

## 8. The encoding of the manipulee in Spanish: the competition between dative and accusative

BOJANA MIKELENIĆ, METKA BEZLAJ

Keywords: *Spanish; clitic variation; manipulation verbs; leísmo*

In this study, we analyse the alternation between dative and accusative clitics in nine manipulation verbs in Spanish. We use a reference corpus *CORPES XXI* to extract all examples of these verbs in constructions with a prepositional object in order to observe the current status of clitic alternation, or *leísmo*, and its geographical distribution in different Spanish-speaking areas. We look into the variables of type of clause, gender, and number of the clitics' referent as well. The analysis has shown that the process of replacing the etymological dative with the accusative with these verbs is uneven, that is, it varies from area to area and from verb to verb.

*Original scientific paper*

### 1. Introduction

The present study deals with selected Spanish verbs that show change in their government, i.e. we focus on a certain number of verbs that can present their pronominal object as accusative or dative. The literature review (e.g. Fernández-Ordóñez 1999) showed that these verbs are fairly frequent in Spanish. However, due to practical reasons, as well as the homogeneity of our analysis, we decided to consider only those that can be syntactically and semantically categorized into a coherent verb group. More specifically, we looked at verbs such as *animar* 'to encourage', *autorizar* 'to authorize', *forzar* 'to force', etc. that could be included in the so-called manipulation verbs (Givón 2001: 151–152). The reinterpretation of the clitic with these verbs is not consistent and it depends not only on dialectological factors but can vary from verb to verb. The data was gathered from a corpus: *Corpus del Español del Siglo XXI – CORPES XXI*, which is representative of the Spanish

language spoken in Spain and in Latin America, so it allows for diatopic analyses.

In that sense, the overall goal of our work is to examine the current status of clitic alternation within this group of verbs, which is a topic that is also related to the relationship between direct and indirect object in Spanish. The aim of our analysis is three-fold: firstly, we want to examine which third-person clitics are used most frequently with manipulation verbs selected for the study; secondly, our goal is to describe various types of this usage, especially with respect to different characteristics of the object (such as case, number, gender, person, and type of complement clause). Finally, our work also sets out to account for substantial regional variation of the phenomenon, which is, as already stated, an important factor when it comes to dealing with this topic. In other words, the study includes language variety as an important variable, and we investigate how different Spanish-speaking countries compare in this regard.

## 2. Theoretical background

The behaviour of verbal clitics in Romance is different than that of a noun phrase in its place; essentially, their distribution is not the same. As Halpern (2017) shows on a French example, but the same can also be applied to Spanish, a clitic precedes the verb it modifies, while a nonclitic object follows it. As a consequence, these are not the so-called *simple clitics*, like reduced auxiliaries and pronouns in English that are basically just accentless words but could be named *special clitics* (Halpern 2017). There is substantial research being done regarding the similarities and differences of these clitics with the inflectional affixes (agreement morphology) and their status in the argument structure of a verb (cf. Halpern 2017; Mayer 2017; Spencer & Luís 2012). This also connects to the relationship between direct and indirect object in general.

In his paper on coding properties of Spanish objects, García-Miguel (2015) investigates clitic doubling, *a*-marking, and clitic case as a means of understanding the boundaries between direct and indirect object. Clitic doubling (or object doubling), a frequent topic in studies about clitics in Spanish, is possible only when the object (direct or indirect) is introduced by the preposition *a* (*a*-marking; see section 3 for more details). The author considers these properties as variables and observes their frequency

in a corpus and their distribution in different types of constructions. He concludes that, even though all three variables are dependent on several factors (animacy and definiteness of the referent for *a*-marking, topicality and accessibility for clitic doubling, animacy, gender, and process type for clitic case), considerable dialectal variation has to be taken into account for clitic case.<sup>106</sup>

This suggests that one of the principal means of distinguishing between the two objects loses its strength and the boundaries between them become unclear. On the one hand, De Mello (2002) claims that *leísmo*, the use of the clitic *le* for direct object, does not exist in Latin America and that the verbs that alternate between usage of dative and accusative actually alternate between usage of indirect and direct object, respectively. Even though De Mello considers only Latin American data, the line of reasoning is obvious: dative marks the indirect and accusative the direct object. On the other hand, Company Company (2001, 2003) uses diachronic data to identify the switch in case marking, going so far as to argue that the dative promotion displaces the accusative as the usual and more frequent marking of a direct object. A third route would be to accept the murky boundaries and view objecthood as a continuum, with direct and indirect object on either side of it, and with only limited generalizations possible (García-Miguel 2015).

In the following sections, we will be looking more closely into the clitic case. The clitic variation in Spanish is described through three different phenomena: *leísmo*, which we have already mentioned, but also *laísmo*, and *loísmo*. However, to understand them, we first need to introduce the clitic system in Spanish.

## 2.1. The clitic system in Spanish and clitic variation

In Spanish, like in some other Romance languages, first and second person (singular and plural) clitics show syncretism for case. However, case distinctions are present among third person (singular and plural) clitics. Gender is clearly marked in accusative forms, while there is no gender distinction among third person dative clitics (Table 1).

---

<sup>106</sup> See Giuliani (*this volume*, Chapter 7) for an account of the encoding of the third argument of the Latin verb *mittō* 'send' based on the animacy of the referents.

Table 1. A part of personal pronoun paradigm in Spanish<sup>107</sup>

|       | Strong subject pronouns    | Strong object pronouns | Accusative clitics | Dative clitics |
|-------|----------------------------|------------------------|--------------------|----------------|
| 1SG   | <i>yo</i>                  | <i>mí</i>              |                    | <i>me</i>      |
| 2SG   | <i>tú</i>                  | <i>ti</i>              |                    | <i>te</i>      |
| 3M.SG | <i>él</i>                  |                        | <i>lo</i>          | <i>le</i>      |
| 3F.SG | <i>ella</i>                |                        | <i>la</i>          |                |
| 1PL   | <i>nosotros / nosotras</i> |                        |                    | <i>nos</i>     |
| 2PL   | <i>vosotros / vosotras</i> |                        |                    | <i>os</i>      |
| 3M.PL | <i>ellos</i>               |                        | <i>los</i>         | <i>les</i>     |
| 3F.PL | <i>ellas</i>               |                        | <i>las</i>         |                |

The forms *lo*, *la*, *los*, *las* typically correspond to the direct object and *le*, *les* to the indirect one, a usage known as *etymological*, because it follows the case system of Latin demonstratives (Fernández-Ordóñez 1999: 1319). Nevertheless, in present-day Spanish there are various uses that diverge from this original accusative/dative distinction. This variation can be accounted for by three phenomena mentioned before: *leísmo*, *laísmo*, and *loísmo*. *Leísmo* consists of the use of dative clitics for the direct object (s. Vinja 1998: 149; Alarcos Llorach 1999: 202). For example, in the sentence (1) the clitic *le* is used to double the direct object *A Pedro*. On the other hand, the use of *lo* for indirect object is standardly called *loísmo*,<sup>108</sup> and this usage can be seen in the example (2), where *lo* is used instead of the indirect *le*. Finally, *laísmo* refers to the usage of *la* instead of *le*, as in (3), where *la* refers to the indirect object.

- (1) *A Pedro le vi ayer.*  
 ‘Peter, I saw [**him**] yesterday.’
- (2) *Lo dije que se fuera.*  
 ‘I told **him** to go away.’

<sup>107</sup> This table only shows the part of this paradigm that interests us here; for a more detailed account, including regional variations, cf. *Nueva gramática de la lengua española* – NGL (2009: §16.3).

<sup>108</sup> Fernández-Ordóñez (1999: 1320) mentions that for some authors *loísmo* means the usage of *lo* for the direct object as opposed to *le*. We will not be using the said term in this sense.

(3) *La di un libro.*‘I gave **her** a book.’

Out of these three phenomena, *leísmo* is the most widespread and can be found both in Spain and in Latin America, while *loísmo* and *laísmo* are restricted only to certain areas of Spain (NGLE 2009: §16.10; Fernández-Ordóñez 1999: 1322). This is the reason why we have chosen *leísmo* as the central focus of our research, even though it can be difficult to separate the three, as they have frequently been studied together.

Fernández-Ordóñez (1999, 1993) gives an overview of clitic variation in Spanish, summarizing the most important hypotheses about its origins and the research done during the 19<sup>th</sup> and 20<sup>th</sup> century. She explains that the authors contrast *etymological* and *non-etymological* or *anti-etymological* usage, depending on whether it follows the regimen of the verb in question in Latin or not (Fernández-Ordóñez 1999: 1320). The first traditional hypothesis that was proposed to explain these phenomena is based on the idea that case was eliminated in favour of gender; in other words, the syncretism between first person and second person clitics was extended to third person clitics. In that sense, Fernández Ramírez (1987: 43) and Lapesa (1968: 541) see the origins of clitic variation in a tendency to mirror the demonstrative pronouns model (*este, esta, esto*), where *le* would be used for masculine, *la* for feminine, and *lo* for neutral gender. The second hypothesis suggests that Spanish tends to differentiate between personal and non-personal entities, and this favoured the emergence of *leísmo* (Fernández Ramírez 1964: 280). However, these explanations are not sufficient to account for different facets of these phenomena. For example, the first hypothesis cannot explain why *leísmo* occurs more frequently with personal direct objects, while the second cannot account for examples of *leísmo* with inanimate non-personal objects, nor origins of *laísmo* and *loísmo* (Fernández-Ordóñez 1999: 1321–1322).

This prompted linguists to consider dialectological differences in the Spanish-speaking world, which proved to be a turning point in the interpretation of clitic variation (Fernández-Ordóñez 1999: 1322). Working independently on Spanish spoken data, García González (1981) and Klein-Andreu (1993) concluded that continuity and discontinuity of the object's referent, or it being countable or uncountable, was as important for the choice of clitic as gender. This system is known as *referential*, as opposed to the *etymological*. In light of this, Fernández-Ordóñez (1999: 1322) concludes that we should distinguish between three situations: 1) *leísmo* in the areas that keep the case distinction, 2) *leísmo* and perhaps *loísmo* in the areas where

Spanish comes in contact with non-Indo-European languages (Quechua, Aymara, Guaraní, and Basque), and 3) clitic confusions (*leísmo*, *laísmo*, and *loísmo*) in the areas where the referential system is at work. Concerning the first situation, Fernández-Ordóñez (1999: 1323–1339) distinguishes different types of verbs that alternate the clitics and have been described in literature as cases of *leísmo*, but do not, strictly speaking, fall into that category. With these verbs dative is not being used where the accusative should be, but on the contrary, these were originally intransitive constructions that are presently being transformed into transitive ones (Fernández-Ordóñez 1999: 1323). In this study we consider this type of verbs: we will explain more about them in section 3.

It is also worth mentioning that there are different types of *leísmo*, depending on animacy, gender and number of clitic's referent (Fernández-Ordóñez 1999: 1319; NGLÉ 2009: §16.8c). The most frequent type of *leísmo* has the clitic referring to a singular masculine and personal direct object (4). It is followed by the singular and masculine clitic which refers to an inanimate object (5), that is in turn followed by two types of plural *leísmo*: clitic in masculine gender referring to a person (6) and clitic in masculine gender referring to an inanimate object (7). Finally, the rarest of all, *leísmo* with a feminine referent, usually a person, both singular (8) and plural (9).

- (4) *Mario viene esta noche, así que le llevaré a cenar.*  
 'Mario is coming this evening, so I'll take **him** to dinner.'
- (5) *Mi madre me envió un paquete y le recogeré mañana.*  
 'My mother sent me a package and I'll pick **it** up tomorrow.'
- (6) *Hace tiempo que no veo a tus hijos, pero Juan les acaba de ver.*  
 'It's been some time since I've seen your children, but Juan just saw **them**.'
- (7) *Compré los pasteles y les probé enseguida.*  
 'I've bought the cakes and tried **them** immediately.'
- (8) *¿Conoces a Marta? No, no le conozco.*  
 'Do you know Marta? No, I don't know **her**.'
- (9) *Las chicas están cantando, les oigo de mi cuarto.*  
 'The girls are singing, I hear **them** from my room.'

A separate subtype of *leísmo* where the referent is a person is the so-called *leísmo de cortesía*, where the clitic *le/les* is used for courtesy address to the second person, that is, where the referent of the object is not the third

person (NGLE 2009: §16.8d). In these cases, the dative clitic agrees with second person pronouns *usted/ustedes*, as in *Yo le conozco a usted* ('I know **you**'). In our research, it was important to tag these examples manually so as to exclude them from the analysis.

## 2.2. Previous research

Apart from numerous studies on written data (s. Fernández-Ordóñez 1993), there are various studies looking into *leísmo* in Spain (Klein-Andreu 1993; Díaz Montesinos 2017), and Latin America (DeMello 2002), which analyse speech samples or are corpus-based. Due to the difficulties in collecting and working with spoken data, those studies tend to be limited to a specific city or region. For example, during three months of fieldwork, Klein-Andreu (1993) collected 150 hours of conversation in five Castilian provinces (Valladolid, Burgos, Soria, Logroño, and Toledo). This became somewhat easier with the development of language resources and technologies, but oral corpora are scarce to this day and tend to be specialized and collected specifically for a certain investigation or project. In this sense, for his study of *leísmo* in Málaga, Díaz Montesinos (2017) used a corpus of 72 semi-structured interviews and DeMello (2002) a corpus of approximately 1.3 million words of educated speech samples from ten Latin American capitals.

One of the important questions is also how different types of this phenomenon can be classified and whether all of them could even be considered *leísmo* (Fernández-Ordóñez 1999; DeMello 2002). As mentioned before, among them are the verbs that interest us: manipulation verbs that can alternate between dative and accusative clitics.

In his study of four of these verbs (*enseñar* 'to teach', *ayudar* 'to help', *obligar* 'to make someone do something', and *invitar* 'to invite'), Rivas (2013: 156) explains that the question of *leísmo* with respect to manipulation verbs has not received a lot of attention in literature. Previous research seems to show that most of the Peninsular Spanish prefers dative clitics with these verbs, while in Andalusia, Canary Islands, and Latin America the accusative clitic is more common (Roegiest 1990: 243;<sup>109</sup> Fernández-Ordóñez 1999: 1326). However, Díaz Montesinos (2017) concludes for Málaga, a city in Andalusia, that with some of these verbs (*obligar* 'to make someone do something',

---

<sup>109</sup> Roegiest (1990) analysed written data from twentieth-century literary works in the Peninsular Spanish.

*animar* ‘to encourage’, *ayudar* ‘to help’ and *enseñar* ‘to teach’) dative is more frequent. In NGLE (2009: §16.9o) it is noted that these verbs show a considerable dialectal variation in Spain, the Antilles, Mexico, and a part of Central America, while for the rest of the areas the accusative form is more frequent. On the other hand, Rivas (2013: 162), who used *Corpus del español*<sup>110</sup> to extract approximately 5 million words of spoken Spanish, observed that the Peninsular Spanish prefers the original dative marker, while the Latin American Spanish more often opts for the accusative. However, his (Rivas 2013: 162) research indicates that this can vary from verb to verb, since in present-day Spanish verbs *obligar* ‘to make someone do something’ and *invitar* ‘to invite’ prefer the accusative form, while *enseñar* ‘to teach’ and *ayudar* ‘to help’ favour the presence of the dative clitic. In our analysis, we will be returning to these previous studies to compare the results for each verb in more detail.

To the best of our knowledge, there are no studies of *leísmo* on a large reference corpus.<sup>111</sup> Even though the corpus we used, *Corpus del Español del Siglo XXI – CORPES XXI*,<sup>112</sup> contains mostly written data, it is comprised of all different kinds of texts that are representative of the usage of Spanish language. Our study is based on the version 0.94 of the corpus which was published in July 2021 and contains more than 327 000 documents and 350 million words.<sup>113</sup> As already noted, the *CORPES XXI* covers mostly written registers (90%), but it also includes oral data (10%). The written registers in the corpus are categorized as books (40%), press (40%), Internet sources (7,5%) and miscellaneous sources (2,5%). When it comes to geographical distribution, 70% of the texts are from Latin America and 30% were produced in Spain. The material from Latin America is classified according to different linguistic regions (see Table 4). The version which was used in the study

---

<sup>110</sup> The structure of the corpus, as well as access to its search engine, can be found here: <https://www.corpusdelespanol.org>.

<sup>111</sup> García-Miguel (2015) uses the *ARTHUS* corpus in his research. This corpus contains almost 1,5 million words and, while it does contain texts of different types and origin, it is quite small. More information about this corpus can be found here: <https://www.bds.usc.es/corpus.html>.

<sup>112</sup> The detailed structure of the corpus, as well as access to its search engine, can be found here: <https://www.rae.es/banco-de-datos/corpes-xxi>.

<sup>113</sup> Since then, two updated versions of the corpus have been published. The current version (the version 1.0) is available from June 2023 and it represents an increase of more than 45 million forms compared to the version that was used in this study.

covers the period from the beginning of this century up to the year 2020. Furthermore, it is worth mentioning that all the texts in the corpus are lemmatized and morphosyntactically annotated. For these reasons, we believe the *CORPES XXI* to be the state-of-the-art resource to examine the current status of clitic alternation within this group of verbs.

### 3. Data and methodology

The possible verb candidates for the study were first gathered from Fernández-Ordóñez's (1999) two interrelated sections about clitic reinterpretation with different sets of verbs (1999: 1325–1328, 1330–1335).<sup>114</sup> In these sections, the author deals with a number of suasive verbs and verbs of influence, as well as other verbs that show a change in government. She (1999: 1330) explains that all of these verbs governed the dative case in Medieval Spanish and in the present-day Spanish they are in the process of replacing it with the accusative. In this way 32 verbs were selected and further analysed in order to see whether they matched the criteria for inclusion in the group of manipulation verbs.<sup>115</sup>

For a verb to be considered a manipulation verb, we relied on Givón's criteria (2001: 151–152). Semantically, these verbs typically include a person who manipulates (or tries to manipulate) the behaviour of another human, the so-called manipulee. The agent of the complement clause should be coreferential with the manipulee from the main clause and the complement clause expresses the target event which should be performed by the manipulee.

Syntactically, the manipulee of the main clause can be either the direct or indirect object and it is also the coreferential-subject of the complement clause, where it is left unexpressed. The verbs we have chosen all follow the same pattern: *verb + object + preposition + clause* (which can appear in a

---

<sup>114</sup> It is worth noting that the author also mentions many verbs for which alternation causes a change in structure and/or meaning. She calls these the cases of “apparent *leísmo*”. Such verbs were omitted from the analysis.

<sup>115</sup> As mentioned in the previous section, morphosyntactic variation in the complementation of four of these verbs (i.e. *enseñar*, *ayudar*, *obligar*, and *invitar*) has already received substantial interest in Rivas (2013). However, our study is based on a different set of data and more attention is paid to dialectological differences.

nominalized or a non-finite form, as shown below). Moreover, it is crucial to highlight the importance of the preposition in this structure, since it seems that its presence favours the reinterpretation of the clitic, i.e. those manipulation verbs which require a preposition tend to show clitic variation (Fernández-Ordóñez 1999: 1333).

Because of the nature of these verbs, the object expressing the manipulee is [+animate] and [+human], while the prepositional object expressing the target event to be performed by the manipulee is necessarily inanimate. This and the fact that in these constructions the animated object pronominalized in dative can be transformed into the subject of a passive clause (usually reserved only for direct objects) could explain why the original dative is being substituted for the accusative (Fernández-Ordóñez 1999: 1333–1334). The passivization of the sentence (10a) (where the object *a Juan* is doubled by the dative clitic *le*) can be found in (10b).

- (10) a. *A Juan le invitaron a la conferencia.*  
 DAT John CL.3SG invite.PST.3PL to DET.F.SG conference(F).SG  
 ‘They invited John [him] to the conference.’
- b. *Juan fue invitado a la conferencia.*  
 John be.PST.3SG invite.PTCP to DET.F.SG conference(F).SG  
 ‘John was invited to the conference.’

Each of the 32 verbs were examined semantically and syntactically and, in the end, only 9 verbs (*animar* ‘to encourage’, *autorizar* ‘to authorize’, *convencer* ‘to convince’, *forzar* ‘to force’, *incitar* ‘to incite’, *invitar* ‘to invite’, *obligar* ‘to make someone do something’, *ayudar* ‘to help’, and *enseñar* ‘to teach’) were included in the study. For example, the verb *animar* ‘to encourage’ was evaluated on the basis of the example (11a) where Juan is the manipulee and María is the one who tries to manipulate his behaviour. In addition, Juan is the object of the main verb *animar* ‘to encourage’, but at the same time it is the subject of the complement clause, i.e. of the verb *estudiar* ‘to study’. The complement clause clearly represents the target action that should be done by Juan.

- (11) a. *María animó a Juan a estudiar más.*  
 ‘Mary encouraged **John** to study harder.’
- b. *María lo / le animó a estudiar más.*  
 ‘Mary encouraged **him** to study harder.’

It is also important to clarify that in Spanish the syncretic form *a* (as in *a Juan* in (11a)) can introduce both direct objects differentially and indirect objects obligatorily. In other words, when not presented as clitics, all indirect objects are necessarily introduced by *a*, while only direct objects that are [+human], [+animate] and [+definite] are marked by differential object marking, which is homonymous of the dative marker *a* (Mayer 2017: 5). In consequence, the change in the government of the manipulation verbs selected for our study can only be appreciated when their objects are expressed through 3<sup>rd</sup> person clitic pronouns which are marked for case (*lo* for the accusative and *le* for the dative, as in (11b)).

Furthermore, the prepositional object is also introduced by the preposition *a* (except in the case of *convencer + de* ‘to convince’) and it can present itself as a finite noun clause<sup>116</sup> (as in (12)), non-finite noun clause in the infinitive form (as in (13)) or a noun phrase (as in (14)).

- (12) *María animó a Juan a que cantara una canción.*  
 Mary encourage.PST.3SG DAT/ACC John to that sing.SBJV.3SG INDF.F.SG song(F).SG  
 ‘Mary encouraged John to sing a song.’

- (13) *María obligó a Juan a venir a su fiesta.*  
 Mary make.PST.3SG DAT/ACC John to come.INF to POSS.3SG party.FSG  
 ‘Mary made John come to her party.’

- (14) *María invitó a Juan a un café.*  
 Mary invite.PST.3SG DAT/ACC John to INDF.M.SG coffee(M).SG  
 ‘Mary invited John to get a cup of coffee.’

On the other hand, some verbs were excluded from the analysis because they lacked the necessary syntactic characteristic, namely the preposition (e.g. *prohibir* ‘to forbid’ or *ordenar* ‘to order’ as in (15) and (16)), while others did not fit in the category semantically (e.g. *acusar* ‘to accuse’ or *reñir* ‘to scold’ as in (17) and (18), where the subject of the main clause does not seem to manipulate another person’s behaviour nor the complement clause could be seen as coding the target event to be performed by the manipulee).

<sup>116</sup>In Spanish grammaticography, it is customary to consider the element introducing these clauses as a preposition (e.g. NGLÉ 2009: §36.1e), even though other grammatical traditions may see it as a complementizer.

- (15) *María prohibió a Juan que estudiara más.*  
‘Mary forbade John to study harder.’
- (16) *María ordenó a Juan estudiar más.*  
‘Mary ordered John to study harder.’
- (17) *María acusó a Juan de robar los libros.*  
‘Mary accused John of stealing the books.’
- (18) *María riñó a Juan por robar los libros.*  
‘Mary scolded John for stealing the books.’

Subsequently, we proceeded to look for the selected verbs in a reference corpus for Spanish language, *Corpus del Español del Siglo XXI – CORPES XXI*, which contains more than 350 million words and is representative of the modern 21st century Spanish language spoken in Spain and in Latin America (see section 2.2. for more details about the corpus). The data was collected in the following manner: We looked for the lemma of each verb and its required preposition in an interval 2 to the right followed by either the accusative clitic (i.e. the lemma *lo*) or the dative clitic (i.e. the lemma *le*)<sup>117</sup> in an interval 2 to the left and right.<sup>118</sup> This procedure was selected in order to obtain separate results for *verb + the dative clitic* and *verb + the accusative clitic*. Only the third person clitics (i.e. *lo, la, los, las*) were chosen due to the fact that, as we have seen previously, clitics for first and second person (both singular and plural) show syncretism for case in Spanish. All the concordances were then downloaded, as well as the data about the distribution by region and country for each verb, which is provided automatically by the corpus search engine.

In the next step, a percentage of dative and accusative examples for each verb was calculated from the total number of concordances. However, since we set out to explore different types of clitic alternation according to various above-mentioned variables that could not be searched for automatically in the corpus, a more detailed manual annotation and analysis of the examples

<sup>117</sup> Even though *lo* and *le* are a part of the same paradigm, they are in this corpus lemmatized independently. This is why we refer to them as two lemmas.

<sup>118</sup> The interval for the preposition was chosen on the basis of previous studies on prepositional complementation (Mikelenić 2020) and the preliminary analysis of a smaller set of random concordances showed that accusative and dative clitics are highly unlikely to appear further than two spaces away from the verb (either to the left or to the right).

was warranted. In other words, the annotation of each example included (apart from the case and number tags that were obtained directly from the corpus) the following categories: person (we annotated whether the referent of the clitic was a third person or the second person addressed by means of *usted* or *ustedes*, a phenomenon called *leísmo de cortesía*), gender (we distinguished between masculine and feminine forms whenever it was possible to establish this distinction based on the referent), and type of clause (we marked each complement clause as infinitival, nominalized or finite). This annotation was performed on 15% of random examples for each clitic type and for each verb.<sup>119</sup>

## 4. Results and analysis

We begin with the overall distribution of the clitic alternation with these verbs, to pass onto a general geographical distribution, where we distinguish between frequency of usage of each clitic for each verb in Spain and in Latin America as a whole. We then look into a more detailed diatopic analysis of five most frequent verbs. Lastly, we examine other manually annotated variables, namely type of clause, number, and gender.

### 4.1. Overall distribution of the accusative/dative alternation

Table 2 shows the first set of results obtained from the corpus. The verbs selected for the analysis appeared in 34,881 sentences in total and their overall distribution is shown in Table 2. The verbs are listed in descending order according to their frequency with both the accusative and the dative clitic, and each verb is accompanied by the corresponding preposition. The last two columns display the percentage of these examples that appear with the accusative and the dative clitic. This data already seems to confirm that the process of replacing the etymological dative with the accusative is not nearly finished and, what is more, these changes do not appear to affect all the verbs in the same way. Some verbs, like *ayudar* ‘to help’, *animar* ‘to encourage’ and *autorizar* ‘to authorize’, hesitate between the two possibilities,

---

<sup>119</sup> Since the concordances for each verb were not lacking in number (for verbs like *ayudar* we found almost 10.000 examples, as we will show later on), this percentage was chosen as the highest number of sentences that we could feasibly annotate manually.

i.e. they are used as frequently with the accusative and the dative. On the other hand, the rest of the verbs show a more obvious preference for one of the options, in most cases the accusative object (with the clear exception of the verb *enseñar* ‘to teach’, which favours the dative). For the verbs *obligar* ‘to make someone do something’, *invitar* ‘to invite’, *enseñar* ‘to teach’, and *ayudar* ‘to help’, the results are in line with what Rivas (2013) concluded: The first two verbs occur more often with the accusative, while the last two prefer the dative clitic, even though with *ayudar* ‘to help’ both options are equally frequent. However, as we have mentioned in the theoretical introduction to our paper, the status of these verbs cannot be accounted for without registering their diatopic differences, which are presented in the following section.

Table 2. Overall distribution of the accusative/dative alternation with selected manipulation verbs

| verb + preposition                                 | total number of examples | with accusative <i>lo</i> | with dative <i>le</i> |
|--|--------------------------|---------------------------|-----------------------|
| <i>ayudar a</i><br>‘to help’                       | 9,943                    | 45.65%                    | 54.35%                |
| <i>obligar a</i><br>‘to make someone do something’ | 9,363                    | 69.34%                    | 30.66%                |
| <i>invitar a</i><br>‘to invite’                    | 6,906                    | 76.74%                    | 23.26%                |
| <i>enseñar a</i><br>‘to teach’                     | 3,458                    | 14.11%                    | 85.89%                |
| <i>convencer de</i><br>‘to convince’               | 2,653                    | 80.63%                    | 19.37%                |
| <i>animar a</i><br>‘to encourage’                  | 1,135                    | 56.30%                    | 43.70%                |
| <i>forzar a</i><br>‘to force’                      | 668                      | 72.75%                    | 27.25%                |
| <i>incitar a</i><br>‘to incite’                    | 388                      | 67.78%                    | 32.22%                |
| <i>autorizar a</i><br>‘to authorize’               | 367                      | 58.04%                    | 41.96%                |

## 4.2. Geographical distribution of the accusative/dative alternation

In Table 3, the geographical distribution of the variation dative/accusative is presented. The table displays the normalized frequencies for each verb with lemmas *lo* or *le* in Spain and in Latin America. In order to first present a more general view of the issue, we have not divided the frequencies for Latin America by different regions; instead, they are displayed as ranges. When presented in this way, the frequencies show that most of the examined verbs prefer the dative construction (*le*) in Spain (with the exception of *convencer* ‘to convince’), while Latin America usually opts for the accusative (*lo*). This is in accordance with the overview we find in other studies (s. section 2.2.) that also confirm that Spain, as a whole, favours the older conservative option, that is, the presence of the dative clitic, whereas Latin America disfavours its use.<sup>120</sup>

Table 3. Geographical distribution of the accusative/dative alternation with selected manipulation verbs in Spain and Latin America

| verb +<br>preposition                                 | Spain  |                                 | Latin America                       |                                 |
|---|--|---------------------------------|-------------------------------------|---------------------------------|
|   | norm. freq. <sup>121</sup><br>accusative <i>lo</i> | norm. freq.<br>dative <i>le</i> | norm. freq.<br>accusative <i>lo</i> | norm. freq.<br>dative <i>le</i> |
| <i>ayudar a</i><br>‘to help’                          | 8.45   | 16.68                           | 13.99–18.60                         | 5.57–20.48                      |
| <i>obligar a</i><br>‘to make someone<br>do something’ | 10.07  | 16.3                            | 20.52–25.45                         | 2.43–6.31                       |
| <i>invitar a</i><br>‘to invite’                       | 7.81   | 8.38                            | 15.78–23.41                         | 1.53–3.64                       |
| <i>enseñar a</i><br>‘to teach’                        | 1.67   | 6.43                            | 0.55–2.39                           | 7.33–10.92                      |
| <i>convencer de</i><br>‘to convince’                  | 4.31   | 3.35                            | 6.09–8.76                           | 0.36–0.71                       |
| <i>animar a</i><br>‘to encourage’                     | 2.21   | 3.41                            | 0.88–2.63                           | 0.32–0.59                       |
| <i>forzar a</i><br>‘to force’                         | 0.92   | 1.14                            | 1.17–1.99                           | 0.15–0.2                        |
| <i>incitar a</i><br>‘to incite’                       | 0.49   | 0.83                            | 0.59–1.08                           | 0.12–0.14                       |
| <i>autorizar a</i><br>‘to authorize’                  | 0.16   | 0.6                             | 0.56–1.45                           | 0.32–0.48                       |

<sup>120</sup> For an analysis by areas, see the next section.

<sup>121</sup> Normalized frequency per 1 million words.

This is most obvious with verbs such as *obligar* ‘to make someone do something’ and *invitar* ‘to invite’, while the strength of the preference for one option over the other decreases with verbs at the end of the table, namely *animar* ‘to encourage’, *convencer* ‘to force’, *incitar* ‘to incite’, and *autorizar* ‘to authorize’. However, we must consider that the last four verbs are generally not as frequent in the corpus, which prompted us to focus on the first verbs in more detail. Another reason for looking at the first five verbs in this table more closely is the fact that some of the data observed here diverges from the general tendencies, as we can see with verbs like *ayudar* ‘to help’, *enseñar* ‘to teach’, and *convencer* ‘to convince’. In addition, some of the frequency ranges for Latin America are wider, which shows us there is a discrepancy between different areas, and only a closer look into the data can help clarify the extent of regional variation.

### 4.3. Analysis of the most frequent verbs by areas

When it comes to the verb *ayudar* ‘to help’ (Table 4), four regions of Latin America exhibit a preference for the accusative. This preference is especially strong in the Río de la Plata region (namely, Argentina). However, it is interesting to note that the continental Caribbean region (i.e. Colombia and Venezuela), as well as Mexico and Central America prefer the dative, which is a more conservative option that is typically characteristic of Spain.<sup>122</sup> This is especially important because previous research shows that Caracas seems to opt for the accusative in these constructions (Aleza Izquierdo 2009: 266–268), and the data on Mexico is mixed. Fernández-Ordóñez (1999: 1331) notes that with this verb Mexico and Central America prefer the dative, while Aleza Izquierdo (2009: 266–268) reports accusative for Mexico City and DeMello’s (2002: 264) data for Mexico City displays no preference between the two clitics.<sup>123</sup> DeMello (2002: 269) goes even further and claims that there is a difference in meaning between *ayudarlo* and *ayudarle*, where *le* signifies a more active participant than *lo*. Rivas (2013: 162–163) uses this distinction to explain the overall preference for dative with the verb *ayudar* ‘to help’

<sup>122</sup> As it was mentioned in section 2.2., a difference exists regarding dative/accusative alternation with these verbs between different Spanish regions. However, since the corpus we used does not provide the data to allow for that kind of distribution, we will not make any assumptions about this in our analysis.

<sup>123</sup> It should be noted that DeMello (2002) does not look into *ayudar* ‘to help’ only with the prepositional object, but in all occurrences. In addition, his corpus is smaller, and he recorded a total of eight examples of this verb, four with *lo* and four with *le*.

in these constructions, saying that «dative marking is preferred in order to show that the referent of the object has an active participation in the activity expressed by the matrix verb». For our data, however, this explanation is incomplete, as it should give more importance to the differences between the varieties of Spanish language.

Table 4. Geographical distribution by areas of the accusative/dative alternation with *ayudar + a* 'to help' and *enseñar + a* 'to teach'

| Ayudar a 'to help'          |  |                                       | Enseñar a 'to teach'        |  |                                       |
|-----------------------------|--|---------------------------------------|-----------------------------|--|---------------------------------------|
| Area <sup>124</sup>         | norm. freq.<br>accusative<br><i>lo</i> | norm.<br>freq.<br>dative<br><i>le</i> | Area                        | norm. freq.<br>accusative<br><i>lo</i> | norm.<br>freq.<br>dative<br><i>le</i> |
| Río de la Plata             | 18.6                                   | 5.57                                  | Continental<br>Caribbean    | 1.82                                   | 10.92                                 |
| The Antilles                | 17.23                                  | 11.28                                 | Chilean                     | 0.69                                   | 10.17                                 |
| Continental<br>Caribbean    | 14.67                                  | 19.77                                 | Río de la Plata             | 0.6                                    | 10.16                                 |
| Andean                      | 14.41                                  | 11.55                                 | Mexico &<br>Central America | 1.46                                   | 9.42                                  |
| Chilean                     | 14.33                                  | 10.12                                 | Andean                      | 0.55                                   | 8.28                                  |
| Mexico & Central<br>America | 13.99                                  | 20.48                                 | The Antilles                | 2.39                                   | 7.33                                  |
| Spain                       | 8.45                                   | 16.68                                 | Spain                       | 1.67                                   | 6.43                                  |

The verb *enseñar* 'to teach' (Table 4) prefers the dative both in Spain and in Latin America. This preference is registered in the literature, but the data so far was inconclusive regarding its distribution. Even those countries like Colombia, Uruguay, or Argentina, for which some authors claimed they exhibit a higher degree of alternation (s. Fernández-Ordóñez 1999: 1331),

<sup>124</sup> Countries included in each area: Andean: Bolivia, Ecuador, Peru; the Antilles: Cuba, Puerto Rico, Dominican Republic; Continental Caribbean: Colombia, Venezuela; Chilean: Chile; Mexico & Central America: Costa Rica, El Salvador, Guatemala, Mexico, Nicaragua, Panama; Río de la Plata: Argentina, Paraguay, Uruguay. The corpus also includes data from the United States of America, the Philippines, and Equatorial Guinea, but we chose not to include it in this distribution because of an extremely small number of documents and, by extension, examples for each of these varieties.

seem to prefer the dative. In other words, our data shows that the use of *le* is a general tendency with *enseñar* ‘to teach’ in present-day Spanish. This is in line with Rivas (2013: 162), who also reports a preference for dative in these constructions.

When it comes to its distribution in Latin America, the verb *invitar* ‘to invite’ (Table 5) displays a strong preference for the accusative that is in line with claims made by Fernández-Ordóñez (1999: 1326) and Rivas (2013: 162). Nevertheless, this preference is not always corroborated in previous research, especially in the case of the Antilles, Mexico, and Central America, that, according to NGLE (2009: §16.9o), should show a higher degree of alternation between the two forms. The data for Spain, however, does corroborate exactly this, even though the dative is slightly more frequent.

For the verb *obligar* ‘to make someone do something’ (Table 5), all the regions of Latin America favour the accusative, while in Spain the dative is more frequent. However, similarly to *invitar* ‘to invite’, previous research on the subject seems to indicate that Spain, the Antilles, Mexico, and Central America hesitate more between two options (Fernández-Ordóñez 1999: 1326; NGLE 2009: §16.9o). The same as for the verb *invitar* ‘to invite’, Rivas (2013) concluded that *obligar* ‘to make someone do something’ generally prefers the accusative clitic. Comparing the verbs *obligar* ‘to make someone do something’ and *invitar* ‘to invite’ with *ayudar* ‘to help’, and *enseñar* ‘to teach’, he (Rivas 2013: 163) claims that *obligar* ‘to make someone do something’ and *invitar* ‘to invite’ «present their objects as passive entities that do not exert any control on the activity expressed by the matrix verb» and, because of this, disfavour the dative form. Again, this explanation has to be expanded to include the dialectal situation, especially in Spain, which Rivas (2013: 170) agrees with.

Table 5. Geographical distribution by areas of the accusative/dative alternation with *invitar + a*, *obligar + a* and *convencer + de*

| Invitar a 'to invite'    |                                  |                              | Obligar a<br>'to make someone do something' |                                  |                              | Convencer de 'to convince' |                                  |                              |
|--------------------------|----------------------------------|------------------------------|---|----------------------------------|------------------------------|----------------------------|----------------------------------|------------------------------|
| Area                     | norm. freq. accusative <i>lo</i> | norm. freq. dative <i>le</i> | Area  | norm. freq. accusative <i>lo</i> | norm. freq. dative <i>le</i> | Area                       | norm. freq. accusative <i>lo</i> | norm. freq. dative <i>le</i> |
| Chilean                  | 23.41                            | 1.98                         | Río de la Plata                             | 25.45                            | 2.68                         | Río de la Plata            | 8.76                             | 0.71                         |
| Continental Caribbean    | 22.28                            | 2.21                         | Chilean                                     | 25.00                            | 2.43                         | Mexico & Central America   | 7.35                             | 0.36                         |
| Río de la Plata          | 19.51                            | 1.53                         | Mexico & Central America                    | 23.87                            | 4.19                         | Chilean                    | 7.09                             | 0.29                         |
| Mexico & Central America | 19.09                            | 3.07                         | Continental Caribbean                       | 23.84                            | 2.9                          | Continental Caribbean      | 6.85                             | 0.36                         |
| The Antilles             | 16.12                            | 2.75                         | Andean                                      | 20.73                            | 6.31                         | The Antilles               | 6.13                             | 0.35                         |
| Andean                   | 15.78                            | 3.64                         | The Antilles                                | 20.52                            | 5.24                         | Andean                     | 6.09                             | 0.48                         |
| Spain                    | 7.81                             | 8.38                         | Spain                                       | 10.07                            | 16.3                         | Spain                      | 4.31                             | 3.35                         |

The verb *convencer* 'to convince' (Table 5), like *invitar* 'to invite', shows a high degree of vacillation between the accusative and the dative in Spain, even though the accusative clitic seems to be somewhat more frequent. Of all the verbs that we investigated, *convencer* 'to convince' is the only one that exhibits a slight preference for the accusative in Spain. In Latin America, the general tendency is observed, and the accusative is the preferred option in all the areas, even though previous research on the subject seems to indicate that, again, the Antilles, Mexico, and Central America hesitate more between the two options (NGLE 2009: §16.9o).

Summarizing, even though the verb *ayudar* 'to help' in these constructions is slightly more frequent with the dative (Table 2), when we look into the geographical distribution, the image changes. The verb *ayudar* 'to help' occurs more often with the accusative marking in the areas of Río de

la Plata, the Antilles, Andean, and Chilean, while dative is preferred in the Continental Caribbean, Mexico and Central America, and Spain. On the other hand, the verb *enseñar* 'to teach' shows a clear preference for the dative both in Spain and in Latin America. The verbs *invitar* 'to invite' and *obligar* 'to make someone do something' prefer accusative in Latin America and dative in Spain, even though both forms are present in Spain, especially in the case of *invitar* 'to invite' that exhibits a high degree of vacillation between the two. The same can be said for *convencer* 'to convince', that is, surprisingly, more frequent with the accusative in Spain. The whole of Latin America prefers the accusative marking with this verb as well.

Our analysis shows that we have to be careful with any general conclusions and that it is crucial to take into account the dialectological differences between Spanish speaking countries (cf. García-Miguel 2015). Furthermore, and in line with Rivas (2013), it has been corroborated that these verbs in this construction differ between themselves, so any conclusions about them as a whole are likely to be incomplete.

Finally, the group of verbs analysed in this study points to a continuous dynamic relationship between the dative and accusative marking, because in these constructions dative is the one being replaced with the accusative. This goes against diachronic evidence gathered by Company Company (2001: 37) and her claim that «whenever a Dat is syntactically in a position to compete with an Acc, the Dat wins».

#### 4.4. Manual analysis of other variables

The second part of our research was the manual analysis of 15% of all examples by marking them for different variables. Each example was marked for five different variables, four of which (case, number, gender, and type of clause) are displayed in Table 6. The fifth variable was person, in order to exclude from the analysis the examples of previously mentioned *leísmo de cortesía*, that amounted to 7% of the total number of dative examples marked manually.

The example (19) contains a clitic *las* that is accusative, plural, and feminine and the verb of its subordinate clause is in the infinitive form, whereas the example (20) has a clitic *le* that is dative, singular, and masculine and the verb of its subordinate clause is in the finite form. Gender with dative clitics was marked by reading through each segment to find a referent and if one was not found, that was labelled by an X. It was also possible that the

referents were of both genders (marked as MF), as in (21).<sup>125</sup> A third type of clause was actually not a clause but a noun phrase (marked as NOM), as in (22).

Table 6. Variables marked manually and examples

| Variables      |     | Example  |
|----------------|-----|--|
| CASE           | ACC | (19) <i>enseñarlas a colgarse de la barra</i><br>'teach <b>them</b> to hang from the bar'  |
|                | DAT | (20) <i>...me costaría convencerle de que suba...</i><br>'...I would have a hard time convincing <b>him</b> to go up...'   |
| NUMBER         | SG  | (20) <i>...me costaría convencerle de que suba...</i><br>'...I would have a hard time convincing <b>him</b> to go up...'   |
|                | PL  | (19) <i>enseñarlas a colgarse de la barra</i><br>'teach <b>them</b> to hang from the bar'  |
| GENDER         | M   | (20) <i>El obispo es un buen amigo mío, me costaría convencerle de que suba hasta el monasterio.</i><br>'The bishop is a good friend of mine, I would have a hard time convincing <b>him</b> to go up to the monastery.' |
|                | F   | (19) <i>enseñarlas a colgarse de la barra</i><br>'teach <b>them</b> to hang from the bar'  |
|                | MF  | (21) <i>poner límites a niñas y niños no sólo les ayuda a tener claro...</i><br>'Setting limits for girls and boys not only helps <b>them</b> to be sure...'   |
| TYPE OF CLAUSE | INF | (19) <i>enseñarlas a colgarse de la barra</i><br>'teach them <b>to hang</b> from the bar'  |
|                | FIN | (20) <i>...me costaría convencerle de que suba...</i><br>'...I would have a hard time convincing him <b>to go up</b> ...'  |
|                | NOM | (22) <i>lo obligó a eso el calor</i><br>'the heat made him <b>do that</b> '  |

<sup>125</sup> It should be mentioned that in Spanish masculine plural forms can be used generically to refer to both genders, so, for example *los niños* can include girls as well as boys. In those cases, the example was marked as MF. However, due to the lack of wider context, this was often difficult to ascertain, leading to marking most of these masculine plurals as M, observing their morphological form. Having in mind that at least during the last two decades and the gender equality awareness the generic use of masculine has been dissuaded in favour of visibility of both genders, it would be interesting to find out the current situation in language use.

Looking into the type of complement clause, our analysis showed that Spanish manipulation verbs tend to come with the infinitive form, which is a conclusion that was already corroborated by other authors (Fernández-Ordóñez 1999: 1325–1326; Rivas 2013: 163). Finite clauses and the nominalized forms seem to be similar in regard to frequency in these constructions (Table 7).

Table 7. Results for manually marked variables: number, gender and type of clause

| CLITIC TYPE: ACC |     |      |     | CLITIC TYPE: DAT |     |      |     |
|------------------|-----|------|-----|------------------|-----|------|-----|
|                  |     | 2595 |     |                  |     | 1798 |     |
| NUMBER           | SG  | 1943 | 75% | NUMBER           | SG  | 1207 | 67% |
|                  | PL  | 652  | 25% |                  | PL  | 591  | 33% |
| GENDER           | M   | 1615 | 62% | GENDER           | M   | 1237 | 69% |
|                  | F   | 949  | 37% |                  | F   | 270  | 15% |
|                  | MF  | 30   | 1%  |                  | MF  | 42   | 2%  |
|                  | X   | 1    | 0%  |                  | X   | 249  | 14% |
| TYPE OF CLAUSE   | INF | 2065 | 80% | TYPE OF CLAUSE   | INF | 1630 | 91% |
|                  | FIN | 259  | 10% |                  | FIN | 96   | 5%  |
|                  | NOM | 271  | 10% |                  | NOM | 72   | 4%  |

As for the variable of gender, as stated before, due to the nature of the manipulation verbs, the object referent in all our examples was a person. According to the literature (Fernández-Ordóñez 1999: 1319; NGLE 2009: §16.8c), we expected a higher number of masculine in comparison to feminine dative clitics, given that masculine *leísmo* tends to be significantly more frequent than feminine. This is corroborated in our data (Table 7), even though for 14% of dative examples, the gender of the referent could not be discerned from the text. Nevertheless, they were equally distributed throughout the corpus and were simply a result of insufficient available context. Therefore, there is no reason to believe that the gender distribution within that group would not be the same as for the rest of the dative examples. To investigate the relationship between the type of clitic and gender, we applied a chi-square test, taking into account only the examples where gender was clear from the context. The test has revealed that the two variables were related ( $\chi^2=164.994$ ,  $df=1$ ,  $p<0.0001$ ,  $\varphi=0.201$ ), that the type of clitic is affected by the gender of the clitic's referent. In other words, the gender of the referent has a role in the choice of clitic, even though the

correlation is weak. We can conclude that the type of clitic can to a small degree be predicted by the gender, more specifically female gender, given that the observed frequency for female accusative clitics was higher than expected, contrary to the observed frequency for female dative clitics which was lower than expected (Table 8).

Table 8. Observed and expected frequencies for variables of gender and number across the type of clitic

|               | ACC      | DAT      |
|---------------|----------|----------|
| <b>M_exp</b>  | 1796.249 | 1055.751 |
| <b>M_obs</b>  | 1615     | 1237     |
| <b>F_exp</b>  | 767.7514 | 451.2486 |
| <b>F_obs</b>  | 949      | 270      |
| <b>SG_exp</b> | 1860.744 | 1289.256 |
| <b>SG_obs</b> | 1943     | 1207     |
| <b>PL_exp</b> | 734.2556 | 508.7444 |
| <b>PL_obs</b> | 652      | 591      |

In regard to the variable of number, we expected singular *leísmo* to be more frequent than plural, which is also confirmed by our data (Table 7). In order to investigate a possible connection of the type of clitic and number, we have conducted a chi-square test here as well. The result ( $\chi^2=31.398$ ,  $df=1$ ,  $p<0.0001$ ,  $\phi=0.084$ ) again showed a statistically significant correlation between the two variables, but with a very weak effect size. The same as with gender, the number of the referent influences the choice of clitic, albeit a very weak influence. The difference is that there was slightly more plural *leísmo* than expected and less singular (Table 8).

Therefore, in addition to concluding that in these constructions with manipulation verbs masculine *leísmo* is more frequent than feminine, as well as singular *leísmo* in comparison to plural, (which is in accordance with previously stated order of types of *leísmo* by frequency, see section 2.1.), we can also ascertain the correlation between the type of clitic and grammatical features of its referent, namely gender and number. However, the effect size is weak for the first and very weak for the second variable.

## 5. Conclusions

To sum up, it can be noted that our analysis confirms that manipulation verbs which govern a prepositional object exhibit alternation between accusative and dative clitics in modern Spanish. When it comes to the geographical distribution, generally speaking, the Peninsular Spanish favours the dative clitic, whereas the Latin American Spanish prefers the presence of an accusative clitic. In other words, the use of clitics is more conservative in the Peninsular Spanish than in the Latin American Spanish. However, it is crucial to add that these tendencies are not distributed equally across Spanish-speaking regions and vary from verb to verb.

This also brings us closer to understanding the grammatical relations of two core participants in Spanish: the direct and indirect object. We are inclined to agree with García-Miguel (2015) that there are no clear boundaries between them, and we believe that the results of our study, where the clitic case is subject to dialectal factors and there is even considerable variation across verbs in this relatively homogenous group, point to that.

As for the analysis of other variables, the most frequent type of clause introduced by the preposition tends to have the verb in the infinitive form. We can also conclude that the most frequent type of *leísmo* in our research is the one for human masculine singular object, that is to say, these constructions with the manipulation verbs follow the general tendency of distribution of types of *leísmo*. Furthermore, the gender and number of the referent seem to weakly affect the choice of clitic.

Future research should establish whether the tendencies we have observed in our analysis can be found in all Spanish varieties and account for more subtle differences between them. For example, according to some studies (s. section 2.2.), we could expect that, for those verbs that present a higher degree of vacillation in our data, some varieties of the Peninsular Spanish do favour one option over the other, but this could not be confirmed in our research since our corpus does not differentiate between various regions in Spain.

## Appendix

### Acknowledgments

We are grateful for the valuable comments given by two anonymous reviewers and the editors of this volume.

### Abbreviations

|      |                    |
|------|--------------------|
| 1    | first person       |
| 2    | second person      |
| 3    | third person       |
| acc  | accusative         |
| cl   | clitic             |
| dat  | dative             |
| det  | determiner         |
| exp  | expected           |
| f    | feminine           |
| fin  | finite             |
| indf | indefinite         |
| inf  | infinitive         |
| m    | masculine          |
| MF   | masculine/feminine |
| nom  | nominative         |
| obs  | observed           |
| pl   | plural             |
| pst  | past               |
| ptcp | participle         |
| sg   | singular           |
| X    | unknown referent   |

### Acronyms

NGLE *Nueva gramática de la lengua española*

## References

- Alarcos Llorach, E. (1999). *Gramática de la lengua española*. Madrid: Espasa Calpe.
- Aleza Izquierdo, M. (2009). El pronombre personal complemento: los clíticos. In: Hernández, C. (ed) *Estudios lingüísticos del español hablado en América. Volume 3.2*, 223–351. Madrid: Visor.
- Company Company, C. (2001). Multiple dative-marking grammaticalization: Spanish as a special kind of primary object language. *Studies in Language*, 25(1), 1–47.
- Company Company, C. (2003). Transitivity and grammaticalization of object. The struggle of direct and indirect object in Spanish. In: Fiorentino, G. (ed) *Romance objects. Transitivity in Romance languages*, 217–260. Berlin/New York: Mouton de Gruyter.
- DeMello, G. (2002). “Leísmo” in Contemporary Spanish American Educated Speech. *Linguistics*, 40(2), 261–283.
- Díaz Montesinos, F. (2017). Leísmo real y leísmo aparente. Sobre el uso de los pronombres átonos (le, les, la, las, lo, los) en Málaga. *Verba: Anuario galego de filoloxía*, 44, 409–480.
- Fernández-Ordóñez, I. (1993). Leísmo, láismo y loísmo: estado de la cuestión. In: Fernández Soriano (ed) *Los pronombres átonos*, 63–96. Madrid: Taurus.
- Fernández-Ordóñez, I. (1999). Leísmo, láismo y loísmo. In: Demonte Barreto, V. & Bosque, I. (eds) *Gramática descriptiva de la lengua española. Volume 1*, 1317–1398. Madrid: Espasa Calpe.
- Fernández Ramírez, S. (1964). Un proceso lingüístico en marcha. In: *Presente y futuro de la lengua española, Actas de la Asamblea de Filología del I Congreso de Instituciones Hispánicas. II*, 277–285. Madrid: Ediciones Cultura Hispánica.
- Fernández Ramírez, S. (1987). *Gramática española. 3.2. El pronombre*. Madrid: Arco Libros.
- García-Miguel, J. M. (2015). Variable coding and object alignment in Spanish: A corpus-based approach. *Folia Lingüística*, 49(1), 205–256.
- García González, F. (1981). /le (lu), la, lo (lu)/ en el Centro-Norte de la Península. *Verba*, 8, 347–535.
- Givón, T. (2001). *Syntax: an introduction. Volume I*. Amsterdam/Philadelphia: John Benjamins.
- Halpern, A. L. (2017). Clitics. In: Spencer, A. & Zwicky, A. (eds) *The Handbook of Morphology*, 101–122. Oxford: John Wiley & Sons.
- Klein-Andreu, F. (1993). Distintos sistemas de empleo de “le”, “la”, “lo” – perspectiva sincrónica, diacrónica y sociolingüística. In: Fernández Soriano, O. (ed) *Los pronombres átonos*, 337–354. Madrid: Taurus.
- Lapesa, R. (1968). Sobre los orígenes y evolución del leísmo, láismo y loísmo. In: Baldinger, K. (ed) *Festschrift Walther von Wartburg*, 523–551. Tübingen: Max Niemeyer.
- Mayer, E. (2017). *Spanish clitics on the move: Variation in time and space*. Berlin, Boston: De Gruyter Mouton.
- Mikelenić, B. (2020). *Kontrastivna korpusna analiza prijedložne dopune u španjolskome i njezinih ekvivalenata u hrvatskome*. PhD thesis, University of Zagreb.
- Real Academia Española y Asociación de Academias de la Lengua Española (2009). *Nueva gramática de la lengua española*. Madrid: Espasa Calpe.
- Rivas, J. (2013). A Usage-Based Analysis of Morphosyntactic Variation in the Complementation of Manipulative Verbs. *Revista Internacional de Lingüística Iberoamericana*, 11(1/21), 155–171.
- Roegiest, E. (1990). La tipología sintáctica del objeto transitivo en español. *Verba*, 17, 239–248.
- Spencer, A. & Luís, A. R. (2012). *Clitics. An Introduction*. Cambridge University Press.
- Vinja, V. (1998). *Gramatika španjolskog jezika*. Zagreb: Školska knjiga.