir* 'follow', and others which select animate objects; but she thinks that "no es fácil hallar un denominador común a los lexemas verbales que favorecen el empleo de la marca" [it is not eeasy to find a common denominator among the verbal lexemes that favor the use of *a*-marking] (Laca 2006: 470). She also notes that other verbs are resistant to *a*-marking, in particular *tener* 'have' and existential *haber* 'there be', although the resistence of *tener* to *a*-marking is debatable (see Miles and Arciniegas 1983) .

The case of relational verbs is especially interesting, in that they also *a*-mark their inanimate object in many cases:

(17)   *El sujeto precede **a**-l objeto.* [LIN: 071]
       the subject precedes to-the object

       'Subject precedes object.'

According to Delbecque, the preposition *a* marks "the relationship between the S entity and O entity as 'bilateral', i.e., instead of having a simple unidirectional force-

dynamics going from the subject entity towards the DO entity, the relationship could just as well be presented the other way around" (Delbecque 2001: 87–88). García García (2007, 2014) provides a similar motivation for the use of *a* with inanimate objects of reversible or symmetrical predicates: *a*-marking is required when the object is equally agentive as or more agentive than the subject. As noted by Laca (1987: 291), the possibility of *a*-marking of inanimate objects with these stative relational verbs contradicts the interpretation of such marking as a signal of high transitivity (Hopper and Thompson 1980). On the other hand, it shows that DOM is not only related to inherent properties of the object but may involve properties of the entire predication, including the verb. In my opinion, the most relevant fact in accounting for variable object marking is that the subject is expected to outrank the object in properties related with agentivity and topicality (as can be deduced from the corpus data of Table 1), and *a*-marking is used where the subject–object asymmetry is less clear. This justifies the higher animacy of the *a*-marked objects per se, but also the marking of inanimate objects with some relational verbs.

These relational verbs, e.g. *preceder*, should be distinguished from those requiring the dative case in cliticization such as *gustar* or *ocurrir* (see Section 3.2 on clitic case below). Dative case has a strong association with animacy, and *gustar*-type verbs almost always select an animate object, and consequently also the preposition *a*. But some verbs of happening, like *ocurrir* 'occur' and *suceder* 'happen', admit an object that requires the preposition *a* and the dative case in cliticization even if it is inanimate as in (18).

(18)   *¿Qué demonio-s **le**   sucedía   **a**   su maldita memoria?* [MIR: 117]
        what devil-PL   3SG.DAT happened to his damned memory
        'What the hell was happening to his damn memory?'

Relational verbs, *gustar*-type verbs, and happening verbs show that *a*-marking in two-participant clauses is not solely associated with verbs expressing prototypical actions. Relational verbs show that *a*-marking is, at least in part, independent of animacy and independent of dative case.

Finally, apart from animacy and the lexical meaning of the verb, some discourse-pragmatic factors are relevant. Leonetti points out that "the contribution of *a* is the encoding of an instruction to process the object as an internal topic, that is, as a prominent and referentially autonomous argument" (Leonetti 2004: 94). The topicality of *a*-marked objects imposes, or favors, strong readings of the NP (specific or generic) and blocks semantic incorporation into the predicate, that is, the use of a NP as a property-denoting expression that modifies the predicate, as in *tener dos hijos* 'to have two children'. In a broader context, Dalrymple and Nikolaeva (2011) argue that "casemarking and agreement patterns in many languages with DOM distinguish topical objects, which are grammatically marked, from nontopical, grammatically unmarked objects" (Dalrymple and Nikolaeva 2011: 219). Their conception of topical object is similar to Givón's view of subject as primary topic and object as secondary topic (Givón 2001: 198). However, they point out that "objects are just as likely to be topics as to be focus" (Dalrymple and Nikolaeva 2011: 167). Furthermore, their analysis does not relate formal markedness to functional markedness for objects, as they think that "topical objects are common in human discourse, [and that] formally marked objects are just as frequent in languages with DOM as formally unmarked objects" (Dalrymple and

Nikolaeva 2011: 166). However, our corpus data of Spanish do not give full support to this idea: it is true that more than 50% of the objects are definite (see Table 1), and that definiteness strongly correlates with topicality. However, DOM in Spanish does not depend primarily on definiteness but on animacy. And both animate objects and formally marked objects are clearly less frequent in texts than inanimate objects and unmarked objects (see Table 4). Moreover, while I agree that formally marked objects are high in animacy, individuation, referentiality, and topicality, I do not think that they represent the canonical or unmarked object in Spanish. The Object grammatical relation is not characterized, in general, by its referential autonomy. It is semantically more dependent on the verb than the subject; there are cognate objects that extend the meaning of the verb (*vivir la vida* 'to live life'), verb object idioms (*estirar la pata* 'to kick the bucket [lit. 'to stretch one's leg']), objects with light verbs (*dar un paseo* 'to take a walk'), etc. Some languages incorporate nouns in the morphology of the verb, the morphologically complex V-N compound becoming an intransitive verb. Spanish does not have morphological processes of incorporation, and *a*-marking is used to signal the individuation and referentiality of the object. Only in this sense is *a*-marking a signal of higher transitivity. But, whereas referentiality and topicality are noteworthy properties of subjects, an autonomous and topical object must be seen as "atypical" (Laca 1987: 309).

Iemmolo (2010: 258) also contends that "DOM could be assumed to iconically signal the fact that the direct object has nontypical pragmatic and semantic properties", a thesis that is fully in line with the one defended in this article. He shows that DOM in several Romance languages "emerges in pragmatically and semantically marked contexts, namely personal pronouns in (mainly left) dislocation contexts" (2010: 247), that is, as a marked topic. In Spanish "topicality is no longer the main parameter triggering DOM", which has been extended to "animate and definite objects regardless of their information status" (2010:265). However, information status is not irrelevant in *a*-marking and is one of the main parameters triggering clitic doubling. Despite the fact that ADESSE was not designed to code information status, it records the sequential order of the participants with regard to the verb. Preverbal position of a constituent is associated with the informative function of Theme, the point of departure of the message (Halliday 2004: 64), and it is the usual position of the Subject syntactic function. In Table 6, we can see that there is a higher relative frequency of *a*-marking in preverbal position; but that clitic doubling is more clearly dependent on discourse factors than *a*-marking and is strongly associated with preverbal position, and in (15) and (16), irrespective of the animacy of the referent.

Table 6: Object [DO/IO] position, doubling, and *a*-marking in two-participant clauses (relative and interrogative pronouns excluded)

	Pre-verbal	**Post-verbal**	**Odds Ratio**[11]
All objects			
Doubling (yes/no)	1295/600	793/49293	134.16
a-marking (yes/no)	890/1005	4592/45494	8.77

[11] The Odds Ratio (OR) has been calculated for marked coding (*a*-marking or doubling) in preverbal position as the ratio of the odds of marked coding occurring in preverbal position to the odds of it occurring in postverbal position. The higher the OR, the stronger the association between marked coding and preverbal position.

Animate object NPs			
Doubling (yes/no)	855/45	588/4495	145.25
a-marking (yes/no)	856/44	4033/1050	5.07
Inanimate objects			
Doubling (yes/no)	440/555	205/44798	173.25
a-marking (yes/no)	34/961	559/44444	2.81

In post-verbal position, doubling is rarely present, except for personal pronouns, for which doubling is almost obligatory.[12] Cases of post-verbal NP doubling are subject to dialectal variation (they are far more common in the Porteño Spanish of Argentina and Uruguay, for example) and to additional discourse factors. In ADESSE the information status of noun phrases referring to participants is not annotated, so it is not possible to provide quantitative global data. However, many examples from the textual corpus show a clear association between doubling and information status:

(19)   *entonces de pronto digo: "<…>¿Conocés **a Elena Garro**?" Y yo veo que a Alejandra los pelos se le paran así <…> "¿Y de dónde **la** conocés vos **a Elena Garro**?*   (BAI: 418)

'and then suddenly, I say: "<…> Do you know Elena Garro?" And I see that Alejandra is completely amazed <…> "and from where do you know Elena Garro?"'

In this example, the first mention of *Elena Garro* is not doubled because it is new information. In subsequent mentions, the referent has been activated; it is therefore no longer new information and becomes doubled. So, clitic doubling of lexical definite NPs is used to refer to referents that are highly accessible from the discourse or situational context. That is, "clitic doubling can be interpreted as the formal correlate of an intermediate level of referent accessibility, along a continuum which has weak pronouns (i.e. clitics) and lexical NPs at either end" (Belloro 2007: 131).

The clitic doubling construction exemplified in (19) is typical of Porteño Spanish, where both clitic doubling and *a*-marking are used to mark the prominence and topicality of the object more extensively than in the Spanish of Madrid (Dumitrescu 1997, 1998), but there are a few examples of postverbal object doubling in peninsular Spanish (20), even when the object is inanimate (20c). The referents in these cases are always highly accessible from discourse context.

(20)   a.   *yo traté muy muy íntimamente a Madariaga, a Salvador de Madariaga y yo **le** conozco mejor **a Salvador** que sus hermanas*. (MAD: 279)
'I treated Madariaga, Salvador de Madariaga, very very intimately and I know Salvador better than his sisters do.'

---

[12] The word "almost" is justified here by the presence in the corpus of three examples of non-doubled personal pronouns. An anonymous reviewer points out that three examples of non-doubling do not invalidate the statistical tendency for doubling to be basically obligatory.

b. *Tras una discusión con el director, José Antonio Lorente, **lo** apuñaló **al responsable**.* (2VO: 017)
'After an argument with the director, José Antonio Lorente, he stabbed the person in charge.'

c. *¿Pero es que **las** dan muy mal **esas optativas**?* (MAD: 388)
'But, do they teach those optional subjects very badly?'

Apart from this, the quantifier *todo*(*s*) and definite numerals usually appear as doubled in all varieties of Spanish (*Los conozco a todos/a los dos* 'I know everybody/both of them'). In any case, indexing of post-verbal object NPs is, in global terms, the exception and not the norm.

However, object doubling is the norm and not the exception with *gustar*-type verbs in all dialects.

(21) *Esos rascacielos que **le** gustan **a la Andrea**,...* (SON: 197)
those skyscrapers that 3SG.DAT like   to the Andrea
'Those skyscrapers that Andrea likes...'

These verbs also prefer a dative clitic, and not accusative. But note that a preference for dative case does not imply an automatic preference for clitic doubling. For example, the verb *avisar* 'warn, inform' is not found in this corpus with an accusative clitic but only with the dative (*avisarle*), as in (22a), yet most *a*-marked full NPs are not clitic-doubled with this verb (22b).

(22) a. *Yo encenderé   mientras tú   le   avisas.* (COA: 023)
I   light.FUT.1SG while   you   3SG.DAT warn.2SG
'I will light (it) while you warn him.'

b. *Ve a avisar a Bagnone.* (COA: 073)
Go to warn   to Bagnone
'Go and warn Bagnone.'

The selection of human *a*-marked objects and the quantitative preference for dative case makes this verb similar to *gustar*-type verbs, but the postverbal position and the absence of clitic doubling in (22b) makes it different from *gustar*-type verbs. This is a good example of the fuzzy limits between DO and IO. The choice of the clitic case is addressed in the next section.

4.2 The case of pronominal clitics

Besides variable object marking involving the preposition *a* vs. ø and doubling vs. no doubling, the variable marking of the Spanish Object becomes apparent through the use of case in pronominal clitics. The system of the pronominal clitics in the singular is shown in Table 7.

Table 7: System of personal clitics (singular) in Spanish

	Accusative		Dative
	Masculine	**Feminine**	
3rd person & Addressee (polite)	*lo*	*la*	*le*
2nd (Addressee)	*te*		
1st (Speaker)	*me*		

As we can see here, only the third-person clitics distinguish between accusative and dative case. First- and second-person clitics do not distinguish between direct and indirect object functions: *me ve* 'he sees me', *me gusta* 'I like it', *me avisa* 'he warns me', *algo me ocurrió* 'something happened to me'. So, third person clitics are usually used as a test for the direct vs. indirect object functions (*lo ve* 'he sees him' vs. *le gusta* 'he likes it').

However, in two-participant clauses some factors distort the view that the clitic form is a good index of syntactic function, since the case system seen in Table 7 does not work in all dialects of Spanish. In some varieties, there is a preference for the use of the dative form *le* as human masculine Direct Object ("leísmo de persona": *verle a él* 'to see him'), that is, in contexts where other varieties use *lo* or where the accusative *la* is preferred for feminine referents (*verla a ella* 'to see her'). An additional complication arises from the fact that in some Peninsular Spanish dialects a "referential" system (Table 8), or a variant of it, is employed; this system is based on gender and discreteness but not on case, and it is used indiscriminately both in two-participant and in three-participant clauses (Klein-Andreu 1981, 2000; Fernández Ordóñez 1993, 1999). Here only properties of the referent are relevant: *le* is simply count masculine and *lo* is used if a noun of either gender is regarded as non-discrete.

Table 8. "Referential" gender-based system of third-person clitics (non-plural forms) (adapted from Fernández Ordóñez 1999: 1360)

Count		Mass
Masculine	Feminine	
le	*la*	*lo*

The influence of this "referential" system may explain why some speakers of Spanish at times use *lo* and *la* for R in ditransitive clauses, and why they often use *le* in two-participant clauses. However, our data do not support the claim that there are two different independent systems at work (the case system and the referential system), each one internally coherent; rather, out data point to one unstable system affected by several competing forces. Of course, this may be attributed to the fact that our corpus includes mainly literary texts by authors from different geographical origins.[13] The standard language seems to be a partial compromise between the case system and the referential system. This compromise, which is sanctioned by the Real Academia, is based on the case system but tolerates *le* for human masculine singular objects. On the other hand, the standard use also shows the emergence of a dynamic system subject to several forces that are manifested in most varieties of Spanish to a greater or lesser extent.

---

[13] This corpus, then, is not especially useful for dialectological studies focusing on vernacular varieties.

The *le* vs. *lo/la* alternation must be seen as a particular case of DOM, alongside with *a*-marking and doubling (Flores and Melis 2007). One of the main factors in the choice of *le* vs. *lo/la* is clearly animacy, as can be seen in Table 9. Dative *le(s)* is almost wholly limited to animate objects, and is the norm when politely referring to addressees. This distribution, plus the neutralization of accusative–dative distinction in first and second person, demonstrates that case in pronominal clitics is related to the animacy hierarchy, as are *a*-marking and doubling, as speaker and hearer are situated at the highest ranks of the animacy and topicality scales.

Table 9: Animacy and case of Obj [DO/IO] clitics in 2-participant clauses [N=13,930] (ADESSE)

	Acc *lo(s)/la(s)*	Dat *le(s)*	% Dat
Addressee (polite)	53	242	82.0%
Animate	3375	3962	54.0%
Inanimate	6163	135	2.1%

Among animate objects, gender and number are also relevant in the choice of case. Dative *le* is much more frequent in the singular masculine than in the plural or in the feminine, as shown in Table 10.

Table 10: Gender, number, and case in 3rd-person animate Obj [DO/IO] clitics [N=7337] (ADESSE)

	Acc *lo(s)/la(s)*	Dat *le(s)*	% Dat
Masculine singular [*lo* vs. *le*]	1147	2906	71.8%
Masculine plural [*los* vs. *les*]	471	475	50.4%
Feminine singular [*la* vs. *le*]	1646	543	24.8%
Feminine plural [*las* vs. *les*]	111	38	25.5%

As the use of feminine forms resembles the case-based etymological paradigm better, it is preferred in Spanish grammars as a test for whether a verb is transitive. A verb is considered transitive if it prefers the accusative for a feminine referent. Therefore *ver* 'to see' is transitive and takes a DO (***La** vieron a ella* 'They saw her'), whereas *gustar* is not and takes an IO (***Le** gusta a ella* 'She likes him/her/it´). Accordingly, "leísmo" is defined as the use of an "anomalous" dative form *le* with a "transitive" verb.

"Leísmo" is much more frequent in Spain than in Latin America, and in fact many linguists claim that in American Spanish there are no proper cases of "leísmo". For example, DeMello (2002) argues that the *le/lo* or *le/la* alternation always reflects a functional contrast between Direct and Indirect Object. A similar line of reasoning is developed by Fernández Ordóñez (1999: 1323–1341) for all case-distinguishing dialects. However, both in Spain and in Latin America, there is considerable variation across texts and across verbs. In other words, many two-participant verbs in our corpus do not exclusively take either the accusative or the dative to express the Object. Therefore, the mere possibility of an accusative form cannot be used mechanically as a test for transitivity, as even non-"leístas" Latin American writers use dative case (23a) with some verbs that in other contexts appear with the accusative (23b):

(23) a. ***Les*** *obedecí* *a ciegas*. (CRO: 042)
　　　3PL.DAT obeyed.PFV.1SG blindly
　　　'I obeyed them blindly.' (García-Márquez [Colombia, America])

　　 b. *Mamá **la*** *obedeció*. (SUR: 039)
　　　mom 3SG.ACC.F obeyed.PFV.3SG
　　　'Mom obeyed her.' (García-Morales [Extremadura, Spain])

Still, the non-subject argument of two-participant clauses is indexed very frequently by the accusative *lo* in texts by Spanish authors of any geographical origin, although there are some clear differences in the frequency figures. Table 11 summarizes data from our corpus in a restrictive way, taking only those verbs that in traditional terms may be considered "transitive" using the criterion that these verbs may be attested with the accusative or present some other formal mark of the DO. In the table, masculine singular animate referents are shown in a separate column because of their strong association with *le*.

Table 11: Third-person DO clitics with animate referent in transitive clauses: Text, author's origin (S=Spain; A=(Latin) America), totals, and % dative (*le* vs. *lo*) for masculine singular (M.SG) and for not masculine singular (¬M.SG)

TEXT	ORIGIN	M.SG	% *le* (M.SG)	NOT M.SG	% *le* (¬M.SG)	TEXT	ORIGIN	M.SG	% *le* (M.SG)	NOT M.SG	% *le* (¬M.SG)
PAS	S	36	100%	35	17%	2VO	S	9	56%	7	43%
CIN	S	42	95%	65	31%	1VO	S	42	52%	15	20%
HOM	S	19	95**%**	42	19%	LAB	S	110	49%	87	22%
SUR	S	60	95%	208	10%	AYE	S	41	46%	16	13%
SON	S	553	93%	266	32%	1IN	S	69	46%	40	8%
MIR	S	170	91%	36	19%	3VO	S	16	44%	16	38%
MOR	S	55	91%	64	11%	RAT	S	59	36%	43	19%
TER	S	211	90%	104	21%	SEV	S	48	35%	40	8%
OCH	S	111	87%	54	11%	CAI	S	86	26%	168	5%
COA	S	47	85%	22	14%	TIE	A	18	11%	33	0%
PAI	S	103	80%	33	33%	ZOR	S	10	10%	23	9%
CAR	S	120	77%	71	17%	BAI	A	91	9%	119	6%
2IN	S	16	75%	14	14%	HOT	S	12	8%	58	7%
LIN	A	4	75%	3	0%	DIE	A	135	6%	161	7%
JOV	S	165	73%	99	17%	GLE	A	96	5%	202	1%
MAD	S	92	59%	130	25%	CRO	A	195	3%	127	3%
USO	S	51	59%	167	10%	HIS	A	159	2%	120	1%

Table 11 shows that, even if we limit ourselves to traditional DOs, there is substantial variation across dialects and writers. This makes it difficult to establish a clear-cut division between two patterns: Subj–DO vs. Subj–IO. Moreover, the *le/lo* alternation occurs with most verbs; most two-participant verbs allow either accusative clitic (*lo*) or dative *le,* even in the same text:

(24) a. *Lo que realmente **lo**   preocupaba era una ceremonia.* (HIS: 131)
          What   actually   3SG.M.ACC worried   was a   ceremony
          'What actually worried him was a ceremony.'

      b. *Esos bienes que tanto   **le**   preocupan*
          these goods that so much 3SG.DAT worry

          no   le   servirán de nada   (HIS: 070)
          don't 3SG.DAT serve   of nothing

          'These goods that worry him so much will be useless to him.'

(25) a. *No **lo**   he   visto…* (LAB: 231]
          not 3SG.M.ACC have.1SG seen
          'I haven't seen him ...'

      b. *Desde que **le**   vi   por primera vez...* (LAB: 158)
          From   that 3SG.DAT see.PST.1SG for first   time
          'Since I saw him for the first time…'

It can be seen from the corpus findings, then, that there are not two distinct classes of verbs, each with its own syntactic pattern, but that verbs may show a graded preference for the accusative case or for the dative case. This preference can be represented on a scale, where verbs with an accusative case percentage of (close to) 100% represent the schema Subject – Direct Object and verbs with a dative case percentage of 100% represent the schema Subject – Indirect Object. All the other verbs that alternate between accusative and dative to different degrees are positioned in between the two ends of the spectrum:

Table 12: 2-participant verbs: % of **dative clitics** (vs. accusative clitics) with animate objects (any gender or number)

0–20%		20– 60%	60-90%	90-100%
tener	*conocer*	*observar*	*odiar*	*gustar*
encontrar	*buscar*	*ayudar*	*alegrar*	*parecer*
matar	*ver*	*convencer*	*asustar*	*hablar*
sentir	*mirar*	*oír*	*engañar*	*pasar*
leer	*llamar*	*seguir*	*amenazar*	*pegar*
	llevar	*dejar*	*preocupar*	

← ─── →

SUBJ-DO	SUBJ-DO/IO	SUBJ-IO

At the ends of this scale, verbs show a strong tendency towards either accusative or dative independently of the gender of the Object referent. But, in the middle, verbs show a clear bias motivated by the gender of the Object, with masculine *le* (rather than *lo*) and feminine *la* (rather than *le*) in most texts of Peninsular Spanish origin.

Within this continuum, it is possible to recognize at least three groups of verbs, with no sharp boundaries between them:

(i) verbs that almost always take the dative, independently of animacy and gender of the Object. That includes the human Experiencer [E] of *gustar* 'like' and other verbs of sensation, the human addressee or Receiver [R] of *hablar* 'talk' and other

verbal processes, but also the (mostly human) affected participant of an action verb like *pegar* 'hit';

(ii)   verbs that alternate between dative and accusative case, independently of the gender and number of the Object. This includes verbs like *asustar* 'frighten', *alegrar* 'make happy', *preocupar* 'worry', *ayudar* 'help'. The selection of case here has clear semantic consequences: generally speaking, dative case tends to be used with (a) more static, uncontrolled relationships, (b) and/or less effective events not completely affecting the object, (c) and/or events with more autonomous activity by the object, (d) and/or events with a subject lower in the animacy hierarchy (Vázquez Rozas 1995, 2006). In other words, dative Object is associated with lower transitivity (Hopper and Thompson 1980);

(iii) verbs that prefer the accusative case, except for animate masculine singular objects where, depending on dialect and individual preferences, *le* and *lo* can be used: *ver* 'see', *mirar* 'look', *conocer* 'know', *matar* 'kill', *tener* 'have', etc. The behavior of these verbs is not homogeneous and some verbs seem to tend more towards dative than others. For example, the frequency of use of the masculine *le* (vs. *lo*) is higher with *observar* 'observe' (64%) than with *conocer* 'know' (32%) and *ver* 'see' (28%). A more detailed study is needed to investigate if these differences are significant and if they are semantically motivated.

4.3 An interim summary

We have discussed three different phenomena related with the variable coding of Objects in two-participant clauses of Spanish. Table 13 provides a summary of the frequency of each of these phenomena.

Table 13: Variable coding of the Object of two-participant clauses (ADESSE)

	ø-marked	*a*-marked	N
A-marking	51673 (90.0%)	5710 (10.0%)	57386
	not-doubled	doubled	
Clitic-doubling	55191 (96.2%)	2195 (3.8 %)	57386
	accusative	dative	
Clitic case	9591 (68.9%)	4339 (31.1 %)	13930

No dialect of Spanish has categorical rules for the use of *a*, clitic doubling, or clitic case in two-participant constructions. Everywhere, a gradient can be observed. The less frequent options are *a*-marking, clitic doubling, and dative clitic. The "canonical" object is a postverbal phrase not marked by a preposition and not indexed in the verb. If the object is cliticized, the canonical case of the object is accusative. But all three phenomena are independent and subject to a large amount of variation. The tendency is to have morphologically marked objects for referents high in the animacy hierarchy and morphologically unmarked nominals for referents low in the animacy hierarchy. Referential properties of the object are the main factor governing *a*-marking. Clitic doubling and clitic case are also related to the animacy hierarchy, but the cut-off point and the frequency cline are different for each phenomenon. Clitic doubling correlates more strongly with information status (topicality and/or accessibility). Case is also governed by dialect variation, gender and number, and type of process. *A*-marking and clitic doubling depend to a lesser extent on the lexical meaning of the verb.

If a verb admits the more frequent options, namely ø-marking, no doubled objects, and accusative case, we would say that it is a transitive verb and that its object is a Direct Object [DO]. If a verb requires *a*-marking and dative case, and prefers clitic doubling, we would say that it is not transitive and it is construed with an Indirect Object [IO]. But it is very difficult to decide the syntactic function of many specific examples, because no overt formal feature gives a definitive indication of syntactic function *per se*, and verbs align along a multidimensional continuum concerning their preferences for one or another variable coding feature.

5. Object coding in ditransitive clauses

The topic of this section is marking of the objects of ditransitive clauses and the alignment patterns that derive from it. In Table 14, we see a very different distribution of the coding properties of T (Transferred object of prototypical ditransitive clauses) and R (Recipient in prototypical ditransitive clauses):

Table 14: Frequency distribution of variable coding in ditransitive clauses
[N=8445] (ADESSE)

	T	R
Preposition *a* (vs. ø)	0.7 %	100 %
Clitic doubling (vs. not doubled)	0.9 %	42.8 %
Dative (vs. accusative) case (3rd-person clitics)	0.3 %	99.7 %

When examining the coding of R and T, the first thing to be taken into account is the high animacy and accessibility of R, as we have seen in Table 3. Most ditransitive examples (5246/8445 = 62.1%) present the pattern <Subj Dat-V NP>, where the dative clitic [= R] has an animate referent and the NP [= T] refers to an inanimate entity. There are relatively few examples where R is a full NP, and relatively few examples with an animate T. As a result, the preconditions for *a*-marking, doubling, and case choice are seldom met. From the inherent properties of participants, we expect a split of object coding between T and R. If T is almost always ø-marked and accusative, this may be due solely to the fact that it is almost always inanimate, and not to its syntactic function. The same applies mutatis mutandis to R, whose *a*-marking and dative case may be attributed to the fact that is almost always animate. Therefore, it is necessary to examine in more detail whether the coding of T and R is determined by their inherent properties, such as animacy, or by their syntactic function (or both).

On the whole, *a*-marking is obligatory with R NPs (100% of the corpus) and almost impossible with Ts (only 0.7%). This clearly suggests an indirective alignment as far as preposition *a* is concerned because only in R, but neither in P nor in T, is *a*-marking obligatory, and because in P as well as in T the option without *a* is the most frequent. We have seen that *a*-marking is strongly conditioned by the animacy of the referent, so the low percentage of *a*-marking in T could simply be due to the fact that T is usually inanimate. However, the fact that inanimate Rs are also *a*-marked, as in (26), shows that Spanish does not strictly follow what Kittilä (2006) labels an animacy-based strategy, but rather that it follows a role-based strategy

(26)   *Es   difícil   atribuir=le   sentido **a** esa opinión.* (LIN: 073)
         Is   difficult   attribute=3SG.DAT sense   to this opinion
         'It is difficult to make sense of this opinion.'

But what happens in those rare cases where both T and R are animate? If both objects are animate, it is even rarer for both to be instantiated by NPs. In most cases, R is instantiated only as a personal clitic index. In the pattern T = NP, R = clitic, 55.3% of T-NPs are *a*-marked, a percentage that is rather lower than that of the animate object of two-participant clauses (75%). However, if we look at the examples in the corpus, almost all the animate ø-marked Ts correspond either to animals or to nominals of an unspecific referent (27b), whereas those with *a* always have an specific individuated referent (27a):

(27) a.   *Da=me   **a-l**   niño, Hilaria – ordenó Mónica.* (DIE: 124)
          give=me [R] to[T]-the boy   Hilaria   ordered Monica
          'Give me the boy, Hilaria –ordered Monica.'

     b.   *¡Si Dios me   hubiese   dado un hijo como él!* (SON: 339)
           if God   me [R] have.PST.SBJV given a   son like   him
          'If only God had given me a son like him!'

It seems as if in this case the choice of *a*-marking for T follows the same principles as for P in monotransitive clauses, even though potential ambiguity between T and R roles may exist.
     In the less frequent combination, T = clitic and R = NP, R is always *a*-marked. These are also potentially ambiguous cases, and they are only disambiguated contextually:

(28) *Tu   madre **te**   entrega   **a**   ese dottore de mierda,*
     your mother 2SG [T] hand over to [R] that dottore of shit

     pero tu   abuelo   te   sacará adelante. (SON: 123)
     but   your grandfather 2SG take   forward
     'Your mother is handing you over to that fucking *dottore*, but your grandfather will keep you going.'

When both R and T are animate **and** full NPs or Pronouns, there is a preference in our data for reserving *a*-marking for R; only in this case do most examples (15 out of 16, including those in which T is indefinite or refers to animals) show a ø-marked T, as in (29a), but with *a*-marking not unattested (29b):

(29) a.   *El viejo   entrega **ø** el niño **a**   Renato.* (SON: 123)
          the old man hand over [T] the boy to[R] Renato
          'The old man hands over the boy to Renato.'

     b.   ***A**   mí   me   quitaron   **a**   mi negra consentida.* (DIE: 082)
          to[R] me 1SG steal.PFV.3PL to[T] my black.F complaisant
          'I was robbed of my complaisant darling.'

27

The variation never affects R marking, which always receives *a*. In the only example of two NPs with a human T and a non-human R, only R is *a*-marked:

(30) *Con el niño en brazos se acerca a la ventana, como exhibiendo su triunfo a Milán entero*,

    o   *presentando* ∅   *el*   *niño* ***a***   *la*   *nieve* *amiga.*   (SON: 218)
    or   presenting   [T]   the   boy   to [R]   the   snow   friendly

    'Holding the boy on his arms, he approaches the window as if he was exhibiting his victory to all of Milan, or presenting the boy to the friendly snow.'

In sum, with respect to animate referents it can be said that *a*-marking follows an indirective alignment, being obligatory with R and variable with T and P.

The second parameter of variation is the case of the third-person clitic. Is there also "leísmo" in ditransitives? With regard to written standard Spanish, the distinction between accusative and dative case in ditransitive clauses is clear: with very few exceptions, a dative clitic always represents the R participant, and an accusative clitic always represents the T participant. Only in three examples (0.3%) in the corpus, out of a total of 974 clitics referring to T, is the dative used where the accusative is expected (see (31)).

(31) *Entrégue =**le***   *a la policía*   *o déje=le*   *marchar.*
    hand over 3SG.**DAT**[T] to the police [R] or let=3SG.DAT go

    'Hand him over to the police or let him go.'

As for R clitics in the third person, the dative is the norm and only 13 examples out of 3547 take the accusative. Half of those occur in the subcorpus of spoken Spanish from the Madrid area, and only *la* and *los* forms are attested, as in (32):

(32) a. *Debo*   *pasar**la***   *setecientas* *libras* *mensuales*
        must.1SG   PASS=3.ACC.F 700   pound   monthly

        de por vida. (CIN: 051)
        of   for life

        'I must transfer her seven hundred pounds monthly for life.'

    b. *él* *siempre* ***la***   *da*   *muchos* *permisos*   (MAD: 349)
        he   always   3.ACC.F gives   many   permissions

        'He always gives his consent to her.'

This use of accusative-instead-of-dative (*laísmo* and *loísmo*) is partly motivated by a dialectal system based on gender, as seen in Table 8. However, the use of *la* in this context has also been seen as a real accusative, added to a V+NP cluster, and not as an anomalous dative (see Romero 2013).

Object clitics may occur alone or doubling a co-nominal. We have seen that clitic doubling is a variable property of objects in two-participant clauses. Table 15 gives a summary of object clitic-doubling in ditransitive clauses.

Table 15: The instantiation of object arguments in ditransitive clauses (ADESSE)[14]

| | **T** | | | **R** | | |
	N	% (i)	% (ii)	N	% (i)	% (ii)
Doubled (clitic+NP)	77	0.9%	1.0%	936	11.1%	42.8%
Not doubled (NP)	7470	88.4%	99.0%	1249	14.8%	57.2%
Clitic only	898	10.6%	--	6260	74.1%	--
	8445	100 %	100 %	8445	100,0%	100 %

Once again, the distribution of clitic-doubling follows indirective alignment rather clearly. The global percentages of P-doubling (3.8%, as seen in Table 4) are closer to those of T (1.0%) than to those of R (42.8%). That said, R indexation either by a clitic alone or by a clitic doubling an NP is, to a certain extent, typologically anomalous. Argument indexing is expected to occur with more salient participants and is expected to follow the grammatical relations hierarchy (Croft 1988; Siewierska 2004: 43), where direct objects [DO] and primary objects [PO] are ranked higher than indirect objects [IO] and secondary objects [SO]:

(33)   Grammatical Relation Hierarchy (GRH): likelihood of indexing as a function of grammatical relations
Subject > Object1 (DO/PO) > Object2 (IO/SO) > Oblique

Given that, in Spanish, indexing of R has preference over indexing of T, this can be taken as a symptom of the primacy of R among the objects (Company 2001; Comrie 2012: 20). However, Givón (1984, 2001) argues that there are two separate and distinct functional hierarchies: the "hierarchy of the grammaticalized pragmatic case roles" or "hierarchy of grammatical relations" Subj > DO > other, and the "hierarchy of semantic case roles" Agent > Dative/Benefactive > Patient, where Dative must be understood as equivalent to R. Both hierarchies, together with hierarchies of animacy, definiteness, and topicality, determine the likelihood of agreement and are motivated "by the text-frequency of anaphoric pronominalization" (Givón 2001: 426). This leads to a preference for R indexing independently of grammatical relation and alignment type. However, if the Grammatical Relation Hierarchy in (33) is correct, what we expect is a tendency whereby, in order for R to be indexed, either R becomes the primary object [PO] (by secundative alignment or by applicative diathesis) or indexing of up to three arguments (as in Basque and Georgian) is made possible. In a sample used by Haspelmath, there is no language with indirective alignment in which R is indexed and P is not (Haspelmath 2005a: 13), and Siewierska (2003: 356; 2004: 44) only mentions the case of Gude, a Chadic language, where the R-index is combined with an applicative morpheme.

What is unexpected about R is its twofold character: on the one hand it is prototypically salient in animacy and definiteness, but on the other hand it is conceptually closer to obliques such as goal and benefactive. T, rather than R, is the most affected entity in a transfer event and this favors indirective alignment and unmarked flagging of T. But R is more topical than T, as we have seen in Table 3, and

---

[14] The percentages in columns (i) take into account the indexing-only option. The percentages in columns (ii) compare "doubled" vs. "not doubled", as was done in Section 4 for objects in two-participant clauses.

this favors its indexing. This is the basis of Givón's hierarchy of semantic roles, and the reason for Haspelmath's generalization: "In ditransitive constructions, indexing shows no strong alignment preference (16 IND: 22 SEC), but flagging strongly prefers indirective alignment (58 IND: 6 SEC)" (Haspelmath 2005a: 7). Of the sixteen languages with indirective alignment indexing in his sample, ten show indexing of P and T, but not of R, and in all ten languages, R is flagged by a case-marker or adposition. Moreover, eight of these languages (Bagirmi, Lango, Lavukaleve, Tzutujil, Slave, Apurinã, Hixkaryana, and Kipea) flag R and not T or P. This clearly suggests that R is treated in those languages as an oblique secondary participant. If we want to match the topicworthiness, animacy, and definiteness of R with the grammatical relations hierarchy, the most obvious strategy is to make R the first object. This can be done either by secundative alignment in the basic diathesis, by applicative constructions, or by object alternation as in English (see also Malchukov et al 2010: 20–21)

Spanish has a different strategy: give priority to R in indexation and maintain the proximity to obliques with obligatory *a*-marking. With indexation, the Indirect Object syntactic function makes an argument salient and it serves to present as central participants those elements which are not valency-bound, such as benefactives, possessors, locatives, etc. For example, the affected possessor of verbs of touching, hitting, and others, as in (34a); may be "raised" to Direct Object, leaving the body part as a marginal oblique (34b), or to dative Indirect Object, retaining the body part as Direct Object (34c). In the latter case, the dative clitic indexes the possessor as an additional core participant and the clitic cannot be omitted, as in (34d) (Gutiérrez Ordóñez 1999: §30.6; Vaamonde 2011).

(34) a. *Golpeé   el   brazo   de   Juan.*
          hit.PFV.1SG   his   arm   of   Juan
          'I hit Juan's arm.'

      b. *A   Juan   lo   golpeé   en   el   brazo.*
          To Juan   3SG.**ACC**.M   hit.PFV.1SG   on   the   arm
          'I hit Juan on the arm'

      c. *Le   golpeé   el   brazo   a   Juan.*
          3SG.**DAT**   hit.PFV.1SG   the   arm   to   Juan
          'I hit Juan on the arm.' (lit. 'I hit Juan the arm.')

      d. */? *Golpeé   el   brazo   a   Juan.*
               hit.PFV.1SG   the   arm   to   Juan

In canonical ditransitive constructions, which denote a transfer event and where the R is valency-bound, it is easier to find examples of alternation between a clitic-doubled R NP (35b) and an NP with no clitic indexing (35a):

(35) a. *Da   la   policía francesa setecientas mil   peseta-s*
         gives   the   police   French   700   thousand   peseta-PL
         a   quien lo   encuentre.   (MAD: 049)
         to   who   3.ACC.M   find.3SG.SBJV
         'French police offer seven hundred thousand pesetas to whoever finds him.'

   b. *Al   subir, **le**   das   ese cheque **a** mi padre*. (OCH: 038)
     to.ART go up 3.DAT give that check to my father
     'When you go up, you will give that check to my father.'

According to Maldonado, "in all dialects of Spanish, the clitic establishes a stronger link between the indirect object and the agent as some entity is transferred. As the link becomes looser, the acceptability of omitting the clitic increases" (Maldonado 2002:18). Doubling, therefore, reinforces the status of R as a central participant in the event, and is also associated with a higher discourse prominence and accessibility than the non-doubled counterpart (Belloro 2007). Clitic-doubling optionality has even been compared to object alternation in English and other languages (Demonte 1995; Nishida 2012). However, it must be noted that whereas there might be some functional analogy between the English double object construction and the Spanish clitic-doubled indirect object construction, there is absolutely no formal equivalence between these two constructions.

Besides flagging and indexing, we might briefly consider additional properties of grammatical relations such as order and passivization. Word order, by itself, cannot be used to identify the alignment type if both participants are in preverbal or postverbal position. However, there is a general preference for R–T order, except in indirective constructions, which prefer T–R order when the R is flagged by an adposition (Heine and König 2010; Malchukov et al. 2010: 16–17). Therefore, a preference for the order V–R–T in Spanish can be taken as a move towards primary object type (Company 2001: 9–10). However, word order is relatively free in Spanish and does not provide a clear picture here. First, ditransitive clauses prefer an argument structure with at most one lexical argument and a tendency for the IO to be instantiated as a clitic only. In the cases where there are two lexical objects (T and R), the ADESSE database still shows a preference for V–T–R order (978 clauses) over V–R–T order (552 clauses), that is, word order is also more coherent with indirective alignment.

Finally, passivization clearly shows indirective alignment in Spanish: only P and T can become passive subjects. Some authors have adduced a number of examples in which an IO seems to be passivized, but the status of the passivized argument is not at all clear in these cases. Moreover, the R participant of canonical transfer ditransitive constructions can never be passivized (**María fue entregada un galardón* 'Mary was given an award'). Certainly, in typological terms, the topicality of R should favor its passivization: "R-passivization is generally preferred over T-passivization. This would make sense given that the function of passives is to topicalize the object, because the R tends to be more topical in the ditransitive construction" (Malchukov et al. 2010: 30). Yet in Spanish there are alternative strategies for the topicalization of R: relatively free word order allows the topicalization of IO without it having to become a subject; and both the periphrastic passive and the *se*-impersonal or *se*-passive may serve to shade the agent (*A María le fue entregado un galardón*, *A María se le entregó un galardón*).

6. Objecthood in Spanish: Summary and discussion

In the previous sections, the use of three main variable coding devices of Spanish objects has been discussed. Table 16 summarizes the overall distribution of each coding option:

Table 16: Frequency distribution of variable coding of Objects (ADESSE)

	Non-subject in 2-participant clauses	*T* = DO in ditransitives	*R* = IO in ditransitives
Preposition *a* (vs. ø)	10.0 %	0.7 %	100 %
Doubling by clitic (vs. no doubling)	3.8 %	0.9 %	42.8 %
Dative (vs. accusative) case (3rd person clitics)	31.1 %	0.3 %	99.7 %

As Table 16 makes clear, the distribution of variable coding is sensitive to the number of core participants. In ditransitive clauses, T and R can be singled out clearly, in that there is a nearly categorical contrast in their use of the preposition *a* and the dative case: the former is almost the sole option for R, and the latter is almost never used for T. Doubling is less significant as an index of syntactic function, but it is very significant in terms of status as a prominent participant, as it is more tied to R than to any other object. Each of the coding properties of R is also an option for the coding of the object of two-participant clauses, but it is always a less frequent option than the coding properties of T.

Therefore, Spanish cannot be considered a primary object language (a language with secundative alignment in some coding system). It is true that R aligns with P in some cases: mostly, when P is animate and definite for *a*-marking, when it is topicalized for doubling, and when it is animate, masculine, and singular for dative third-person clitic. However, all those cases are less frequent than those where P is coded like T. My view on this is in keeping with the received view on Spanish: "Here our practice is to adopt the most typical transitive construction, with an inanimate, indefinite P as the major monotransitive construction […]. Thus we say that Spanish has indirective alignment of flagging" (Malchukov et al. 2010: 7).

Nevertheless, variable coding of P or DOM is a very relevant phenomenon in the grammar of Spanish, and must be seen in relation to the coding of participants in ditransitive clauses. The tendency is to have morphological marking for Rs in ditransitives and for Ps high in the animacy hierarchy, whereas we usually have no marking for Ts in ditransitives and for Ps low in the animacy hierarchy. Additionally, clitic doubling correlates more strongly with discourse status (topic and given), and "leísmo" – the use of a dative case clitic in transitive clauses – is governed by dialect variation, gender, and type of process. But no dialect of Spanish has categorical rules for the use of *a*, clitic doubling, or "leísmo". Rather, syntactic functions in two-participant clauses are graded, with no clear-cut distinguishing criteria. The variable coding of objects creates a problem that cannot be solved within the traditional distinction between Direct Object and Indirect Object. While, on this traditional view, one function is defined to the exclusion of the other, the distinction between the two functions can, on the whole, only be assured in ditransitive clauses. In two-participant clauses, variable marking of the object can be understood as a means to approximate the object semantically and formally either to the T-pole or the R-pole. As a result, there would not be two clearly distinct syntactic functions, but rather one Object macrofunction affording a variation space. When the object of a two-participant clause largely resembles IO coding, it bears some semantic similarity to ditransitive Rs, in that it shares with the participant R semantic or pragmatic properties (animacy, topicality, etc.) or a semantic role, as with receivers [R] of verbal processes like *hablar* 'talk'.

Table 17: Objecthood in Spanish

Construction	3-participant	2-participant		3-participant
Participant	T	P	E/R	R
Form	ø ←			→ *a*
	non-doubled ←			→ (doubled)
	accusative ←			→ dative
Syntactic function	DO ←			→ IO

This claim has a corollary for the interpretation of grammatical relations not only at a language-specific level, but also at a construction-specific level (Croft 2001). First, coding in three-participant constructions behaves differently from coding in two-participant constructions: in the former, the priority is to differentiate the roles of R and T, whereas in the latter only P must be differentiated from A; consequently, in two-participant clauses the coding resources available may be used for expressing additional semantic and pragmatic nuances, reflecting whether P is closer to T or to R. Second, the coding and behavioral properties of constructions, including *a*-marking, clitic doubling, case, passivization, word order, etc., are related but independent phenomena. Each one occupies a different region in the syntactic-semantic space and has its own functions and motivations, and it is also partly independent of the number of core participants. There are a few cases of unexpected coding of T and R, such as *a*-marked Ts or dative *le* for Ts, or accusative *la* for Rs, etc., but nothing especially anomalous in any of these rare uses. They are, rather, extensions of the general rules guiding the functioning of variable coding of arguments in Spanish. That said, this does not mean that we can dispense with grammatical relations in the grammar of Spanish or that the notions and terms "Direct Object" [DO] and "Indirect Object" [IO] are useless. They remain valid as reference points where certain behavioral and control properties cluster, provided that we bear in mind that such properties may vary independently and that, ultimately, they serve as clues to identifying entire clause constructions and not only parts of constructions.

The final issue I want to discuss here concerns the motivations of participant coding devices in general and of variable coding in particular. Among the several proposals in the literature as to why grammatical coding of participants is the way it is, the most common have recourse to concepts of iconicity, discriminating argument functions, markedness, and transitivity.

The term iconicity denotes a kind of isomorphism between form and meaning; for example, the use of an adposition such as Spanish *a*, which adds extra material between verb and noun, can be argued to reflect greater conceptual distance between verb and object (Haiman 1983: 790–793). A related sense of iconicity is associated with the concept of markedness: "the more marked a direct object *qua* object, the more likely it is to be overtly case-marked" (Aissen 2003). Haspelmath (2008) argues that grammatical asymmetries are better explained in terms of frequency than in terms of iconicity, and below I will also utilize an explanation based on frequency.

Iconicity may be correlated with economy (Haiman 1983; Aissen 2003), i.e. the preference to avoid the use of marking unless necessary. What motivates coding is its discriminating value; accordingly, an object is morphologically marked so as to be distinguished from a subject when subject and object present similar features (human, topical, etc.), or the morphological coding of R is motivated by the need to distinguish it both from T and from the subject. However, cross-linguistically, functional

disambiguation does not serve as an explanation for all cases of variable object marking: there are similar object marking patterns in ergative languages, like Dolakha Newari (Genetti 1997), where there is no need for a disambiguating morpheme in the patient because the transitive agent always carries an ergative marker. Neither does the discriminating function always work for Spanish. DOM in monotransitives does not block all potential ambiguities, and DOM is not fully absent in ditransitives, giving rise to cases of potential ambiguity between T and R: for instance, there are examples of *a*-marking and *leísmo* for T, and some examples of accusative case (*laísmo* and *loísmo*) for R.

I think that the main motivating factors for variable marking of the object derive from the corpus data we have seen above in Section 2, which give support to the hierarchies of comparative syntactic functions introduced in (13), and which are reproduced here again as (36): there is a polarization between A and P in transitive clauses in terms of inherent agency potential (animacy) and inherent topicality, and there is an analogous polarization between R and T in ditransitive clauses:

(36)   Animacy and topicality hierarchies of comparative syntactic functions
   a.   A > S > P
   b.   R > P > T

P is expected to be less animate, less topical, and less accessible than A, and it is expected to show intermediate values between R and T in animacy and topicality. These properties conform quite well to Comrie's generalization, which he supports with data about inverse marking and DOM from several languages: "The most natural kind of transitive construction is one where the A is high in animacy and definiteness, and the P is lower in animacy and definiteness; and any deviation from this pattern leads to a more marked construction" (Comrie 1989: 128). It is difficult to ascertain what exactly might be understood by "natural kind of transitive construction", but the principle might be stated simply in terms of frequency of use as observed in our corpus data: more complex morphological means are used for less frequent patterns, those contrary to usual expectations. Applied to DOM, that means that animate objects tend to be explicitly coded because animate referents occur more frequently as subjects and inanimate referents occur more frequently as objects (Haspelmath 2008: 14). Using the terms "marked" and "unmarked" as equivalent to "less frequent" and "more frequent", the phenomena of variable morphological marking of objects in Spanish always correlate with marked options, although each variation phenomenon correlates with object properties to different degrees.

(37)   Objects

(*unmarked*)	(*marked*)	
Non-human	Human	→ *a*-marking, dat *le*
Less definite	Definite	
(Part of) Rheme	Topic	→ clitic doubling
Low accessibility	More accessible	→ clitic
Patient	Less affected	→ dat *le*

Note that the marked options characterize semantically and formally both the R participant in ditransitive clauses and the smaller subset of P participants in monotransitive clauses. As these are marked options in Spanish, the alignment must be considered as indirective.

This view relates formal markedness of objects with their functional markedness assessed in terms of frequency or typicality, and differs in this and some other important aspects from Dalrymple and Nicolaeva (2011), who state:

> Most work on DOM assumes that object marking originates from the need to differentiate the object from the subject. However, we claim that DOM actually marks similarities rather than differences between subjects (canonical topics) and topical objects: topics tend to bear grammatical marking, no matter what their grammatical function. Thus, our analysis does not relate the formal markedness of objects with their functional markedness, at least if the latter is assessed in terms of frequency or typicality (Dalrymple and Nikolaeva 2011: 220)

I am not contending that the primary function of object marking is to differentiate the object from the subject, but just that more complex morphology is used preferably for less frequent or more unexpected patterns. Further, the function of each coding device is different and its function is in line with what is generally expected from differential object marking and differential object indexing (Iemmolo 2011). Indexing – subject agreement and object clitics – is related to reference-tracking and all the arguments that may be indexed in Spanish have in common their semantic and pragmatic prominence and their high accessibility and topicality. In this sense, I have argued in the first pages of this article that subject and objects are core or central participants; from this point of view as well, it can be argued that indexing marks similarities between subject and topical objects. Note, however, that direct objects in Spanish may be indexed but most times they are not, so it is only those objects functionally marked, in the sense of being less frequent, that are also formally marked by clitic doubling. On the other hand, *a*-marking is used with certain kinds of objects, which are precisely those having referential and pragmatic properties closer to typical subject properties. Given that Spanish subjects are never marked by *a*, we can consider this marking to be an unequivocal index of non-subject. Once again, most objects do not take *a*, so that it is only those objects functionally marked, in the sense of being less frequent, that are also formally marked by *a*. Finally, the *le/lo* variation is specific to the object and it is not relevant whether it may serve to differentiate the object from the subject. What is relevant here is that the dative form *le* is more tied to the R slot and to the animacy properties typical of this slot, whereas accusative *lo* is more directly related with T and with those referential properties of P more typically associated with T. Once again, *lo* is a more frequent object marker, and *le* remains as a marked option valid for R and for objects with referential properties typically associated with R. In sum, Spanish data do not seem to support the view that "formally marked objects are just as frequent in languages with DOM as formally unmarked objects" (Dalrymple and Nikolaeva 2011: 166)

This view of unmarked objects in Spanish also contrasts with the concept of cardinal transitivity (Hopper and Thompson 1980), which requires a definite, highly individuated object. Along these lines, Næss defines a prototypical transitive clause as "one where the two participants are maximally semantically distinct in terms of the roles in the event described by the clause" (Næss 2007: 30). Focusing on the roles of the

participants and not on their inherent properties, she explains that the maximal contrast between agent and patient roles presupposes that both are individuated. She also points out that "prototypicality and markedness are two entirely distinct and independent concepts, and the one neither entails nor excludes the other" (Næss 2007: 25) and that the association between affectedness and a high degree of individuation can account for the DOM data (Næss 2004). From this point of view, accusative case marking should be considered a marker of a high degree of individuation and affectedness in objects, that is, a marker of prototypical transitivity.

Actually, the problem is not so much one of defining a transitive prototype as deciding if there is a uniform correlation between object marking and transitivity. This is independent of how individuation is represented in canonical transitivity. As far as Spanish is concerned, it is true that the preposition *a* and personal indexing always mark a highly individuated object. It is also true that in terms of pure frequency, definite objects are more frequent than indefinite objects but, still, definite human objects are marked. Individuation is a strong factor for object marking in Spanish. However, an individuated object does not imply in itself that the object is affected or that the subject of the clause is an agent. So, an individuated object does not imply high transitivity, and object marking in Spanish is more dependent on individuation than on transitivity. To state this more clearly, semantic transitivity in the sense of Hopper and Thompson (1980) does not directly correlate with object marking in Spanish two-participant clauses. Many formally transitive clauses, that is, V+N constructions, present a non-individuated object (*tener tiempo* 'have time', *dar paseos* 'take walks', *comer carne* 'eat meat', ...) and show very low semantic transitivity. They may be equivalent to simple intransitive verbs or to formally intransitive constructions of other languages such as object-incorporation, antipassives, etc. However, in Spanish such examples are not formally distinct from V+NP constructions with a definite individuated phrase (*tener el tiempo necesario* 'have enough time', *comer la carne* 'eat the meat'). As we move up the animacy and definiteness scale, we get highly individuated objects that tend to get marked coding (*Vio a su hermano* 'S/he saw her/his brother', *Mató a César* 'He killed Caesar') in most cases within semantically transitive clauses with two maximally distinct participant roles. But *a*-marking is also extended to many oblique arguments, whereby the probability of having *a*-marking, clitic doubling, and dative-*le* with objects increases as the subject is lower in animacy and agency and the event is less dynamic and less effective, resulting in low-transitivity *gustar*-type Subj–IO clauses (*A María le gustan los helados* 'María likes ice-cream') and in between many verbs alternating accusative and dative to varying degrees (*le/la preocupa, le/la molesta*, *le/la alegra*, etc...). In sum, zero-marking may be associated with low transitivity if the object is not individuated as distinct from the process denoted by the verb, and marked coding of the object, although in principle more associated with individuation and higher transitivity, may also serve to code independent oblique arguments and low-transitivity two-participant clauses.

7. Conclusion

Using corpus data, this study has explored the factors affecting the distribution of three related variable coding properties of objects in Spanish: marking, indexing, and case in pronominal clitics (including "*leísmo*"). Each formal property may independently vary to some extent, with many cases of alternation in the same context. For this reason, it

has been argued that Direct Object and Indirect Object are not two clear-cut grammatical relations in Spanish, but rather constitute a macro-function Object as a space where several gradient phenomena of variation take place.

It has been shown that all three types of variation are dependent on several factors that contribute to the agency potential, saliency, and accessibility of an argument. So, while *a*-marking is more dependent on inherent properties of the referent (animacy and definiteness), clitic doubling is more dependent on information status (topicality and accessibility), and clitic case depends mostly on animacy, gender, and process type and is subject to considerable dialectal variation.

Variable coding of participants in Spanish has been linked up with markedness, in the sense that less frequent options get more morphological marking, and also with the polarization between A and P arguments in transitive clauses in terms of inherent agency potential (animacy) and inherent topicality. As the P argument is expected to be less animate, less topical, and less accessible than the A argument, only the unexpected alternatives get either *a*-marking or doubling or favor the use of dative *le*. In addition, it has been argued that the frequency of variable coding in transitive and ditransitive clauses shows that any coding property of Spanish follows an indirective alignment type rather than a secundative or primary object alignment type, as the coding of the R argument is formally and functionally marked. Finally, it has been argued that variable coding of the object does not correlate directly with transitivity in Spanish.

Acknowledgments

Abbreviations

1/2/3 = first/second/third person; ACC = accusative; DAT = dative; DEF = definite; F = feminine; FUT = future; IPFV = imperfective; M = masculine; PFV = perfective; PL = plural; PST = past; SG = singular; SBJV = subjunctive

Appendix: References for the texts of the ARTHUS corpus

1IN: Olmo, Lauro & Pilar Enciso (eds.) 1987. *Teatro infantil, I*, Madrid: Antonio Machado.
2IN: Olmo, Lauro & Pilar Enciso (eds.) 1987. *Teatro infantil II*, Madrid: Antonio Machado.
1VO: *La Voz de Galicia,* 30 October 1991. A Coruña
2VO: *La Voz de Galicia,* 22 November 1991. A Coruña

3VO: *La Voz de Galicia,* 23 November 1991. A Coruña

AYE: Díaz, Jorge 1988. *Ayer, sin ir más lejos*. Madrid: Antonio Machado.

BAI: Barrenechea, Ana María (ed.) 1987. *El habla culta de la ciudad de Buenos Aires. Materiales para su estudio* (tomo 2). Buenos Aires: Instituto de Filología y Literaturas Hispánicas 'Dr. Amado Alonso' .

CAI: Buero Vallejo, Antonio 1981.*Caimán*. Madrid: Espasa-Calpe.

CAR: Colinas, A. 1986. *Larga carta a Francesca*. Barcelona: Seix Barral.

CIN: Reina, María Manuela 1989. *La cinta dorada*. Madrid: Antonio Machado.

COA: Fernán Gómez, Fernando 1987. *La coartada*. Madrid: Antonio Machado.

CRO: García Márquez, Gabriel 1987. *Crónica de una muerte anunciada*. Madrid: Mondadori.

DIE: Poniatowska, Elena 1987. *Querido Diego, te abraza Quiela y otros cuentos*. Madrid: Alianza/Era.

GLE: Cortázar, Julio 1981. *Queremos tanto a Glenda*, Madrid: Alfaguara, 4th edn.

HIS: Bioy Casares, Adolfo 1986. *Historias desaforadas*. Madrid: Alianza.

HOM: Salom, Jaime 1984. *Un hombre en la puerta*. Madrid: Preyson.

HOT: Gala, Antonio 1988. *El hotelito*. Madrid: Antonio Machado.

JOV: Aldecoa, Josefina 1986. *Porque éramos jóvenes*. Barcelona: Seix Barral.

LAB: Mendoza, Eduardo 1982. *El laberinto de las aceitunas*. Barcelona: Seix Barral.

LIN: Bunge, Mario 1983. *Lingüística y filosofía*. Barcelona: Ariel.

MAD: Esgueva, Manuel & Margarita Cantarero (eds.) 1981. *El habla de la ciudad de Madrid. Materiales para su estudio*. Madrid: CSIC.

MIR: Guelbenzu, José María 1987. *La mirada*. Madrid: Alianza.

MOR: Alonso de Santos, José Luis 1987. *Bajarse al moro*, 2nd ed. Madrid: A. Machado.

OCH: Diosdado, Ana 1990. *Los ochenta son nuestros*. Madrid: Antonio Machado.

PAI: Goytisolo, Juan 1982. *Paisajes después de la batalla*. Barcelona: Montesinos.

PAS: Reina, María Manuela 1988. *El pasajero de la noche*. Madrid: Antonio Machado.

RAT: Sánchez Ferlosio, Rafael 1986. *La homilía del ratón*. Madrid: El País.

SEV: Pineda, Miguel Angel de (ed.) 1983. *Sociolingüística andaluza 2: Material de encuestas para el estudio del habla urbana culta de Sevilla*. Sevilla: Universidad de Sevilla.

SON: Sampedro, José Luis 1985. *La sonrisa etrusca*. Madrid: Alfaguara.

SUR: García Morales, Adelaida 1985. *El sur (seguido de Bene)*. Barcelona: Anagrama.

TER: Martínez de Pisón, Ignacio 1988. *La ternura del dragón*, 3rd ed. Barcelona: Anagrama.

TIE: Paz, Octavio 1983. *Tiempo nublado*. Barcelona: Seix Barral.

USO: Martín Gaite, Carmen 1988. *Usos amorosos de la postguerra española,* 8th ed. Barcelona: Anagrama.

ZOR: Nieva, Francisco 1989. *Te quiero, zorra*. Madrid: Ed. Antonio Machado.

References

ADESSE. *Base de datos de verbos, alternancias de diátesis y esquemas sintáctico-semánticos del español*. Universidade de Vigo. http://adesse.uvigo.es/ (accessed on 25 May 2013).

Aissen, Judith. 2003. Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory* 21. 435–483.

Alarcos Llorach, Emilio. 1994. *Gramática de la lengua española*. Madrid: Espasa Calpe.

Andrews, Avery. 1985. The major functions of the noun phrase. In Timothy Shopen (ed.), *Language typology and syntactic description,* Vol. I, 62–154. Cambridge: Cambridge University Press.

Ariel, Mira. 1990. *Accessing noun-phrase antecedents*. London: Routledge.

Belloro, Valeria. 2007. Spanish clitic doubling: A study of the syntax-pragmatics interface. Buffalo, NY: University of New York doctoral dissertation.

Bickel, Balthasar. 2011. Grammatical relations typology. In Jae Jung Song (ed.), *The Oxford handbook of linguistic typology*, 399–444. Oxford: Oxford University Press.

Bickel, Balthasar & Johanna Nichols. 2009. Case marking and alignment. In Andrej Malchukov & Andrew Spencer (eds.), *The Oxford handbook of case*, 304–321. Oxford: Oxford University Press.

Bossong, Georg. 1998. Le marquage differentielle de l'objet dans les langues d'Europe. In Jack Feuillet (ed.), *Actance et valence dans les langues de l'Europe*, 193–258. Berlin: Mouton de Gruyter.

Bybee, Joan. 2006. From usage to grammar: The mind's response to repetition. *Language* 82(4). 711–733.

Chafe, Wallace. 1994. *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. Chicago: University of Chicago Press.

Company, Concepción. 2001. Multiple dative-marking grammaticalization: Spanish as a special kind of primary object language. *Studies in Language* 25(1). 1–47.

Comrie, Bernard. 1979. Definite and animate direct objets: a natural class. *Linguistica Silesiana* 3. 13–21.

Comrie, Bernard. 1989 [1981]. *Language universals and linguistic typology: Syntax and morphology*, 2nd ed. Chicago: University of Chicago Press.

Comrie, Bernard. 2012. Some argument-structure properties of "give" in the languages of Europe and Northern and Central Asia. In Pirkko Suihkonen, Bernard Comrie & Valery D. Solovyev (eds.), *Argument structure and grammatical relations: A crosslinguistic typology*, 17–35. Amsterdam: John Benjamins.

Croft, William. 1988. Agreement vs. case marking and direct objects. In Michael Barlow & Charles A. Ferguson (eds.), *Agreement in natural language: Approaches, theories, descriptions*, 159–179. Stanford: CSLI.

Croft, William. 2001. *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.

Dalrymple, Mary & Irina Nikolaeva. 2011. *Objects and information structure*. Cambridge: Cambridge University Press.

de Hoop, Helen & Andrej Malchukov. 2007. On fluid differential case marking: A bidirectional OT approach. *Lingua* 117(9). 1636–1656.

Delbecque, Nicole. 2001. A Construction Grammar approach to transitivity in Spanish. In Kristin Davidse & Béatrice Lamiroy (eds.), *The nominative & accusative and their counterparts*, 81–130. Amsterdam: John Benjamins.

DeMello, George. 2002. Leísmo in contemporary Spanish American educated speech. *Linguistics* 40(2). 261–283.

Demonte, Violeta. 1995. Dative alternation in Spanish. *Probus* 7. 5–30.

Dryer, Matthew. 1986. Primary objects, secondary objects and antidative. *Language* 62(4). 808–845.

Dryer, Matthew. 1997. Are grammatical relations universal? In Joan Bybee, John Haiman & Sandra A. Thompson (eds.), *Essays on language function and language type*, 115–143. Amsterdam: John Benjamins.

Du Bois, John W. 1987. The discourse basis of ergativity. *Language* 63. 805–855.

Du Bois, John W. 2003. Discourse and grammar. In Michael Tomasello (ed.), *The new psychology of language: Cognitive and functional approaches to language structure,* Vol. 2, 47–87. Mahwah, NJ: Lawrence Erlbaum.

Du Bois, John W., Lorraine E. Kumpf & William J. Ashby (eds.). 2003. *Preferred argument structure: Grammar as architecture for function*. Amsterdam: John Benjamins.

Dumitrescu, Domnita. 1997. El parámetro discursivo en la expresión del objeto directo lexical: Español madrileño vs español porteño. *Signo y seña* 7. 305–354.

Dumitrescu, Domnita. 1998. "A" personal, duplicación clítica y marcadez: español porteño vs. español madrileño. In Aengus Ward (ed.), *Actas del XII Congreso de la Asociación Internacional de Hispanistas*, 140–152. Birmingham: University of Birmingham.

Fernández Ordóñez, Inés. 1993. Leísmo, laísmo y loísmo: Estado de la cuestión. In Olga Fernández Soriano (ed.), *Los pronombres átonos*, 63–96. Madrid: Taurus.

Fernández Ordóñez, Inés. 1999. Leísmo, laísmo y loísmo. In Ignacio Bosque & Violeta Demonte (eds.), *Gramática descriptiva de la lengua española*, Vol. 1, 1317–1398. Madrid: Espasa.

Flores, Marcela. 2002. *Leísmo, laísmo y loísmo: Sus orígenes y evolución*. México: UNAM.

Flores, Marcela & Chantal Melis. 2007. El leísmo desde la perspectiva del "marcado diferencial del objeto". *Revista de Historia de la Lengua Española* 2. 83–107.

García, Erica C. 1975. *The role of the theory in linguistic analysis: The Spanish pronoun system*. Amsterdam: North-Holland.

García, Erica C. & Ricardo Otheguy. 1977. Dialect variation in leísmo: A semantic approach. In Ralph W. Fasold & Roger W. Shuy (eds.), *Studies in language variation*, 65–87. Washington: Georgetown University Press.

García García, Marco. 2007. Differential object marking with inanimate objects. In Georg A. Kaiser & Manuel Leonetti (eds.), *Proceedings of the Workshop "Definiteness, Specificity and Animacy in Ibero- Romance Languages"* (Arbeitspapier 122), 63–84.. Konstanz: Universität Konstanz.

García García, Marco. 2014. *Differentielle Objektmarkierung bei unbelebten Objekten im Spanischen*. Berlin: De Gruyter Mouton.

García-Miguel, José M. 1991. La duplicación de complemento directo e indirecto como concordancia. *Verba* 18. 375–410.

García-Miguel, José M. 1995. *Las relaciones gramaticales entre predicado y participantes*. Santiago de Compostela: Universidade de Santiago de Compostela.

García-Miguel, José M. 1999. Grammatical relations in Spanish triactant clauses. In Leon G. de Stadler & Christoff Eyrich (eds.), *Issues in cognitive linguistics*, 447–470. Berlin: Mouton de Gruyter.

Genetti, Carol. 1997. Object relations and dative case in Dolakha Newari. *Studies in Language* 21(1). 37–68.

Givón, Talmy (ed.). 1983. *Topic continuity in discourse: A quantitative cross-language study*. Amsterdam: John Benjamins.

Givón, Talmy. 1984. Direct object and dative shifting: Semantic and pragmatic case. In Frans Plank (ed.), *Objects: Towards a theory of grammatical relations*. London: Academic Press.

Givón, Talmy. 2001. *Syntax: An introduction*, Vol. 1. Amsterdam: John Benjamins.

Goldberg, Adele E. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Gutiérrez Ordóñez, Salvador. 1999. Los dativos. In Ignacio Bosque & Violeta Demonte (eds.), *Gramática descriptiva de la lengua española*, Vol. 1, 1855–1930. Madrid: Espasa.

Haiman, John. 1983. Iconic and economic motivation. *Language* 59(4). 781–819.

Haiman, John. 1985. *Natural syntax: Iconicity and erosion*. Cambridge: Cambridge University Press.

Halliday, M.A.K. 2004 [1984]. *An introduction to functional grammar*, 3rd ed. (revised by Christian M.I.M. Matthiessen). London: Arnold.

Haspelmath, Martin. 2005a. Argument marking in ditransitive alignment types. *Linguistic Discovery* 3(1). 1–21.

Haspelmath, Martin. 2005b. Ditransitive constructions: The verb "give". In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The World Atlas of Language Structures – WALS: An interactive cross-linguistic database for typological research*. Oxford: Oxford University Press.

Haspelmath, Martin. 2007. Ditransitive alignment splits and inverse alignment. *Functions of Language* 14(1). 79–102.

Haspelmath, Martin. 2008. Frequency vs. iconicity in explaining grammatical asymmetries. *Cognitive Linguistics* 19(1). 1–33.

Haspelmath, Martin. 2011. On S, A, P, T, and R as comparative concepts for alignment typology. *Linguistic Typology* 15(3). 535–567.

Haspelmath, Martin. 2013. Argument indexing: A conceptual framework for the syntax of bound person forms. In Dik Bakker & Martin Haspelmath (eds.), *Languages across boundaries: Studies in memory of Anna Siewierska*, 197–226. Berlin: De Gruyter Mouton.

Heine, Bernd & Christa König. 2010. On the linear order of ditransitive objects. *Language Sciences* 32(1). 87–131.

Hopper, Paul J. & Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56(2). 251–299.

Iemmolo, Giorgio. 2010. Topicality and differential object marking: Evidence from Romance and beyond. *Studies in Language* 34(2). 239–272.

Iemmolo, Giorgio. 2011. Towards a typological study of differential object marking and differential object indexing. Pavia: Università degli Studi di Pavia doctoral dissertation.

Kittilä, Seppo. 2006. Object-, animacy- and role-based strategies: A typology of object marking. *Studies in Language* 30(1). 1–32.

Klein-Andreu, Flora. 1981. Distintos sistemas de empleo de "le", "la", "lo". Perspectiva sincrónica, diacrónica y sociolingüística. *Thesaurus* 36. 284–304.

Klein-Andreu, Flora. 2000. *Variación actual y evolución histórica: Los clíticos le/s, la/s, lo/s*. München: Lincom Europa.

Klein, Udo & Peter de Swart. 2011. Case and referential properties. *Lingua* 121(1). 3–19.

Kliffer, Michael D. 1984. Personal "a", kinesis and individuation. In Philip Baldi (ed.), *Papers from the XIIth Linguistic Symposium on Romance Languages*, 195–215. Amsterdam: John Benjamins.

Laca, Brenda. 1987. Sobre el uso del acusativo preposicional en español. *Romanistisches Jahrbuch* 38. 290–312.

Laca, Brenda. 2006. El objeto directo: La marcación preposicional. In Concepción Company (ed.), *Sintaxis histórica de la lengua española. Primera parte: La frase verbal*, Vol. 1, 423–475. México: Fondo de Cultura Economica.

Lazard, Gilbert. 1984. Actance variations and categories of the object. In Frans Plank (ed.), *Objects: Towards a theory of grammatical relations*, 269–292. London: Academic Press.

Lazard, Gilbert. 2002. Transitivity revisited as an example of a more strict approach in typological research. *Folia Linguistica* 36(3–4). 141–190.

Leonetti, Manuel. 2004. Specificity and differential object marking in Spanish. *Catalan Journal of Linguistics* 3. 75–114.

Leonetti, Manuel. 2008. Specificity in clitic doubling and in differential object marking. *Probus* 20(1). 33–66.

Malchukov, Andrej, Martin Haspelmath & Bernard Comrie. 2010. Ditransitive constructions: A typological overview. *Studies in ditransitive constructions: A comparative handbook*, 1–63. Berlin: Walter de Gruyter.

Malchukov, Andrej & Peter de Swart. 2009. Differential case marking and actancy variations. In Andrej Malchukov & Andrew Spencer (eds.), *The Oxford handbook of case*, 339–355. Oxford: Oxford University Press.

Maldonado, Ricardo. 2002. Objective and subjective datives. *Cognitive Linguistics* 13(1). 1–65.

Miles, Cecil & Romelia Arciniegas. 1983. Tener a: A Spanish myth. *Hispania* 66(1). 84–87.

Næss, Åshild. 2004. What markedness marks: The markedness problem with direct objects. *Lingua* 114(9–10). 1186–1212.

Næss, Åshild. 2007. *Prototypical transitivity*. Amsterdam: John Benjamins.

Newman, John. 1996. *Give: A cognitive linguistic study*. Berlin: Mouton de Gruyter.

Nichols, Johanna. 1983. On direct and oblique cases. *Berkeley Linguistic Society* 9. 170–192.

Nichols, Johanna. 2011. *Ingush Grammar*. Berkeley: University of California Press.

Nishida, Chiyo. 2012. A corpus study of Mexican Spanish three-participant constructions with and without clitic doubling. *Linguistic Discovery* 10(3), 208-240. doi:10.1349/PS1.1537-0852.A.422..

Pensado, Carmen (ed.). 1995. *El complemento directo preposicional*. Madrid: Visor.

RAE & AALE. 2009. *Nueva gramática de la lengua española*, 3 vols. Madrid: Espasa.

Roegiest, Eugeen. 1990. La tipología sintáctica del objeto transitivo en español. *Verba* 17. 239–248.

Romero, Juan. 2013. Accusative datives in Spanish. In Beatriz Fernández & Ricardo Etxepare (eds.), *Variation in datives: A micro-comparative perspective*. New York: Oxford University Press.

Siewierska, Anna. 2003. Person agreement and the determination of alignment. *Transactions of the Philological Society* 101(2). 339–370.

Siewierska, Anna. 2004. *Person*. Cambridge: Cambridge University Press.

Silverstein, Michael. 1976. Hierarchies of features and ergativity. In R.M.W. Dixon (ed.), *Grammatical categories in Australian languages*, 112–171. Canberra: Australian Institute of Aboriginal Studies.

Thompson, Sandra A. 1997. Discourse motivations for the core-oblique distinction as a language universal. In Akio Kamio (ed.), *Directions in functional linguistics*, 59–82. Amsterdam: John Benjamins.

Torrego Salcedo, Esther. 1999. El complemento directo preposicional. In Ignacio Bosque & Violeta Demonte (eds.), *Gramática descriptiva de la lengua española*, Vol. 2, 1779–1805. Madrid: Espasa.

Vaamonde, Gael. 2011. La alternancia posesiva con nombres de partes del cuerpo: Un estudio descriptivo del español a partir de datos de corpus. Vigo: Universidade de Vigo doctoral dissertation.

Vázquez Rozas, Victoria. 1995. *El complemento indirecto en español*. Santiago de Compostela: Universidade de Santiago de Compostela.

Vázquez Rozas, Victoria. 2006. Gustar-type verbs. In Clancy J. Clements & Jiyoung Yoon (eds.), *Functional approaches to Spanish syntax: Lexical semantics, discourse and transitivity*, 80–114. New York: Palgrave Macmillan.

von Heusinger, Klaus. 2008. Verbal semantics and the diachronic development of DOM in Spanish. *Probus* 20(1). 1–31.

von Heusinger, Klaus & Georg A. Kaiser. 2011. Affectedness and differential object marking in Spanish. *Morphology* 21(3–4). 593–617.

Witzlack-Makarevich, Alena. 2010. Typological variation in grammatical relations. Leipzig: Universität Leipzig doctoral dissertation.