Senta Zeugin

DOM in Modern Catalan varieties

An empirical study based on acceptability judgment tasks

Abstract: This paper examines the current situation of Differential Object Marking (DOM) in the major Catalan varieties, focusing on the animacy of the direct object. In the past, Catalan standard grammar has advocated for very conservative rules for applying DOM, and thus a substantial divergence between spoken language and standard grammar was created. Nowadays, standard Catalan seems to accept DOM more readily, but still falls short of the wide range of uses found in Catalan dialects. Hence, this paper relies on experimental methods (i.e. an acceptability judgment task with a 2x3 within-subjects design) to analyze the situation found in the major Catalan dialects, i.e. Central Catalan, North-Western Catalan, Valencian and Majorcan. The experiment focuses on direct objects in definite noun phrases and the test sentences were manipulated for presence/absence of DOM and three degrees of animacy (human, animal, inanimate). Results suggest that for human direct objects in noun phrases the use of DOM is as well accepted as the unmarked version, which contradicts what is claimed by Catalan standard grammars. However, the unmarked version is slightly preferable in the case of animals and clearly preferable for inanimate objects.

Keywords: Differential Object Marking (DOM), Catalan, experimental linguistics, acceptability judgment tasks, diatopic variation

1 Introduction

For some languages, research into DOM is already quite advanced, whereas for others, much ground remains to be covered. One such language is Catalan. Although an increasing number of studies on DOM have been published in recent years, there is still more to discover about the phenomenon. This is one of the

Acknowledgements: This article is based largely on the author's MA-Thesis (Zeugin 2017). I'd like to thank two anonymous reviewers as well as Albert Wall, Johannes Kabatek and John Barlow for their comments and suggestions. I am also very grateful to all the native speakers who reviewed the material for the experiment as well as everyone who participated in the experiment itself.

Senta Zeugin, University of Zurich, senta.zeugin@uzh.ch

main reasons why the experiment described in this article takes Catalan and its main varieties as its focus.

Historically, the view of Catalan authors on DOM, which is often seen as a prototypical feature of neighbouring Spanish, has been rather biased and less than favourable. For the most part it was regarded as a mere alien structure, existing only because of language contact, and thus, further studies were not deemed necessary.¹ However, this view has not been shared by everyone, and there have been a few early defenders of DOM as an inherent Catalan feature. Meier (1945) shows that already in the earlier stages of Catalan, DOM had spread to varying degrees throughout its territories, citing examples from authors such as Bernat Metge and Sant Vicent Ferrer (both 14th/15th century) and Jacint Verdaguer (19th century). Geographical differences can be explained, as Müller (1971) notes, as a result of the gradual loss of the Latin case system (favouring the accusative as the universal case), which for Catalan seemed to start in the south of the territory and spread slowly towards the north. Thus, according to Müller (1971), it should come as no surprise that a text from Valencia (Sant Vincent Ferrer) contains more occurrences of DOM than a contemporary work from Barcelona (Metge). The author goes on to argue that the late starting point for the loss of the Latin case system seems to lie at the root of the much narrower expansion of DOM usage in Catalan compared to Spanish, where this loss began about a century earlier. Further confirmation of the autonomous origins of Catalan DOM can be found in more recent studies, for example by Irimia/Pineda (2019). The authors underline fundamental differences between Spanish and Catalan DOM and cite examples from texts as early as the 14th century.

As a result of this change of perception, the study of Catalan DOM has become more detailed and objective over the years. For example, Rohlf (1971) and later Pensado (1995), Wheeler/Yates/Dohls (1999) and Sancho Cremades (2008) present detailed characterizations, adding numerous examples of DOM in contexts that were not described in earlier works. Thus, they demonstrate, contrary to what until then had often been claimed by Catalan linguists and to what Catalan normative works suggested, that DOM was actually being used much

¹ One of the defenders of this viewpoint was Par (1923, 153), who stated that: “En català modern jamay lo règim directe porta la preposició ‘a’, quan aquell es un nom. Mes ab los pronoms personals tònichs s’hi es introduhida, sens dubte per influència castellana; barbarisme qui es acceptat, àdhuch per qualques gramàtics” (‘In Modern Catalan, the direct object in the form of a noun is never marked by the preposition *a*. However, it is used with strong personal pronouns, without a doubt due to Castilian influence; a barbarism that is, nonetheless, accepted by some grammarians’ [translation SZ]). Badia Margarit’s (1962) opinion is not quite so extreme and, similarly to Pompeu Fabra (1933), he describes various permissible contexts for DOM. However, both authors do still urge to avoid it and to use any other means whenever possible.

more widely. The view that DOM in Catalan is nothing more than a contamination from Spanish has been questioned for some time and is considered outdated. In separating the two phenomena and recognizing the historic and synchronic individuality of Catalan DOM, a need for a detailed description of its evolution and present occurrence arises. And it is in this latter area of investigation that the current paper seeks to contribute to the bigger picture of DOM in Catalan in general and its major dialects in particular.

For this purpose, an experiment was conducted, consisting of an acceptability judgment task in which participants were asked to evaluate 60 sentences, presented to them in a random order. Before describing the experiment, some background information on the situation of DOM in the major Catalan dialects will be provided. Therefore, Section 2 contains an overview of the status of DOM in standard Catalan as well as a brief introduction to the diatopic study of DOM in four major Catalan dialects. The remaining Sections will focus on the experiment itself, starting with its method and general setup in Section 3, and the design of the test sentences in Section 4. Subsequently, Section 5 covers the procedure of the experiment as well as relevant information about the participants themselves. Lastly, in Section 6, the results from the different parts of the experiment will be discussed, in each case by first focusing on the regions individually and then looking at all the dialects together. Section 7 then presents conclusions, including some observations on possible future research.

2 DOM: standard Catalan versus Catalan dialects

When studying the behaviour of dialects, what is predominantly of interest is the spoken language, considering that this is where diatopic variation usually manifests itself. In order to establish such variations, the dialects themselves can be compared. An additional source of insight is the contrast of each dialect with the established standard version of the respective language. In most languages, if not all, there is a certain divergence between standard and everyday language as a natural result of their individual purposes. Additionally, in the case of Catalan, the standard language has served as a means of distinguishing Catalan itself more clearly from neighbouring Spanish, as can be seen, for example, in the works of Fabra (1933) and Badia Margarit (1962). Thus, differences between the standard and the spoken language might be more pronounced in Catalan² than

² Sancho Cremades (2008) confirms the existence of a considerable gap between standard and spoken language in his chapter in the *Gramàtica del català contemporani*.

in other languages, and this suggests that it might be useful to take a brief look at the process of linguistic normalization and its important role in the creation of standard Catalan.³

During a long time the normative version of Catalan, i.e. standard Catalan, was defined in three works by Pompeu Fabra from the beginning of the 20th century: the *Normes ortogràfiques*, published in 1913, the *Gramàtica catalana* from 1918 and the *Diccionari general de la llengua catalana* from 1931 (cf. Costa Carreras 2001). Especially the grammar from 1918, and later on its updated version from 1933, was regarded as normative up until very recently, as Costa Carreras (2001) notes.

Fabra deemed only a very limited number of contexts acceptable territory for DOM, and suggested to avoid it whenever possible. Among the very limited permitted uses in Fabra (1933) is the marking of strong personal pronouns, elements like *tots*, *tothom* and *el qual*, as well as expressions of reciprocity (*l'un a l'altre*) plus cases of possible ambiguity. A rather problematic aspect of Fabra's standard Catalan was the fact that he based it predominantly on the Central Catalan dialect and thus did not pay heed to the wide spectrum of Catalan dialects.⁴ Furthermore, standard Catalan, having its origins in about 1918, did not receive a

3 Linguistic normalization (*normalització lingüística*) is a phrase coined by Catalan sociolinguists, referring to the joint processes of establishing a linguistic standard, or norm, for a certain language (*normativització*), and the extension of this language to all registers and contexts (*extensió social*) (cf. Kremnitz 1979; Vallverdú 1979; Lagarde 2009). In the case of Catalan, the process of normalization was not without its difficulties. At the beginning of the 18th century after the War of the Spanish Succession, Catalan was substituted by Spanish as the official language of Catalonia. During the following two centuries Catalan first experienced a significant decline in usage, and later on near the end of the 19th century, led by a cultural renaissance, a surge in popularity (cf. Vallverdú 1979). Fueled by constant political tension, the Catalan-speaking, antimonarchist bourgeoisie saw the potential of conserving and reanimating their mother tongue, in order to gain the support of the working class population. With this historical background, the linguistic normalization of Catalan started at the beginning of the 20th century, particularly with the works of Pompeu Fabra. Shortly after, during the dictatorship of Primo de Rivera (1923–1930) and from 1939 on under Franco, any languages other than Castilian Spanish were prohibited (cf. Kremnitz 1979). In spite of all the setbacks, the Catalan normalization process never came to a complete stop. Still, it wasn't until the 1970s that it was able to run much more smoothly. Furthermore, this historical background explains some of the political tension that still surrounds Catalan today and the insistence of some of the older grammars on distancing Catalan as far as possible from Spanish.

4 Fabra's initial idea was to create a much more inclusive standard, striving to use not only the Central dialect, but to let it be influenced and moulded by the varieties of the other regions, all the while getting rid of any vestiges of Spanish (cf. Fabra 1908). Colón Domènech (2009, 23) indicates that, while Fabra's intentions might have been good initially, in the end he developed a standard based only on one dialect: "Malgrat una pretesa voluntat de fer un estàndard acumulatiu, l'estàndard en realitat és restrictiu; és a dir, 'particular'" ('Despite the intention of creating a cumulative standard, the standard is actually restrictive, or rather, 'particular' [translation SZ]).

comprehensive update for almost a century until the publication in 2016 of the *Gramàtica de la llengua catalana*, the new normative grammar of the *Institut d'Estudis Catalans* (IEC).

As far as DOM is concerned, this new normative work generally seems to follow in the footsteps of Fabra, stating that, as a rule, the direct object does not require a preposition. Also, in most of the permitted contexts for DOM, it appears to concur with Fabra, while providing much more detailed explanations and pointing out several exceptions. The authors of this recent grammar include a number of new permitted contexts for DOM, for example, when the order of a sentence has been changed by dislocations. It appears that with this new grammar, the normative authorities (nowadays the IEC) have slowly come around to recognizing the existence and validity of DOM as a Catalan phenomenon. The authors even go so far as to refer to the presence of diatopic variation, albeit without providing further details. Nevertheless, a look at the regional grammars and studies on the subject suggests that the updated standard language is still somewhat removed from the real situation of the spoken language throughout the territory.

Illustration 1: The Catalan dialects⁵.

⁵ Illustration 1: Wikimedia CC BY-SA 2.5; annotations of dialect names [SZ]. Names of the dialects are consistent with those used in this article. In the case of Majorcan it should be noted that this

In the remainder of this Section, a brief overview of the situation of DOM in the four major Catalan dialects will be presented. As already mentioned, the Central Catalan dialect was the general basis for the development of the Catalan standard in the early 20th century. An analysis of spoken Central Catalan can be found in Escandell Vidal (2009). Her study illustrates that even the most closely related dialect diverges to a certain degree from the standard language, in the case of DOM most often concerning its use with proper names, human definite noun phrases and occasional non-human animate ones.⁶ Still, the behaviour of Central Catalan is somewhat similar to standard Catalan, whereas especially in Majorcan and Valencian the extension of DOM is much greater than what is considered permissible in the general Catalan standard and also in the respective regional standards.

For the Balearic dialect, Moll (1979) proposes similar guidelines to those of Fabra's grammar, while at the same time introducing one of the earliest recognized examples of DOM with inanimate objects,⁷ a use that certainly goes beyond any normative prescriptions. In more recent studies, such as Wheeler/Yates/Dols (1999) or Escandell Vidal (2009), the Majorcan situation is addressed in even greater detail. The latter author, echoing a previous study by Rohlf's (1971), emphasizes that DOM in Majorcan is found primarily in dislocated positions. According to Escandell Vidal (2009), it seems that not only intrinsic factors of the direct object, such as animacy and definiteness, have an impact on the use of DOM in Majorcan, but also others like topicality and information structure.

In contrast to Majorcan, to which many studies concentrating on DOM are dedicated, for North-Western Catalan there is much less information available. One of the few studies, by Boladeras Taché (2011), is very inconclusive concerning the contexts where DOM can be found. Nonetheless, the author does allude to one notable aspect, which itself probably hinders a more precise examination: the fact that in North-Western Catalan marked and unmarked objects, in combination with the definite article, can sound the same.

On Valencian, on the other hand, much more information can be found, and it tends to suggest a range of contexts for DOM, which goes beyond the uses permitted in the new normative grammar by the IEC. There are examples such as *Hem vist com perseguia un policia a un lladre* (AVL 2006, 303) that show strong

doesn't include the dialects spoken on Ibiza or Menorca. Considering that all participants in the experiment in question came from the island of Majorca, the more generic term Balearic Catalan seemed inappropriate in this instance.

⁶ Escandell Vidal (2009, 840) cites examples like *Les monges no estimen a/ana les nenes* and *Veuré a la Maria*.

⁷ The example from Moll (1979, 203) is: *Colliules, a les peres que ja són madures*.

parallels to the evolution found in Spanish DOM. Sancho Cremades (1995) confirms the similarity of Valencian to Spanish DOM, and adds that it is usually associated with direct objects denoting a person, the capacity of an object to feel the action described by the verb, or topicality. The only restriction noted by Sancho Cremades (1995, 199) is the need for the object to be clearly determined (*no veig als xiquets* [. . .] vs. *he trobat uns xiquets* [. . .]). However, considering the example cited above, even this restriction doesn't always seem to hold. A further interesting aspect is the fact that as early as the 15th century, there were occurrences of DOM in Valencian, as several examples found in the Sermons of Sant Vicent Ferrer demonstrate (*Senyor, yo bé am mon pare, mas més am a vós* cf. Meier 1945, 239).

3 The experiment: main aspects and methodology

As noted in the Introduction, the experiment discussed in this paper consisted of an acceptability judgment task that examined the present use of DOM in Catalan. Thus, the judgments of native Catalan speakers concerning the presence or absence of DOM in given sentences were studied. One of the important aspects of a more detailed description of DOM usage nowadays, and therefore one of the three main aspects investigated in this experiment, is diatopic variety. The geographic distribution of Catalan-speaking regions is considerable. The largest territory can be found in Spain, but there are also areas in France (Roussillon) and Italy (Alghero, Sardinia), resulting in contact situations with at least four different romance languages. Combined with the relative isolation of the Balearic Islands, this leads to a variety of different geographical situations as well as different language contact situations. The possible regional differences are confirmed by some of the previously cited literature, e.g. regional grammars, focusing on the major dialects of Catalan.

Consequently, 'regional variation' as a first major aspect was examined by distributing the experiment across the regions of the six major Catalan dialects. In addition to the regional varieties, the experiment focused on two further aspects: 'animacy' and, to a lesser degree, 'position'. Including both of these factors in the same test sentence was not feasible, because it would have rendered it impossible to determine which one was more influential. Therefore, the experiment was divided into two parts, each centring around one main aspect and thus avoided any conflict between them. Furthermore, since the sentences for both parts differed in structure, it was expected that participants would not easily deduce what linguistic aspect the experiment was focusing on. Ideally, participants evaluated the test sentences spontaneously and without thinking about them extensively.

As mentioned, a total of 60 randomly allocated sentences were given to each participant, divided into two groups of 30, that is, for the two parts of the experiment. Both parts were designed using the Latin Squares method, with six individual conditions distributed across six lists and five repetitions, leading to 30 test sentences per experiment part.

The first part was created around ‘animacy’, one of the main parameters generally used to determine different DOM-configurations.⁸ The focus lay on the animacy of the direct object in noun phrases. In accordance with the basic animacy scale shown below (cf., for example, von Heusinger/Kaiser 2005, 37), the degree of animacy of the direct object was manipulated to see whether this had an effect on the acceptability of the test sentences.

Animacy scale: HUMAN > ANIMATE > INANIMATE

By combining this scale with the presence or absence of the DOM marker *a*, six categories (conditions) were created, each of which was represented by the same number of test sentences. The distribution of the conditions is shown in Table 1 below.

Table 1: Conditions for ‘animacy’.

	+HUMAN	–HUMAN, +ANIMATE	–ANIMATE
DOM	Condition 1	Condition 2	Condition 3
∅	Condition 4	Condition 5	Condition 6

Whereas the main focus of the whole experiment was on this first part about the animacy of the direct object in combination with DOM, the sentences from the second part were mostly used as distractors. With this goal in mind, the chosen sentence structure clearly differed from the one in the first part, as will be seen later on. Nevertheless, in an effort to glean further information on the behaviour of the phenomenon in Catalan, DOM was also included. Thus, the third main aspect to be considered was ‘position’, or, more specifically, the influence of a change in the position of the pronominal direct object and its resulting combination with clitic doubling and DOM on the acceptability of the test sentence. As presented in Table 2, the first positional category was topicalization, bearing in mind that the literature suggests it to be highly susceptible to co-occurrence

⁸ Cf., for example, Bossong (1991), von Heusinger/Kaiser (2005) or Aissen (2003).

with DOM.⁹ Secondly, dislocation to the right was chosen, which seems to be a frequent partner with DOM, particularly in Majorcan (as indicated by Escandell Vidal 2009). Unfortunately, due to the design of the experiment and the need for a rigorous uniformity of the structure of the test sentences, it was not possible to use a total dislocation to the right, as would have been the ideal case. This has to be taken into account when analyzing the results of the second part of the experiment. In Catalan, both types of dislocation go hand in hand with the introduction of a strong personal pronoun, which is where DOM comes into play. As a third category for the factor ‘position’, the most basic sentence type was used, with neither dislocation nor clitic doubling, and thus also without DOM. In combining these three positional categories with a change between 2nd and 3rd person singular pronouns, six conditions were again created, as can be seen in Table 2.

Table 2: Conditions for ‘position’.

	Topicalization with clitic doubling	Dislocation to the right with clitic doubling	Neither dislocation nor clitic doubling
2nd SG	Condition 11	Condition 12	Condition 13
3rd SG	Condition 14	Condition 15	Condition 16

Following the characteristics that they represent, each condition shown in Tables 1 and 2 corresponds to a sentence, or rather, a variation of a test sentence. Consequently, a test sentence always had six variations. Naturally, a participant was not asked to evaluate all of these six variations but just one of them, thus avoiding interference caused by lexical repetition and cross-sentence judgments. As can be seen in Table 3, for the Latin Squares method with six conditions per test sentence, six different lists have to be created, each containing only one of the conditions per test sentence. Furthermore, every condition needs to appear the same number of times in each list, which guarantees the most equal distribution. Consequently, after six sentences, all of them with six variations, every list includes each condition exactly once.

In order to achieve reliable results, it is necessary to gather as much data as possible, thus making it tempting to present participants with a huge number of test sentences. Doing so would be risky and most likely result in the loss of many participants due to boredom. Therefore, a compromise between maximal data gain and minimal loss of participants had to be reached. Accordingly, a total of

⁹ Cf., for example, Rohlfes (1971) and Pensado (1995).

Table 3: Latin Squares.

	List 1	List 2	List 3	List 4	List 5	List 6
Sentence 1	c1	c2	c3	c4	c5	c6
Sentence 2	c6	c1	c2	c3	c4	c5
Sentence 3	c5	c6	c1	c2	c3	c4
Sentence 4	c4	c5	c6	c1	c2	c3
Sentence 5	c3	c4	c5	c6	c1	c2
Sentence 6	c2	c3	c4	c5	c6	c1

60 sentences (each with its six respective variations) were created, 30 per experiment part. In other words, for each of the experiment parts Table 3 was repeated five times, leaving every participant to evaluate each condition five times. Once created, the sentence variations were randomly distributed among the six lists, following the pattern seen in Table 3, and each participant was again randomly presented with only one of these lists, containing 60 sentences in total. Consequently, for each variation of a test sentence to be evaluated at least once (360 overall, 180 per experiment part), six different participants were needed.

4 The experiment: sentence construction

In accordance with the two major aspects under investigation, ‘animacy’ and ‘position’, two sentence groups were designed with the aim of acquiring as much information about DOM in the Catalan varieties as possible. At the same time, the intention was to control for as high a number of other presumably influential factors as possible. In the following, both sentence groups will be discussed in detail. An important aspect here is the fact that the language used for all test sentences is a more or less standard version of Catalan. Taking into account every regional variation of the major Catalan dialects when constructing the sentence catalogue would not have been practical. It would have meant either designing six individual catalogues, sacrificing the possibility of comparing the results on a statistical level, or using an amalgamation of diatopical traits that could only have led to confused participants. A second aspect of a more general nature concerns word choices. Throughout both experiment parts, no verb was used more than once and none of the subjects or objects of the sentences belonging to the ‘animacy’ part were repeated at any point. Furthermore, only transitive verbs were used.

Concerning the first experiment part, as pointed out, the aspect under scrutiny was the animacy of the direct object in noun phrases and its influence on the perceived acceptability of a given sentence with or without DOM. Following the basic animacy scale presented in Section 3, the direct objects chosen denote humans, animals and inanimate objects. All the sentences of this group exhibit exactly the same basic structure, as can be seen in Table 4 below, consisting of an animated subject, a transitive verb in *perfet perifràstic*, a direct object in the form of a noun phrase and a circumstantial complement. Thus, the only manipulated part of the sentence was the direct object itself in terms of its degree of animacy and presence/absence of the preposition *a*, which marks DOM.

Table 4: Examples of test sentences for experiment part ‘animacy’.

	+HUMAN	–HUMAN, +ANIMATE	–ANIMATE
DOM	El director va filmar <u>a la meva nora</u> per al documental sobre la vila.	El director va filmar <u>a la meva mula</u> per al documental sobre la vila.	El director va filmar <u>a la meva granja</u> per al documental sobre la vila.
∅	El director va filmar <u>la meva nora</u> per al documental sobre la vila.	El director va filmar <u>la meva mula</u> per al documental sobre la vila.	El director va filmar <u>la meva granja</u> per al documental sobre la vila.
DOM	El guàrdia va transportar <u>al meu nét</u> fins a l'estació de ferrocarrils.	El guàrdia va transportar <u>al meu gat</u> fins a l'estació de ferrocarrils.	El guàrdia va transportar <u>al meu sac</u> fins a l'estació de ferrocarrils.
∅	El guàrdia va transportar <u>el meu nét</u> fins a l'estació de ferrocarrils.	El guàrdia va transportar <u>el meu gat</u> fins a l'estació de ferrocarrils.	El guàrdia va transportar <u>el meu sac</u> fins a l'estació de ferrocarrils.

As far as the different categories of direct objects are concerned, it should be highlighted that for the human objects only terms denoting family members were used. On the one hand, this helped to keep the sentence structure as balanced as possible. On the other hand, it has been shown that, traditionally, they are more often found with DOM than other nouns. The choices for the animals of the second category of direct objects were made with the intention of avoiding the more unusual ones. Hence, the selection was limited to animals that are often found as (more or less) common pets or on a farm. Finally, for the category of inanimate direct objects, the nouns selected all designate material objects, thus steering clear of abstract concepts that can sometimes be attributed a certain amount of animacy due to metonymy.

As can be seen in Table 4, each sentence contains the possessive as part of the direct object in order to avoid possible influences of other languages such as, for example, Sardinian where the mere presence of the definite article can block the occurrence of DOM.¹⁰ The co-occurrence of possessive and definite article as seen in these sentences is an inherent characteristic of Catalan grammar.

Generally speaking, it should be emphasized that for each of these blocks of six variations of a test sentence, the direct objects were kept as similar in form as possible. Hence, the direct objects used within each of the blocks of six share the same gender and number of syllables. Each object, as well as each subject, was used only once throughout the experiment and the complements were not repeated. Furthermore, in an effort to keep possible interferences as small as possible, the gender combinations of subject and object were distributed equally.¹¹

As was the case for the test sentences of the ‘animacy’ part, those for the part about the position of the direct object were also designed with care. Although this experiment part was not as central to the study as the first one, the sentences were still designed with an eye to maintaining the possibility of attaining some information about DOM in Catalan. Consequently, the sentences were all constructed with the following basic structure: omitted subject, direct object in pronominal form (depending on the condition with or without a type of dislocation and DOM), transitive verb in *perfet perifràstic*, and circumstantial complement.

Table 5: Examples of test sentences for experiment part ‘position’.

	Topicalization with clitic doubling	Dislocation to the right with clitic doubling	Neither dislocation nor clitic doubling
2. SG	<u>A tu</u> <u>et</u> van saludar des del cafè del jardí botànic.	<u>Et</u> van saludar <u>a tu</u> des del cafè del jardí botànic.	Et van saludar des del cafè del jardí botànic.
3. SG	<u>A ell</u> <u>e</u> van saludar des del cafè del jardí botànic.	<u>El</u> van saludar <u>a ell</u> des del cafè del jardí botànic.	El van saludar des del cafè del jardí botànic.
2. SG	<u>A tu</u> <u>et</u> van vèncer en la final del torneig d’escacs.	<u>Et</u> van vèncer <u>a tu</u> en la final del torneig d’escacs.	Et van vèncer en la final del torneig d’escacs.
3. SG	<u>A ella</u> <u>la</u> van vèncer en la final del torneig d’escacs.	<u>La</u> van vèncer <u>a ella</u> en la final del torneig d’escacs.	La van vèncer en la final del torneig d’escacs.

¹⁰ Blasco Ferrer (1986) and Bossong (1982) both highlight this link between presence of the definite article and absence of DOM in Sardinian.

¹¹ From a total of 30 sentences, 7 have both feminine subject and object, 7 both masculine subject and object, and 8 feminine subject and masculine object, and vice versa.

Compared to those of the first part, these sentences feature a different basic sentence structure, but still allow for the possibility of observing the behaviour of DOM with personal pronouns. Hence, they were regarded as adequate considering the aims of the experiment. As a second manipulation, a change between 2nd and 3rd person singular pronouns was introduced, in order to see whether there were any observable differences in acceptability. In addition, the 3rd person singular forms alternated between feminine and masculine pronouns. Similarly to the sentences of the ‘animacy’ part, each verb occurs only once throughout the whole experiment and the complements vary as well. Finally, it should be noted that, although the second category shows a clear movement to the right of the direct object, it does not constitute a full movement to the right-most position in the sentence, as would be the prototypical case. Examples of this more extreme movement have been found in several Catalan dialects (cf. Rohlfs 1971; Escandell Vidal 2009). Unfortunately, including this specific type of sentence was not possible due to the design of the experiment and the need to keep the overall uniformity of the test sentences intact.

One further detail remains to be mentioned here. All test sentences were designed by the article’s author, who happens not to be a native speaker of Catalan. Nonetheless, prior to conducting the experiment, several native speakers from various regions checked them for their grammatical correctness.

5 The experiment: proceedings and participants

The experiment described in this article was designed for online participation and was set up using OnExp, a free program from the University of Göttingen. Participants began the experiment with a short introduction, in which they were familiarized with the general aims of the survey (without mentioning explicitly the exact topic under investigation). In a second step, they were introduced to the three stages of the experiment itself. To start with, participants were asked to provide some personal information, including age, origin, and speaking habits; this information was given anonymously. They were then shown the evaluation system used for the test sentences and were given the opportunity to become accustomed to it by practicing on eight exercise sentences designed to exemplify the different acceptability levels. The evaluation system consisted of an acceptability scale ranging from 1 to 7, with 1 being equal to ‘not at all natural’, and 7 ‘totally natural’.

Illustration 2: Acceptability Scale.

After these introductory sections, participants continued with the third phase: the experiment itself. They were shown each test sentence, one after another, in a random order, and were asked to give their opinion on its acceptability, using the scale shown above.

Before turning to the results of the experiment, some details about the participants should be discussed briefly. While some participants were obtained through personal contacts, the majority was reached via social media. The link leading to the survey was published on Twitter and in numerous groups and sites on Facebook. In addition, multiple cultural associations and platforms in each of the regions were contacted, as well as universities and language schools. The only prerequisite was that a participant should be a native speaker of one of the major dialects of Catalan.

From the six major regions a total of 305 participants completed the survey. Out of these, unfortunately, some had to be eliminated for various reasons. Some had more than one major dialect as a first language, making it impossible to attribute their responses clearly to just one region. Others skipped several test sentences, leading to them being rejected for the overall results. Some evaluated every sentence with the same extremely high or low value (always giving a 7, for example), leading to a distortion of the final overall scores.¹² Having eliminated these cases, it was also necessary to level the number of participants per list within every region, in order to avoid giving certain sentences more or less weight in the final regional analysis. Thus, a total of 260 participants was finally achieved. In the following Table the distribution of participants among regions and according to the more important demographic factors is displayed.

As can be seen, for the four regions of the bigger dialects, a sufficiently high number of participants were found to allow for a reliable analysis of their responses. Sadly, this was not the case for Alguerese and Rossellonese, where only a very small number of participants were registered, thus allowing for no more than the reflection of a few trends and some mostly speculative observations on their use of DOM. For the purpose of this paper, then, these have been excluded from further study, and will not be discussed in the following Section.

On the whole, it can be observed that the distribution of female and male participants is rather balanced (with the exception of Valencian, where approximately three quarters were male, and Majorcan, where two thirds were female). A significant demographic aspect was the question of whether the participants

¹² Due to their particular configuration of conditions, some of the test sentences were certain to receive very low or very high evaluations. Therefore, if a participant attributed the exact same value to every sentence, the possibility that he didn't take the experiment seriously was very high and thus he was excluded.

Table 6: Demographic distribution of participants.

	Total	Gender		Philoling ¹³		Age			Education		Mother tongue		
		F	M	No	Yes	≤25	26–50	51+	University	other	Cat.	Dial. + dialect	SP, I, F, S
Total	260	125	135	213	46	44	162	52	207	53	148	112	
Alguerese	11	5	6	10	1	0	2	9	9	2	3	8	
Central Cat.	90	39	51	73	17	14	57	19	74	16	52	38	
Majorcan	72	48	24	67	4	17	49	5	51	21	45	27	
N-W Catalan	36	18	18	22	14	6	20	9	33	3	28	8	
Rossellonese	9	3	6	8	1	0	5	4	7	2	2	7	
Valencian	42	12	30	33	9	7	29	6	33	9	18	24	

were students of linguistics or philology in general, which could have led them to deduce the nature of the experiment more easily. As is shown in columns 4 and 5, the great majority of participants, overall and also within each region, had not attended courses pertaining to these fields of study. While most participants were between 26 and 50 years old, it is an interesting fact that for Alguerese 9 out of 11 participants were older, some considerably. Again, the number of participants here is too small to draw any definitive conclusions.

Turning to the level of education of the participants, it should be noted that, both overall and within each region, there seems to be a very similar distribution of those having completed, or being in the process of completing, university studies. Overall 79.6% of participants attended university while 20.4% did not. This general distribution is reflected to a greater or lesser degree in each region (Majorcan exhibits a slightly lower percentage of university studies, and North-Western Catalan a slightly higher one).

The final category reflected in Table 6 concerns the native language of the participants. They were able to select from the following options: the six major Catalan dialects, plus Spanish, Italian, French and Sardinian. As noted above, when participants picked multiple Catalan dialects as their native language they had to be excluded from the final analysis. As for the total numbers, 56.9% (148 out of 260) only chose a Catalan dialect as their native language, whereas 43.1% (112 out of 260) indicated one or more of the other romance languages as well. Considering the individual regions, an interesting finding emerged: when participants from Central Catalan, Majorcan and North-Western Catalan were asked

¹³ Philoling: Abbreviation [SZ] indicating if participants were students of linguistics and/or philology in general.

for their mother tongue, a majority chose only their dialect. Valencian speakers, however, tended to choose not only their variety, but one of the other Romance languages as well (mainly Spanish).

6 The experiment: results

As noted above, each of the two main parts of the experiment should be considered separately, and for this reason the results here will be examined in different sections. The main focus will be on the first part of the experiment, containing the test sentences designed for ‘animacy’, but there will also be a brief description of the results for ‘position’. Due to the reasons described earlier, only the results from the four most important dialects (Central Catalan, North-Western Catalan, Majorcan and Valencian) will be shown, as well as the overall results and comparison of these four regions.

Before discussing the results themselves, a few details about the methods used in the analysis of the data should be noted. The main part of the analysis was carried out using the programs Excel and R (R Core Team 2019), because of the possibilities they offer for the calculation of mean values for each of the conditions and for the comparison of these based on different parameters. The results presented in the following are, as a rule, based on all the data of one region and the perceived effects therein. Furthermore, in addition to the descriptive analysis, a further analytical method was used, this as a first approach to a statistical analysis of the collected data. The model chosen was a repeated measure ANOVA (Analysis of Variance),¹⁴ which made it possible to ascertain whether or not the possible effects detected by the descriptive approach could be confirmed as statistically significant. When such effects could be confirmed to be statistically significant, then we can say that there was a very high probability that they had an impact on the acceptability of the different types of sentences tested in the experiment. Taking into account two of the main sources for possible bias, two sets of ANOVAs were performed for every region – one focusing on the participants (Subject-Analysis) and one for the test sentences (Item-Analysis). For a first statistical analysis of the collected data, ANOVA was deemed an appropriate choice, allowing for a general impression of the data’s statistical value. In future work, alternative models could be consulted as a means of achieving an even bigger picture.

¹⁴ The ANOVAs were conducted in R, using the *ez* package (Lawrence 2016).

A final aspect to note is that for the overall results, when combining data from the four main regions, each mean value was calculated based on the 1,200 judgments that the respective condition had received (as a result of the total of 240 participants, each evaluating every condition five times).

6.1 Animacy

In this Section, the results of the four regions will be introduced separately in order to better illustrate the possible effects of the manipulated factors on the perceived acceptability of the test sentences. The factor of regional variance will then be discussed by presenting the overall results concerning the first main aspect, ‘animacy’, including a comparison of the four major Catalan dialects. Before starting with the dialect with the highest participation (Central Catalan), a reminder of the focus of the first experiment seems useful. Its main manipulated aspect was the degree of animacy of the direct object in combination with the presence or absence of DOM. The sentences presented to the participants in a completely random order were of the following type: *El guàrdia va transportar al/ el meu nét/gat/sac fins a l'estació de ferrocarrils*.

6.1.1 Regional results for ‘animacy’

In the following Graph, the results obtained from the participation of 90 Central Catalan speakers are illustrated, equalling 450 separate evaluations of each condition. All the values depicted in this Graph, and in those that follow, are means of the aggregated z-scores calculated from the acceptability judgments that each condition had received.¹⁵ They are shown with a Confidence Interval of 95%, represented by the error bars of each mean value. On the horizontal axis the six conditions are represented, showing the three degrees of animacy (human, animal, inanimate) and split into two blocks according to the presence or absence of the second factor, DOM.

Focusing on the presence/absence of DOM, overall, Central Catalan speakers clearly prefer the sentences without the preposition *a*, since C4–C6 show higher values than C1–C3. Nevertheless, the three less-favoured conditions still seem

¹⁵ The decision to use mean z-scores instead of normal mean judgments is based on the fact that they allow for a fairly accurate comparison of the participants, taking into account the possible individual interpretations of the acceptability scale.

Graph 1: Central Catalan: results for ‘animacy’.

to be at the least to some extent acceptable, bearing in mind that especially C1 (humans), but also C2 (animals), reach relatively high mean z-scores. Considering the fact that all three pairs of objects (C1 and C4, C2 and C5, C3 and C6) show mean values whose error bars more and more clearly do not overlap, there is a high probability that the presence/absence of DOM has an effect on the acceptability of a given sentence. Although Central Catalan speakers tend towards the versions without DOM, Graph 1 indicates that mainly for human objects (C1 vs. C4) there is a relatively high acceptance of the version with DOM.

Furthermore, it appears that the participants regard the animacy of the direct object as an influential factor for the acceptability of DOM. Looking at the conditions containing DOM (C1–C3), the human object is distinctly favoured over animals, which in turn are preferred over inanimate marked objects. Since once again neither mean values nor error bars overlap, it can be assumed that the direct object’s animacy plays an important part when deciding about the acceptability of a sentence.

The two ANOVAs confirmed both of the effects described for Central Catalan. Furthermore, not only are the effects of DOM and ‘animacy’ manipulation separately statistically significant, but their interaction is as well, implying that they interactively influence the acceptability of any given sentence.¹⁶

16 ANOVA results for Central Catalan:

1. Subject-Analysis for the effects of DOM: $F(1, 89) = 179.483, p < .001$ and ‘animacy’: $F(2, 178) = 57.317, p < .001$ as well as their interaction: $F(2, 178) = 36.327, p < .001$ on the acceptability of the test sentences.

For the analysis of the Majorcan data, the evaluations of 72 native speakers were taken into account, resulting in a total of 360 judgments per condition. An important detail here is the fact that one participant pointed out that he would have given higher votes if the sentences had included the proper definite articles of his dialect (*es, sa*). Earlier, an explanation was given as to why it was not feasible to incorporate such regional variations. Even so, it is worth noting that Majorcan speakers' acceptance of sentences containing DOM could be even higher than it already is, considering the following Graph.

Graph 2: Majorcan: results for 'animacy'.

Again, the factor 'animacy' appears to have a certain amount of influence over the perceived acceptability of the sentences including DOM, since the mean values of C1–C3 clearly differ from each other and their error bars do not overlap at all. Unsurprisingly, the human object was preferred over the animals and inanimate objects. Remarkably, the z-scores for human objects with DOM (C1) and for the same type of object without DOM (C4) are nearly the same (with a difference of 0.001). This might suggest that Majorcan speakers do not have a preference for one or the other, and thus do not care whether or not a human object is marked. They seem to find both versions equally acceptable. Even for the objects representing animals (C2 and C5) the gap between the two values is quite small. Only when looking at the inanimate objects, a strong partiality for the version without DOM (C3 vs. C6) can be seen.

2. Item-Analysis for the effects of DOM: $F(1, 29) = 267.394$, $p < .001$ and 'animacy': $F(2, 58) = 29.354$, $p < .001$ as well as their interaction: $F(2, 58) = 26.432$, $p < .001$ on the acceptability of the test sentences.

ANOVAs were used to determine whether these possible effects found in the descriptive approach were of any statistical value. The calculations confirmed, both for ‘animacy’ and for DOM, the very high probability of an effect on the acceptability of a given sentence, showing these, as well as their interaction, to be statistically highly significant.¹⁷

Continuing with the data from North-Western Catalan, it should be noted that this was the region with the smallest number of participants (apart from Rosellonese and Alguerese), amounting only to 36, which still renders 180 evaluations per condition. Looking at the mean values depicted in Graph 3, speakers of North-Western Catalan seem to accept DOM with objects denoting humans far more often than when they refer to animals.

Graph 3: North-Western Catalan: results for ‘animacy’.

The lowest mean value is attributed to the inanimate objects marked by *a*. In addition, it is highly likely that ‘animacy’ has an effect on the acceptability of DOM for this region because the three values (C1–C3) show plainly visible differences and their error bars do not overlap at all. The three conditions representing absence of DOM (C4–C6) are at a relatively high (and similar) level, indicating that ‘animacy’

17 ANOVA results for Majorcan:

1. Subject-Analysis for the effects of DOM: $F(1, 71) = 52.534, p < .001$ and ‘animacy’: $F(2, 142) = 61.369, p < .001$ as well as their interaction: $F(2, 142) = 52.984, p < .001$ on the acceptability of the test sentences.
2. Item-Analysis for the effects of DOM: $F(1, 29) = 55.894, p < .001$ and ‘animacy’: $F(2, 58) = 28.429, p < .001$ as well as their interaction: $F(2, 58) = 26.39, p < .001$ on the acceptability of the test sentences.

does not come into play regarding the acceptance of sentences without DOM. The non-marked object is clearly favoured overall by North-Western Catalan speakers, with the human objects being the only conditions where the values are somewhat close to each other (C1 vs. C4). Furthermore, the error bars of each pair do not overlap, indicating that presence/absence of DOM very probably affects the acceptability of a sentence.

These tendencies are confirmed by both ANOVAs, producing a statistically very significant result for DOM and for ‘animacy’, as well as their combination.¹⁸ Thus, it seems highly probable that both separately and interactively they have an impact on the acceptance of the sentences.

Lastly, for Valencian a total of 42 participants answered the questionnaire, yielding 210 evaluations per condition overall and allowing for a fairly reliable analysis of the data.

Graph 4: Valencian: results for ‘animacy’.

Valencians seem to distinguish clearly between the individual degrees of animacy when the object is combined with DOM (C1–C3). Considering their z-scores, we note that the human object is deemed quite acceptable, followed

¹⁸ ANOVA results for North-Western Catalan:

1. Subject-Analysis for the effects of DOM: $F(1, 35) = 41.411$, $p < .001$ and ‘animacy’: $F(2, 70) = 25.891$, $p < .001$ as well as their interaction: $F(2, 70) = 15.365$, $p < .001$ on the acceptability of the test sentences.

2. Item-Analysis for the effects of DOM: $F(1, 29) = 106.245$, $p < .001$ and ‘animacy’: $F(2, 58) = 26.484$, $p < .001$ as well as their interaction: $F(2, 58) = 9.229$, $p < .001$ on the acceptability of the test sentences.

by animals, which are still seen as somewhat acceptable. Then there is a considerable gap before the marked inanimate objects, these being regarded as decidedly less normal. Given that the error bars do not overlap, there appears to be an effect of animacy manipulation on the acceptability of the sentences. As was the case in the other regions, the conditions without DOM (C4–C6) are judged to be quite equally acceptable by Valencian speakers overall.

Similarly to Majorcan, Valencians do not exhibit a clear preference when the direct object involves humans and presence/absence of DOM (C1 vs. C4). The respective z-scores differ only by the smallest of margins (0.001), indicating the participants' indifference. For the other two categories, the values show that Valencians lean towards the versions without DOM (although in case of animals only slightly). For C2 and C5 (animals) as well as C3 and C6 (inanimate things) the error bars do not overlap, indicating a possible effect of DOM manipulation on the acceptance of a given sentence.

For Valencian, the ANOVA calculations gave statistically significant results for both the effect of 'animacy' and DOM, as well as their interaction, thus confirming the high probability of their influence on the acceptability of the test sentences.¹⁹

6.1.2 Overall results for 'animacy' and a comparison of the four regions

Having discussed the results for each region separately, including the possible and probable effects of 'animacy' and DOM on the acceptability of sentences, we turn now to the overall results and a comparison of the four regions.

The comparison of the results of the four regions shows that all of them follow a similar general pattern of acceptance, suggesting no large regional effect on the acceptability of the sentences. Nonetheless, some variation can be found. The participants evidently favour human marked objects (C1) over animals (C2) and over inanimate objects (C3), whereas the conditions without DOM (C4–C6) all appear to have quite similar levels of acceptance. Interestingly, Majorcan and Valencian are the dialects that more freely accept DOM not only with humans but

¹⁹ ANOVA results for Valencian:

1. Subject-Analysis for the effects of DOM: $F(1, 41) = 40.534, p < .001$ and 'animacy': $F(2, 82) = 27.768, p < .001$ as well as their interaction: $F(2, 82) = 34.642, p < .001$ on the acceptability of the test sentences.

2. Item-Analysis for the effects of DOM: $F(1, 29) = 91.891, p < .001$ and 'animacy': $F(2, 58) = 18.248, p < .001$ as well as their interaction: $F(2, 58) = 24.895, p < .001$ on the acceptability of the test sentences.

Graph 5: Overall comparison of the results for ‘animacy’.

also with animals. Overall, Majorcan shows the highest mean values, even for the inanimate objects (C3). This confirms to some extent what can be found in the literature on the subject. According to several authors, in Majorcan even inanimate direct objects with Differential Object Marking can be found.²⁰

Central and North-Western Catalan behave very similarly to one another, not only in their lower acceptance of sentences with DOM, but also in their higher acceptance of those without (C4–C6). Furthermore, both dialects clearly favour the non-marked objects in every category, whereas Majorcan and Valencian display a different behaviour. As noted earlier, and illustrated in Graph 5, these latter two dialects do not seem to mind whether human direct objects are marked or not – they accept both versions equally (C1 and C4 are virtually at the same level).

A curious detail is the fact that all four dialects seem to consider non-marked animals (C5) slightly less acceptable than their human and inanimate counterparts (C4 and C6). This might be due to the animals chosen for the test sentences, although when designing the experiment, the intention was to avoid less common animals. This aspect could be explored further by dividing the animal category into subcategories and thus observing whether there was any correlation between animal type and acceptability level.

²⁰ Examples of these objects with DOM are found, among others, in Rohlfs (1971, 323): *a ses patates, else pelarem/else pelarem*, *a ses patates* and Wheeler/Yates/Dols (1999, 462): *A ses tovalloles, posa les dins es calaix*.

Generally speaking, and as it pertains to the acceptability of DOM, the four Catalan varieties appear to behave more or less as might be expected²¹ – favouring the most animated objects. Nonetheless, it is worth noting that whereas the literature on Catalan DOM usually emphasizes the importance of ambiguity as a deciding factor in the use of DOM, these results suggest that it actually might not be so important, at least for the human object. None of the sentences tested during the experiment show any ambiguity whatsoever as to which elements hold the function of subject and direct object. Yet participants still judged the sentences with DOM to be highly acceptable – sometimes even as acceptable as the non-marked ones, as in the case of Majorcan and Valencian.

Before moving on to the second part of the experiment, a few details about the overall results for ‘animacy’ ought to be noted. In Graph 6 the mean z-scores represent the combination of the evaluations from all four regions, which resulted in a total of 240 participants, equalling 1,200 judgments per condition.

Graph 6: Overall results for ‘animacy’.

Unsurprisingly, the general shape and form of the overall distribution of mean values is similar to the regional ones. The overall results present the same scalar distribution for the marked objects (C1–C3), with human objects being seen as more acceptable than animals and inanimate ones. The error bars do not overlap, which is a strong indication of the possible effect of ‘animacy’ on perceived acceptability.

²¹ From a general theoretical viewpoint, it seems that the higher up on the animacy-scale, the more frequently a direct object is marked (cf. Bossong 1991; Aissen 2003).

Regarding the three pairs (C1 and C4, C2 and C5, C3 and C6) globally, there is also a more or less clear preference for the unmarked object, though for the human object (C1 vs. C4) the predilection is not as pronounced as for the others. Looking at the respective error bars for all pairings, the fact that they do not overlap suggests an effect of DOM on the acceptability of a given sentence. An interesting finding in terms of this global distribution is the fact that C1 and C4 have very similar levels of acceptance, as indicated earlier. This implies that in nominal clauses with human direct objects, DOM is rather well accepted among Catalan speakers on the whole. In fact, they accept it nearly as much as the version that is prescribed by the normative grammar – the one without DOM. Thus, it can be concluded that in this type of sentence, participants did not mind whether or not the object was marked; they seemed to find both options normal.

For the overall analysis, the ANOVAs yielded statistically significant values, and thus confirmed the high probability of an effect of DOM and of ‘animacy’, as well as their interaction.²² In this global study of the four regions, the ANOVAs were used to test for another effect: regional variance. Although, as seen in Graph 5 under the descriptive approach, some slight regional differences can be perceived, the statistical analysis does not show a significant value for this effect in both Subject- and Item-Analysis.²³ This implies, with very high probability, that the factor ‘region’ did not have a huge impact on the evaluations and hence that the acceptability of DOM in this type of sentence is not subject to many diatopic factors.

As discussed earlier, the figures shown here only include the results for the four most prevalent Catalan dialects and thus exclude Rossellonese and Alguerese. Considering their geographic situation, it is possible that with further data from these two regions there could be a higher diatopic impact on the acceptability of DOM. But at the moment this remains no more than speculation, and would have to be investigated with the help of more participants.

22 ANOVA results for overall analysis:

1. Subject-Analysis for the effects of DOM: $F(1, 236) = 253.784, p < .001$ and ‘animacy’: $F(2, 472) = 149.916, p < .001$ as well as their interaction: $F(2, 472) = 127.685, p < .001$ on the acceptability of the test sentences.

2. Item-Analysis for the effects of DOM: $F(1, 29) = 246.266, p < .001$ and ‘animacy’: $F(2, 58) = 49.692, p < .001$ as well as their interaction: $F(2, 58) = 47.062, p < .001$ on the acceptability of the test sentences.

23 ANOVA results for the overall analysis:

1. Subject-Analysis for the effect of ‘region’: $F(3, 236) = 6.025, p < .001$ and its interaction with DOM and ‘animacy’: $F(6, 472) = 1.693, p = 0.121$ on the acceptability of the test sentences.

2. Item-Analysis for the effect of ‘region’: $F(3, 87) = 1.816, p = 0.150$ and its interaction with DOM and ‘animacy’: $F(6, 174) = 1.125, p = 0.349$ on the acceptability of the test sentences.

6.2 Position

As noted above, the second experiment part focused on ‘position’, and participants were presented with sentences of the following structure: *A tu et van vèncer/Et van vèncer a tu/Et van vèncer en la final del torneig d’escacs*. Here, the main manipulated factors were the position of the direct object given as a personal pronoun, and its corresponding clitic doubling combined with a change between 2nd and 3rd person singular pronouns.

6.2.1 Regional results for ‘position’

The graphs for these results are very similar in layout to those in the previous Section; the only difference lies in the horizontal axis, on which the corresponding conditions (C11–C16) were plotted with their respective combination of the manipulated characteristics.²⁴ First, in Graph 7 the results for Central Catalan are presented, where participants did not seem to show any concern for the change in personal pronouns. When comparing the respective pairs of conditions (C11 vs. C14, C12 vs. C15 and C13 vs. C16) very little difference can be seen in the levels of their mean z-scores, implying that the change between 2nd and 3rd person singular pronouns does not have any effect on the acceptability of a sentence.

Graph 7: Central Catalan: Results for ‘position’.

²⁴ Abbreviations on the horizontal axis: topicalization (top), dislocation to the right (disr), clitic doubling (cd), absence of any type of dislocation and thus also clitic doubling (ø). The boxes show the split between 2nd and 3rd person singular pronouns.

On the other hand, when looking at the position of the direct object and the corresponding clitic doubling, there appears to be a certain effect on acceptability, since the values are clearly separated and their error bars do not overlap (comparing C11, C12, C13 and C14, C15, C16). Central Catalan speakers overwhelmingly prefer the sentences without any type of dislocation (C13 and C16). This in itself is not very surprising, considering that C13 and C16 represent the most baseline kind of sentence tested in this part of the experiment, consisting of short sentences without any type of modification. Of the two dislocations, Central Catalan speakers lean towards the sentences with topicalization, although they apparently accept both versions containing DOM to a somewhat limited degree.

The conducted ANOVAs indicate, as expected, that the change in the personal pronoun did not have an effect on the acceptability of the test sentences (as reflected by its non-significant result). However, the statistically significant results confirm the probability of an effect of the factor ‘position’.²⁵

Continuing with the Majorcan data, the following Graph again plainly highlights that overall participants favoured the most basic type of test sentence

Graph 8: Majorcan: Results for ‘position’.

25 ANOVA results for Central Catalan:

1. Subject-Analysis for the effects of ‘position’: $F(2, 178) = 191.386, p < .001$ and ‘ppron’: $F(1, 89) = 2.483, p = .119$ as well as their interaction: $F(2, 178) = 0.487, p = .615$ on the acceptability of the test sentences.

2. Item-Analysis for the effects of ‘position’: $F(2, 58) = 201.39, p < .001$ and ‘ppron’: $F(1, 29) = 0.474, p = .497$ as well as their interaction: $F(2, 58) = 0.316, p = .73$ on the acceptability of the test sentences.

without any manipulation. Nevertheless, the other two possible sentence structures show relatively high z-scores as well.

The change in personal pronoun does not seem to affect acceptability, given that the respective pairs of mean values are on very similar levels and their error bars overlap. Regarding the different types of position of the direct object, it is possible that there is a certain effect. Comparing only the two dislocations (both with DOM), Majorcan speakers show a slight preference for topicalization (C11 and C14) while they do not reject C12 and C15 as much as the Central Catalan speakers. This would confirm, at least partly, what was found in Escandell Vidal (2009) about DOM in Majorcan, although she points out that in this dialect dislocation to the right is even more accepted than it appears to be the case for the current study. This is very likely due to the fact that in the experiment presented here, the sentences contained only partial movements to the right and not prototypical dislocations to the extreme right of the sentence, due to the need for strict uniformity of sentence structures.²⁶

By conducting ANOVAs, the possible effect of ‘position’ can be confirmed, given that a statistically significant result was found. Meanwhile, the change in personal pronoun and the interaction of the two factors did not seem to have an effect on the acceptability of the test sentences.²⁷

Similarly, to the Majorcan and Central Catalan data, the speakers of North-Western Catalan prefer C13 and C16 without any manipulations, but the difference to the other two sentence variations is not as great as it was for the Central region. The average values, however, are sufficiently separate to assume the existence of an effect of the direct object’s position on the acceptability of the sentences (comparing C11, C12, C13 and C14, C15, C16). Both types of dislocation (with clitic doubling and DOM) seem to be deemed quite acceptable by the North-Western Catalan speakers, with a slight preference towards topicalization. As was the case for the first two dialects, the changed personal pronouns evidently appear to have no effect on the judgments of the participants.

26 A further experiment, including sentences with dislocations to the extreme right, would show more clearly to what degree Majorcans accept this type of dislocation.

27 ANOVA results for Majorcan:

1. Subject-Analysis for the effects of ‘position’: $F(2, 142) = 68.357, p < .001$ and ‘ppron’: $F(1, 71) = 0.986, p = .324$ as well as their interaction: $F(2, 142) = 0.029, p = .971$ on the acceptability of the test sentences.

2. Item-Analysis for the effects of ‘position’: $F(2, 58) = 117.522, p < .001$ and ‘ppron’: $F(1, 29) = 0.766, p = .389$ as well as their interaction: $F(2, 58) = 0.034, p = .966$ on the acceptability of the test sentences.

Graph 9: North-Western Catalan: Results for ‘position’.

This is confirmed by the non-significant result that the ANOVAs yield for this factor. For the aspect ‘position’, then, the model confirms the very high probability of an effect on the acceptability of the sentences.²⁸

Valencian speakers, apparently, judged all conditions to be more or less acceptable, as can be seen in the following Graph. Again, the participants did not seem to pay attention to the change in personal pronoun, since the three respective pairs of mean values are at very similar levels.

On the other hand, ‘position’ did appear to have a certain influence on the acceptability of the sentences, given that in neither group the values, or their error bars, overlap. Finally, it seems that Valencian speakers accepted both types of dislocation (with clitic doubling and DOM) quite readily, although with a tendency towards topicalization. As was the case with the other regions, participants favoured C13 and C16, which is unsurprising bearing in mind that they represent short sentences without any manipulation.

28 ANOVA results for North-Western Catalan:

1. Subject-Analysis for the effects of ‘position’: $F(2, 70) = 40.835$, $p < .001$ and ‘ppron’: $F(1, 35) = 0.63$, $p = .433$ as well as their interaction: $F(2, 70) = 0.513$, $p = .601$ on the acceptability of the test sentences.

2. Item-Analysis for the effects of ‘position’: $F(2, 58) = 69.143$, $p < .001$ and ‘ppron’: $F(1, 29) = 0.544$, $p = .467$ as well as their interaction: $F(2, 58) = 0.662$, $p = .52$ on the acceptability of the test sentences.

Graph 10: Valencian: Results for ‘position’.

Furthermore, both ANOVAs once again only confirmed the effect of the change in position on the acceptability of the test sentences.²⁹

6.2.2 Overall results for ‘position’ and a comparison of the four regions

Having discussed the regional results concerning the factor ‘position’ separately, we will now take a look at the four dialects together. One of the most obviously notable things here is the fact that these distributions are not quite as uniform and continuous as those in the comparison graph for the ‘animacy’ results (Section 6.1.2, Graph 5), which suggests a possible regional effect.

As we have already seen in the regional analysis, the change in personal pronoun (between 2nd and 3rd person singular) did not seem to have any effect on the acceptability of the sentences, considering that the respective z-scores are all at very similar levels (this is even better illustrated in Graph 12, containing the global results).

²⁹ ANOVA results for Valencian:

1. Subject-Analysis for the effects of ‘position’: $F(2, 82) = 44.876, p < .001$ and ‘ppron’: $F(1, 41) = 1.67, p = .203$ as well as their interaction: $F(2, 82) = 0.076, p = .926$ on the acceptability of the test sentences.
2. Item-Analysis for the effects of ‘position’: $F(2, 58) = 77.204, p < .001$ and ‘ppron’: $F(1, 29) = 0.473, p = .497$ as well as their interaction: $F(2, 58) = 0.125, p = .883$ on the acceptability of the test sentences.

Graph 11: Overall comparison of the results for ‘position’.

On the other hand, an interesting finding here is that all four major Catalan dialects concur as to which sentences they prefer the most: C13 and C16, both without either form of dislocation and both only differing in the form of the personal pronoun. The mean values for C16 are almost identical, whereas for C13 they are also very close. This predilection for the most basic sentence is, of course, not unexpected, although it still seems quite intriguing that the results of all four regions are so similar.

Concerning the other sentence variations (all conditions with DOM), there is slightly more regional variation. Graph 11 illustrates that Valencians, compared to the other speakers, tended to give the test sentences higher evaluations, whereas the Central Catalan participants, overall, rated them the lowest. However, the exceptions here are C13 and C16. Interestingly, for these two conditions the averages of Central Catalan surpass those of the other regions, whereas for every other condition they are clearly lower than the rest. In other words, the difference between the two conditions with low average values (C11 and C12) and C13 is far greater for Central Catalan than for the other regions. This is also true for the comparison of C16 to C14 and C15. Of all the regions, participants from Central Catalan appear to be the firmest in their preference for the sentences that are free of any type of dislocation of the direct object.

As for the conditions with topicalization (C11 and C14) and dislocation to the right (C12 and C15), Graph 11 illustrates that the distribution of Valencian and Central Catalan is more or less parallel, while Majorcan and North-Western Catalan present more variation. As can be seen, for C11 (topicalization) the North-Western z-scores are somewhat higher than those for Majorcan, whereas Majorcan presents slightly higher values for C15 (dislocation to the right). For C12

and C14, both dialects have very similar mean scores. A final point to raise here is the fact that of the two possibilities with dislocations (and DOM), all the regions clearly lean towards the sentences including topicalization. Again, there is a high probability that this is due in part to the fact that the dislocations to the right were not total, but only partial ones. Interestingly, these partial dislocations still attain a certain degree of acceptance, bearing in mind that they are not the prototypical dislocations to the right, for which there are numerous examples, as indicated in Rohlf's (1971) and Escandell Vidal (2009), among others.

Following the comparison of the four regions, Graph 12 below depicts the overall mean values for the combination of the total of 240 participants throughout the four regions (equalling 1,200 judgments per condition). The Graph illustrates perfectly that the change in personal pronoun did not appear to have any effect on the acceptability of the individual conditions, given that the corresponding averages are virtually the same (C11 vs. C14, C12 vs. C15, C13 vs. C16). This, not surprisingly, is confirmed by the ANOVAs, which did not show a statistically significant value for the change in personal pronoun.³⁰

Graph 12: Overall results for 'position'.

30 ANOVA results for overall analysis:

1. Subject-Analysis for the effects of 'position': $F(2, 472) = 268.8, p < .001$, 'ppron': $F(1, 236) = 1.633, p = .203$ and 'region': $F(3, 236) = 6.025, p < .001$ as well as the interaction between 'region' and 'position': $F(6, 472) = 4.718, p < .001$ on the acceptability of the test sentences.
2. Item-Analysis for the effects of 'position': $F(2, 58) = 239.729, p < .001$ and 'ppron': $F(1, 29) = 0.367, p = .55$ and 'region': $F(3, 87) = 20.872, p < .001$ as well as the interaction between 'region' and 'position': $F(6, 174) = 7.343, p < .001$ on the acceptability of the test sentences.

On the other hand, looking at the position of the direct object, there does seem to be an effect. Within both groups (C11, C12, C13 and C14, C15, C16) the averages differ clearly and their error bars do not overlap. Hence, it is possible that the position of the direct object may have a certain influence on the acceptability of a given sentence. The sentences, it appears, were more acceptable when they contained a topicalization of the direct object than when there was a dislocation to the right. However, as can be seen in Graph 12, in overall terms, participants preferred the most basic version without dislocation.

These findings derived from the descriptive approach are confirmed by the statistical analysis by means of the two ANOVAs. The resulting values for the factor ‘position’ are statistically highly significant, suggesting that there is indeed such an effect on the evaluation of the sentences. The final aspect explored through the ANOVAs was the possible effect of the factor ‘region’ on the acceptability of the sentences. As implied in Graph 11 by the visible differences in the four distributions of the mean values, the statistically significant results of both ANOVAs confirm that the factor ‘region’ is highly likely to have an impact on the acceptability of a given sentence. Furthermore, the combination of ‘region’ and ‘position’ is confirmed to have had some influence on the perceptions of the various conditions presented to the participants.

7 Conclusions

The aim of this article has been to illustrate that Catalan in general, and its dialects in particular, are an independent part of the puzzle constituted by the individual types of DOM found in languages across the world. After a long period of being dismissed as a mere Spanish imposition, DOM is now beginning to be seen as an inherent part of the Catalan language itself. Nonetheless, there still are certain discrepancies between the uses permitted by Catalan normative grammars and the current situation found in Catalan dialects, as the experiment described in this article was able to highlight. Based on the results discussed in the previous Sections, it can be said that all Catalan dialects, to a greater or lesser extent, accept DOM in more contexts than the standard language commonly permits.

Furthermore, the factor ‘animacy’ appears to have quite a substantial impact on the acceptability of DOM with noun phrases, both for Catalan in general as well as for each of the dialects. In each case there was a pronounced scalar distinction in the perceived acceptability of the direct objects’ three animacy levels, with the human object reaching a considerably high level throughout the regions. In addition, in this type of sentence, and contrary to what standard Catalan dictates, it

appears that the total lack of ambiguity does not impede speakers' acceptance of DOM in the least. The only major regional variation in this part of the experiment was found for the human objects. Both Valencian and Majorcan seem to perceive the use of DOM with human direct objects in noun phrases as facultative, accepting its presence and absence at equally high levels. Central and North-Western Catalan, by contrast, generally prefer the non-marked version.

Concerning the second factor, 'position', on the whole the position of the direct object had an effect on the acceptability of the sentences, with all the dialects tending towards the most basic sentences as the most acceptable ones and topicalization as the more acceptable version containing DOM. Interestingly, in this part of the experiment, the regional differences were more pronounced than in the first part.

In conclusion, both main factors analyzed in this experiment, 'animacy' and 'position', appear to have an effect on the acceptability of DOM in the Catalan varieties, both on a general as well as on a regional level. Moreover, from a statistical point of view, the ANOVAs confirm the existence of both effects, as well as an interaction effect between DOM and 'animacy'. The third main factor, regional variance, was only found to be of statistical significance for 'position', thus indicating that in some contexts diatopic factors have an impact on the acceptability of DOM in Catalan, whereas in others they do not.

As mentioned in the opening sections of this paper, some of the literature suggests that Catalan and Spanish DOM share several characteristics. The results of the present experiment seem to indicate that this is true to a certain degree, especially for the three marked objects in the analysis of 'animacy', which show a similar distribution to Spanish.³¹ However, the results also reaffirm the inherently independent evolution of Catalan DOM, considering that the three unmarked objects in the 'animacy' section behave in a wholly dissimilar way to Modern Spanish DOM (cf. Wall 2015). In this study the results suggested that in Modern Spanish, unmarked human objects in particular are considered not at all acceptable; and this, in turn, underlines the individuality of the Modern Catalan DOM configuration.

There are several ways in which the current research might be extended and complemented. Based on the existing data, the possible influence of some of the demographic aspects described could be explored, although for statistically rel-

31 That is not to say that the Catalan configuration is solely due to Spanish influence, although geographic proximity suggests at least some interaction between the two languages (as can be expected in any situation of adjacent languages). Rather, it seems that the evolution of Catalan, for this type of marked objects, is slowly going in the same direction as that of Spanish, reflecting a perfectly normal evolution of a DOM system.

evant conclusions a larger dataset from each region would be required. Indeed, in terms of collecting more data, further material on Rossellonese and Alguerese would help to complete the whole picture of DOM in the Catalan dialects. This would be especially interesting, taking into account that in these regions it might be possible to observe influences from other Romance languages which do not show DOM, or which do so only on a more limited scale. As noted earlier on, further subcategories of the animacy of direct objects, such as multiple animal categories or further inanimate categories, could also be investigated, as well as other categories for the second part of the experiment, including total dislocations to the right. In conclusion, it is safe to say that there is still far more to explore towards gaining a complete picture of DOM in Catalan, both in general and in all its different dialects.

Bibliography

- Aissen, Judith, *Differential Object Marking. Iconicity vs. economy*, *Natural Language and Linguistic Theory* 21 (2003), 435–483.
- AVL = Acadèmia Valenciana de la Llengua, *Gramàtica normativa valenciana*, València, Publicacions de l'Acadèmia Valenciana de la Llengua, 2006.
- Badia Margarit, Antoni M., *Gramàtica catalana*, Madrid, Gredos, 1962.
- Blasco Ferrer, Eduardo, *La lingua sarda contemporanea. Grammatica del logudorese e del campidanese. Norma e varietà dell'uso. Sintesi storica*, Cagliari, Della torre, 1986.
- Boladeras Taché, Josep M., *Les preposicions en català nord-occidental*, Lleida, Pagès, 2011.
- Bosson, Georg, *Der präpositionale Akkusativ im Sardischen*, in: Winkelmann, Otto/Braisch, Maria (edd.), *Festschrift für Johannes Hubschmid zum 65. Geburtstag*, Bern/München, Francke, 1982, 579–599.
- Bosson, Georg, *Differential Object Marking in Romance and beyond*, in: Wanner, Dieter/Kibbee, Douglas A. (edd.), *New analyses in Romance linguistics. Selected papers from the Linguistic Symposium on Romance Languages XVIII, Urbana-Champaign, April 7–9, 1988*, Amsterdam/Philadelphia, John Benjamins, 1991, 143–170.
- Colón Domènech, Germà, *Català. Del general al particular*, in: Kabatek, Johannes/Pusch, Claus D. (edd.), *Variació, poliglòssia i estàndard. Processos de convergència i divergència lingüístiques en català, occità i basc*, Aachen, Shaker, 2009, 21–32.
- Costa Carreras, Joan, *Norma i procés d'estandardització*, in: Pradilla, Miquel Àngel (ed.), *Societat, llengua i norma. A l'entorn de la normativització de la llengua catalana*, Benicarló, Alambor, 2001, 41–55.
- Escandell Vidal, Victoria, *Differential Object Marking and topicality. The case of Balearic Catalan*, *Studies in Language* 33:4 (2009), 832–885.
- Fabra, Pompeu, *Sobre diferents problemes pendents en l'actual català literari*, *Anuari de l'Institut d'Estudis Catalans MCMVII* (1908), 352–369.
- Fabra, Pompeu, *Gramàtica catalana*, Barcelona, Institut d'Estudis Catalans, ⁷1933.

- Institut d'Estudis Catalans, *Gramàtica de la llengua catalana*, Barcelona, Institut d'Estudis Catalans, 2016.
- Irimia, Monica A./Pineda, Anna, *Differential Object Marking and scales. Insights from Romance diachrony*, Proceedings of the Linguistic Society of America 4:57 (2019), 1–15, doi. org/10.3765/plsa.v4i1.4561 [last access: 29.04.2019].
- Kremnitz, Georg (ed.), *Sprachen im Konflikt. Theorie und Praxis der katalanischen Soziolinguisten*, Tübingen, Gunter Narr, 1979.
- Lagarde, Christian, *De la norma a la queta. L'evolució de la normativització a Catalunya*, in: Kabatek, Johannes/Pusch, Claus D. (edd.), *Variació, poliglòssia i estàndard. Processos de convergència i divergència lingüístiques en català, occità i basc*, Aachen, Shaker, 2009, 313–326.
- Meier, Harri, *O problema do acusativo preposicional no catalão*, Boletim de Filologia 8 (1945), 237–260.
- Moll, Francesc de B., *Gramàtica catalana. Referida especialment a les Illes Balears*, Palma de Mallorca, Moll, 1979.
- Müller, Bodo, *Das morphemmarkierte Satzobjekt der romanischen Sprachen (Der sogenannte präpositionale Akkusativ)*, Zeitschrift für romanische Philologie 87:5 (1971), 477–519.
- Par, Anfós, *Sintaxi catalana. Segons los escrits en prosa de Bernat Metge (1398)*, Halle, Niemeyer, 1923.
- Pensado, Carmen, *La creación del complemento directo preposicional y la flexión de los pronombres personales en las lenguas románicas*, in: Pensado, Carmen (ed.), *El complemento directo preposicional*, Madrid, Visor, 1995, 179–233.
- Rohlf, Gerhard, *Autour de l'accusatif prépositionnel dans les langues romanes*, Revue de Linguistique Romane 35 (1971), 312–334.
- Sancho Cremades, Pelegrí, *La categoria preposicional*, Valencia, Universitat de València, 1995.
- Sancho Cremades, Pelegrí, *La preposició i el sintagma preposicional*, in: Solà, Joan, et al. (edd.), *Gramàtica del català contemporani*, vol. 2, Barcelona, Empúries, 2008, 1689–1796.
- Vallverdú, Francesc, *La normalització lingüística a Catalunya*, Barcelona, Laia, 1979.
- von Heusinger, Klaus/Kaiser, Georg A., *The evolution of Differential Object Marking in Spanish*, in: von Heusinger, Klaus/Kaiser, Georg A./Stark, Elisabeth (edd.), *Proceedings of the workshop "Specificity and the Evolution/Emergence of Nominal Determination Systems in Romance"*, Konstanz, Fachbereich Sprachwissenschaft der Universität Konstanz, 2005, 33–69.
- Wall, Albert, *The role of variation in the processing of Differential Object Marking (DOM) in Spanish*, Poster at the 48th annual meeting of the SLE, Leiden, 2015.
- Wheeler, Max W./Yates, Alan/Dols, Nicolau, *Catalan. A comprehensive grammar*, London/New York, Routledge, 1999.
- Zeugin, Senta, *La Marca Diferencial del Objeto en las variedades catalanas. Animación y posición del objeto directo como factores de influencia en la aceptabilidad de la MDO*, MA Thesis, University of Zurich, 2017.