

# Javier Caro Reina, Marco García García and Klaus von Heusinger

## Differential Object Marking in Cuban Spanish

**Abstract:** Recent research has shown that Differential Object Marking (DOM) is less frequent in some varieties of Latin American Spanish than in European Spanish. This is the case in Caribbean Spanish, which includes Cuban, Dominican, and Puerto Rican Spanish. We will investigate whether these varieties have preserved an older language stage or, rather, whether this is a more recent development resulting from DOM retraction. In this paper, we will focus on DOM in Cuban Spanish. Following on from Alfaraz (2011), who studied DOM on the basis of sociolinguistic interviews, we will examine DOM both from a diachronic and from a synchronic perspective. The diachronic approach is based on a corpus analysis encompassing the 19<sup>th</sup> and 20<sup>th</sup> centuries while the synchronic approach is based on grammaticality judgment tasks. The corpus analysis points to a slight retraction which evolved with indefinite human nouns. The results of the grammaticality judgment tasks reveal that Cuban Spanish speakers accept the absence of DOM with definite human nouns, which is unacceptable in European Spanish. They also rate the absence of DOM with indefinite human nouns as highly acceptable, as opposed to their European counterparts. We compare the findings provided from the corpus analysis and the judgment task by discussing the importance of considering both production and acceptability data. Thus, this paper makes an important empirical and theoretical contribution to the patterns of DOM in Caribbean Spanish.

---

**Acknowledgements:** This paper was presented at the Workshop “Diachrony of Differential Object Marking” (Institut National des Langues et Civilisations Orientales, Paris, November 16–17, 2017). We would like to thank the audience of the workshop for their comments. We would also like to thank Johannes Hofmann for the fieldwork in Havana, Julio Manero González for his help with the corpus analysis, and Diego Romero Heredero for his help with the design and analysis of the judgment task questionnaires. Last but not least, we would like to express our special gratitude to two anonymous reviewers for very useful comments and to the editors, Johannes Kabatek, Philipp Obrist, and Albert Wall, for their patience and great editorial work. The research for this paper has been funded by the German Research Foundation (DFG) as part of the SFB 1252 *Prominence in Language* (Project-ID 281511265) in the project B04 *Interaction of nominal and verbal features for Differential Object Marking* at the University of Cologne.

---

**Javier Caro Reina**, University of Cologne, e-mail: javier.caroreina@uni-koeln.de

**Marco García García**, University of Cologne, e-mail: marco.garcia@uni-koeln.de

**Klaus von Heusinger**, University of Cologne, e-mail: klaus.vonheusinger@uni-koeln.de

**Keywords:** animacy, Caribbean Spanish, corpus analysis, Cuban Spanish, definiteness, Differential Object Marking, European Spanish, judgment tasks, spoken data

## 1 Introduction

The term Differential Object Marking (DOM) is used to describe the phenomenon by which case marking of the direct object depends on certain semantic-pragmatic conditions such as animacy, referentiality, and topicality, as well as agentivity, affectedness, and telicity (Bossong 1985; Aissen 2003; García García/Primus/Himmelmann 2018; Witzlack-Makarevich/Seržant 2018, among others). DOM in Spanish is a very well attested and widely studied phenomenon. Traditionally, research has concentrated on European Spanish, where DOM has experienced a considerable expansion from Old to Modern European Spanish (Laca 2006; von Heusinger/Kaiser 2011; García García 2018), as will be shown in Section 2. However, there is a growing body of research that has examined the patterns of DOM in varieties of Spanish spoken in Argentina (Barrenechea/Orecchia 1977; Dumitrescu 1997; Tippets 2010; 2011; Montrul 2013; Hoff/Díaz-Campos 2015; Hoff 2018), Cuba (Alfaraz 2011), Mexico (Dulme 1986; Company Company 2002a; Lizárraga Navarro/Mora-Bustos 2010; Tippets 2010; 2011; Ordóñez/Treviño 2016), Peru (Mayer/Delicado Cantero 2015), Venezuela (Domínguez et al. 1998; Balasch 2011), Uruguay (Barrios 1981), and USA (Montrul 2014). In these studies, DOM has been approached from various perspectives, such as language attitudes (Hoff/Díaz-Campos 2015), heritage languages (Montrul 2014), language change (Company Company 2002a; 2002b), and language variation. With regard to language variation, most studies are embedded in the *Habla Culta* project (see Lope Blanch 1986). This is the case with Alfaraz (2011) for Cuban Spanish, Domínguez et al. (1998) and Balasch (2011) for Venezuelan Spanish, and Tippets (2010; 2011) for Buenos Aires, Madrid, and Mexico City Spanish.

Interestingly, the occurrence of DOM has been found to differ in varieties of European and Latin American Spanish, suggesting two opposed tendencies: DOM expansion and DOM retraction. In this respect, Caribbean Spanish has been reported to exhibit a lower frequency of DOM than other varieties of Spanish (Jiménez Sabater 1975, 169–170; Álvarez Nazario 1992, 237; López-Morales 1992, 141; Lunn 2002; Alba 2004, 140–141; Bullock/Toribio 2009; Alfaraz 2011, among others).<sup>1</sup>

---

<sup>1</sup> Additionally, similar patterns have also been observed in Bolivia (Mendoza Quiroga 1992, 459), La Palma (Régulo Pérez 1970, 82–83), and Venezuela (Domínguez et al. 1998).

An example from Dominican Spanish is given in (1), where the definite human direct objects *esa hija* ‘that daughter’ and *las personas de Francia* ‘the people from France’ are not differentially marked. Note that in European Spanish the absence of DOM would result in an ungrammatical utterance.

(1) Lack of DOM in Dominican Spanish

a. *Luba quería mucho ∅ esa hija.*

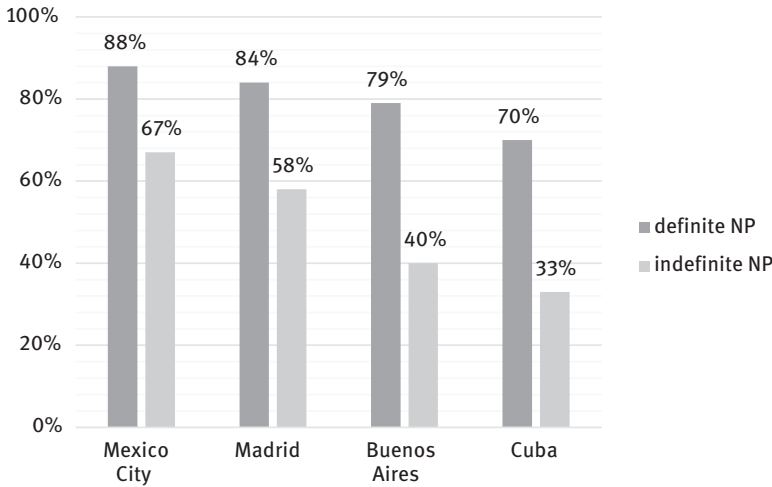
‘Luba loved that daughter very much.’

b. *Para entender ∅ las personas de Francia.*

‘In order to understand the people from France.’

(Bullock/Toribio 2009, 59)

Figure 1 illustrates the occurrence of DOM with definite and indefinite human direct objects in the varieties of Spanish spoken in Mexico City, Madrid, Buenos Aires, and Cuba.<sup>2</sup> With regard to definite human direct objects, DOM has a relative frequency of 88% (153/174) in Mexico City Spanish, 84% (87/104) in Madrid Spanish, 79% (162/205) in Buenos Aires Spanish, and 70% (168/240) in Cuban Spanish. With regard to indefinite human direct objects, DOM has a relative fre-



**Figure 1:** DOM with definite and indefinite human direct objects in selected varieties of Spanish (Tippets 2010, 134, 147, 156; Alfaraz 2011, 224).

<sup>2</sup> The data from Mexico City, Madrid, and Buenos Aires are taken from Tippets (2010), who used the corpora of the *Habla culta de la Ciudad de México* (1971), *Habla culta de Madrid* (1981), and *Habla culta de Buenos Aires* (1987), respectively. Similarly, the data from Cuba is partly based on *Habla culta de Miami* (1968–1969) (cf. Section 3 for details).

quency of 67% (32/48), 58% (21/36), 40% (21/53), and 33% (10/30), respectively. We can observe that DOM occurs more frequently in Mexican Spanish than in the other varieties of Spanish. More importantly, DOM is less frequent in Cuban Spanish than in the other varieties of Spanish.

While there have been a series of sociolinguistic studies on phonological, morphosyntactic, and lexical variation in Cuban Spanish (cf. Cuza 2017 for a comprehensive overview), DOM has not received much attention, with the exception of the work of Alfaraz (2011), which will be presented in Section 3. The aim of the present study is to give a detailed account of DOM in Cuban Spanish drawing on both diachronic and synchronic data. We will address the question of whether the lower incidence of DOM in Cuban Spanish constitutes a remnant from older stages of Spanish or, rather, whether it constitutes a recent development pointing to a retraction of DOM. This is followed by a methodological discussion of the adequate means for investigating variation in DOM, including sociolinguistic interviews, corpus analyses, and grammaticality judgments.

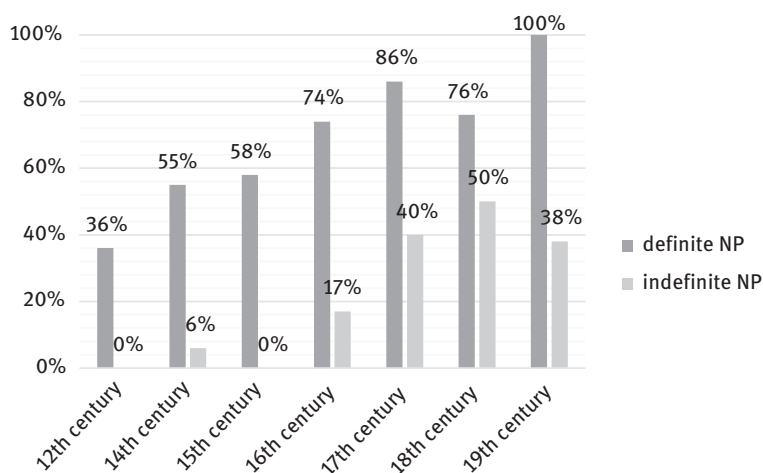
The paper is structured as follows: Section 2 deals with the diachronic development of DOM in European Spanish. Section 3 presents the study conducted by Alfaraz (2011) on the basis of spontaneous speech. Section 4 contains the diachronic corpus analysis conducted for 19<sup>th</sup>- and 20<sup>th</sup>-century Cuban Spanish. Section 5 is dedicated to the grammaticality judgment tasks for Cuban and European Spanish. Section 6 summarizes the main findings of our study and concludes by discussing DOM expansion and retraction.

## 2 Diachronic development of DOM in European Spanish

The diachronic development of DOM has been examined in a number of empirical studies which have assessed the relevance of animacy and definiteness (Company Company 2002a; Laca 2006), topicality (Melis 1995; Pensado 1995), affectedness (von Heusinger 2008; von Heusinger/Kaiser 2011), and telicity (Romero Heredero, this volume). More recently, the development of DOM has been considered with regard to both monotransitive and ditransitive constructions (Ortiz Ciscomani 2005; 2011; von Heusinger 2018). While most studies have concentrated on human and animate direct objects, some have been devoted to inanimate direct objects (Company Company 2002b; Barraza Carbajal 2008; García García 2014; 2018).

Laca (2006) provides the most fine-grained analysis. For this reason, we will refer to her findings in the ensuing sections. Before summarizing her results, we will address two critical empirical issues. First, Laca's (2006) corpus covers a large period of time reaching from the 12<sup>th</sup> to the 19<sup>th</sup> century. However, her data base is rather small, being generally confined to just one or two text samples per century. This is all the more problematic when we try to account for differences between European and Latin American Spanish since the data for each linguistic area is only based on one text per century (cf. Kabatek 2016, 216, 230 for further critical aspects). Second, Laca's (2006) findings cannot be directly compared to other diachronic studies such as Company Company (2002a) and von Heusinger/Kaiser (2011). The comparison with Company Company (2002a) is difficult because the author does not distinguish between definite and indefinite human direct objects. Similarly, the diachronic study of von Heusinger/Kaiser (2011) is less fine-grained since it only contains data from the 15<sup>th</sup>, 17<sup>th</sup> and 19<sup>th</sup> centuries.

In the absence of a more extensive and comparable corpus study, we take Laca's (2006) findings as a point of departure for our diachronic investigation (cf. Section 4). Where possible, her data will be complemented and compared with the findings from other empirical studies. On the basis of Laca's (2006) corpus study, Figure 2 shows the diachronic development of DOM with human direct objects according to definiteness and century.



**Figure 2:** DOM with human direct objects according to definiteness and century in European Spanish (adapted from Laca 2006, 442–443).

Note that in contrast to the more fine-grained distinctions put forward by Laca (2006), Figure 2 only includes the results regarding full lexical NPs (e.g. *una/la mujer* ‘a/the woman’). That is, it excludes pronouns, proper names, NPs without lexical heads (e.g. *los más conocidos* ‘the best known’), definite-like NPs with universal quantifiers (e.g. *cada persona* ‘each person’), indefinite-like NPs with existential (or weak) quantifiers (e.g. *algunas personas* ‘some people’), and bare nouns (e.g. *personas* ‘people’). In addition, it only contains the occurrences of DOM in European Spanish. Laca’s (2006) findings can be summarized as follows: First, there is a clear rise of DOM with both definite and indefinite NPs, whereby the percentages of DOM are clearly and constantly higher with definite NPs than with indefinite NPs. Second, with regard to definite NPs, DOM increases greatly and ends up becoming categorical. More specifically, we observe 36% (13/36) of DOM in the 12<sup>th</sup> century, 55% (36/66) in the 14<sup>th</sup> century, 58% (38/65) in the 15<sup>th</sup> century, 74% (26/35) in the 16<sup>th</sup> century, 86% (117/136) in the 17<sup>th</sup> century, 76% (22/29) in the 18<sup>th</sup> century, and 100% (28/28) in the 19<sup>th</sup> century. Third, with regard to indefinite NPs, DOM also increases considerably, though it never becomes categorical. The development seems to begin between the 15<sup>th</sup> and 16<sup>th</sup> centuries (cf. also von Heusinger/Kaiser 2011, 611). We observe a sharp rise of DOM, reaching 17% (1/6) in the 16<sup>th</sup> century, 40% (21/53) in the 17<sup>th</sup> century, and 50% (8/16) in the 18<sup>th</sup> century. Later, there is a slight decrease to 38% (3/8) in the 19<sup>th</sup> century.<sup>3</sup>

In summary, Laca’s (2006) results show that in Old Spanish, DOM is optional with definite human direct objects, but absent from indefinite human direct objects. In Modern European Spanish, by contrast, DOM is obligatory with definite human direct objects, but optional with indefinite human direct objects. As we will see in Section 3, spoken Cuban Spanish resembles 16<sup>th</sup>-century European Spanish. This raises the question of whether Modern Cuban Spanish has retained this prior language stage. This issue will be discussed in more detail in Section 4.

### 3 Alfaraz’s (2011) spoken data

In this section, we will report on the study carried out by Alfaraz (2011), who examined DOM in Cuban Spanish on the basis of recordings made in Miami in

---

<sup>3</sup> In this respect, Laca (2006, 460) argues that the relatively high percentage of DOM in the 18<sup>th</sup> century is due to the disproportionately high number of causative constructions compared to previous centuries. These constructions have been shown to have a positive influence on the frequency of DOM (cf. García García 2018, 235–336).

the 1960s and 1990s. These two sets of data allow for a real time study. The first corpus consists of a subset of the recording collected in 1968–1969 for the *Habla Culta* project. It is based on 24 speakers aged between 30 and 50. The second corpus was collected by Alfara (2011) in 1996–98. It comprises 26 speakers that were classified into two age groups (30–43 and 62–77), which allowed for an apparent time study. Importantly, both corpora are comprised of (semi-)directed interviews that were conducted with monolingual speakers of Spanish upon their arrival in the USA. Therefore, we can exclude a contact-induced change resulting from contact with English (cf. Carter/Lynch 2015 for Miami Cuban Spanish).

Alfara (2011) found in her two data sets a total of 502 human direct objects, of which 368 (73%) contained DOM. She further analyzed the instances of DOM according to linguistic and social factors. The linguistic factors include referentiality (pronoun, proper name, definite NP, indefinite specific NP, and non-specific NP) and word order (postverbal, preverbal). The social factors include time period (1968–1969, 1996–1998) and age group (30–43 and 62–77 years for the 1990s period).

With regard to referentiality, pronouns and proper names are always differentially marked, both of which have a relative frequency of 98% (94/96 and 88/90, respectively). In contrast, DOM gradually decreases with definite NPs, indefinite specific NPs, and non-specific NPs, which have a relative frequency of 70% (168/240), 33% (10/30), and 17% (8/46), respectively. As for indefinite NPs, the author distinguishes between specific and non-specific NPs. The latter category is not homogeneous since it is comprised of non-specific indefinite NPs and bare nouns. For this reason, we will exclusively refer to specific indefinite NPs when talking about indefinite NPs. In summary, DOM was found to occur more frequently with definite rather than indefinite NPs (70% vs. 33%), as shown in Figure 3.

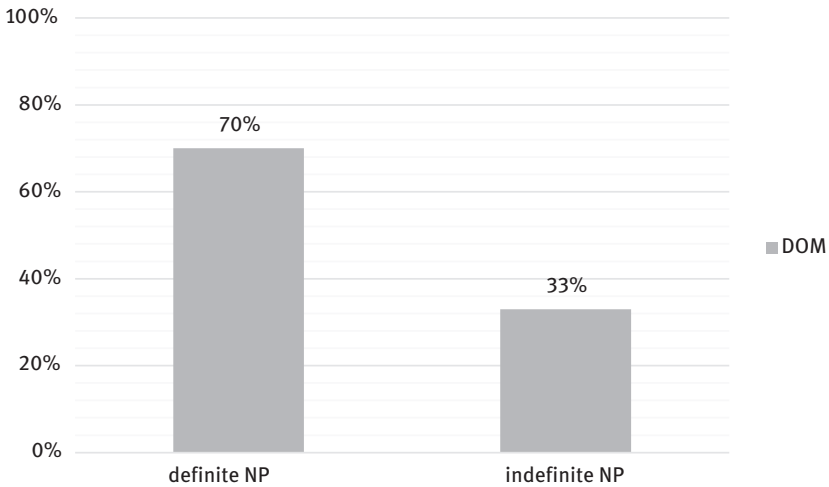
Examples from the sample are shown in (2), where the definite direct objects *la vieja aquella* ‘that old woman’ and *la abuela de Tetico* ‘Tetico’s grandmother’ are not differentially marked. The same applies even for the left-dislocated direct object *esa gente* ‘those people’. In this respect, Cuban Spanish differs from European Spanish, where *a*-marking is required in all of these cases.

(2) Lack of DOM with definite human direct objects in Cuban Spanish

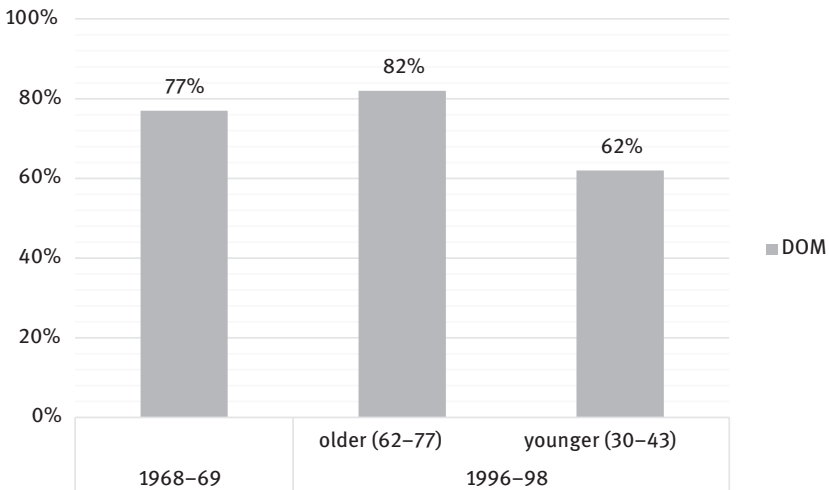
- a. *Tú no viste Ø la vieja aquella fajándose con el viejo aquel.*  
‘You didn’t see that old woman fighting with that old man.’
- b. *Ella no conoció Ø la abuela de Tetico.*  
‘She didn’t meet Tetico’s grandmother.’
- c. *Ø Esa gente tú la manipulas.*  
‘You’re manipulating those people.’

(Alfara 2011, 228–229)

Interestingly, the occurrence of DOM varies according to time period, as depicted in Figure 4. In the first period (1968–1969), DOM was found to occur in 77% of cases involving definite and indefinite NPs. In the second period (1996–1998), however, the author observed a decrease in the occurrence of DOM. More specifically, the



**Figure 3:** DOM with human direct objects according to definiteness in Cuban Spanish (adapted from Alfaraz 2011, 224).



**Figure 4:** DOM in Cuban Spanish with all human direct objects according to age group (adapted from Alfaraz 2011, 224).



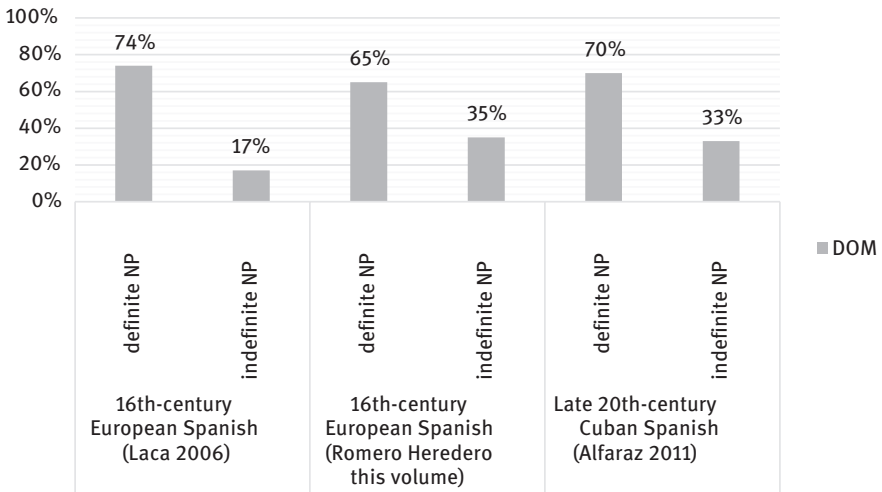
younger generation employed DOM less frequently than the older one (62% vs. 82%). In other words, the results gleaned from the real time and apparent time studies show DOM retraction. Note that the results provided by Alfaraz (2011, 224) do not allow for the combination of age group with definiteness, which would have resulted in a more fine-grained picture of the development of DOM across generations according to definite and indefinite NPs.

## 4 Diachronic corpus analysis

In order to study the patterns of DOM in Cuban Spanish in the 19<sup>th</sup> and 20<sup>th</sup> centuries, we conducted a corpus analysis based on the *Corpus diacrónico del español* (CORDE). Section 4.1 presents the hypotheses according to the patterns of DOM described in Section 2 and Section 3. Section 4.2 is dedicated to the study design. Section 4.3 contains the results according to definiteness and animacy in SVO sentences. Section 4.4 discusses the results.

### 4.1 Hypotheses

The patterns of DOM laid out in Section 2 and Section 3 enable us to detect similarities between 16<sup>th</sup>-century European Spanish and Modern Cuban Spanish. More specifically, Modern Cuban Spanish resembles 16<sup>th</sup>-century European Spanish with respect to definiteness, as illustrated in Figure 5. The values for 16<sup>th</sup>-century European Spanish are based on Laca (2006, 442) and Romero Heredero (this volume) (cf. also Keniston 1937, 10–11, 14 for DOM with definite and indefinite human direct objects in 16<sup>th</sup>-century Castilian prose). With regard to definite human direct objects, DOM has a relative frequency of 74% (26/35) and 65% (468/720) in 16<sup>th</sup>-century European Spanish and 70% (168/240) in Modern Cuban Spanish. With regard to indefinite human direct objects, DOM has a relative frequency of 17% (1/6) and 35% (85/240) in 16<sup>th</sup>-century European Spanish and 33% (10/30) in Modern Cuban Spanish. The different percentages found with indefinite human direct objects in 16<sup>th</sup>-century European Spanish (17% vs. 35%) result from the number of tokens examined by Laca and Romero Heredero (1/6 vs. 85/240) (cf. Section 2 for a critical discussion).



**Figure 5:** DOM in 16<sup>th</sup>-century European Spanish and 20<sup>th</sup>-century Cuban Spanish with definite and indefinite human direct objects (Laca 2006, 442; Romero Heredero, this volume; Alfaraz 2011, 224).

The similarities between 16<sup>th</sup>-century European Spanish and Modern Cuban Spanish (especially with definite human direct objects) raises the question of whether Cuban Spanish is undergoing a process of retraction or whether it has just retained a prior language stage. In order to answer this question, we carried out a diachronic corpus-based study that involves an analysis of DOM in the 19<sup>th</sup> and 20<sup>th</sup> centuries. Our hypotheses are summarized in (3). Note that scholars such as Pérez Guerra (1992, 489) have explained the lower occurrence of DOM in Dominican Spanish in terms of a remnant feature. However, this assumption has not been empirically tested yet.

(3) Hypotheses for 19<sup>th</sup>- and 20<sup>th</sup>-century Cuban Spanish

H1: Cuban Spanish is undergoing DOM retraction.

H2: Cuban Spanish has retained a prior language stage of DOM.

## 4.2 Study design

The analysis is based on the *Corpus diacrónico del español* (CORDE). The corpus allows searching for single words (or combination of words) according to author, work, time span, text type (book, journal, etc.), country, and genre (cf. Octavio de

Toledo y Huerta 2006 for a critical discussion). In this respect, the corpus differs from others such as the *Corpus del español*, which does not allow for a diachronic search according to country. The sources for Cuban Spanish are well suited for a diachronic study since they contain a total of 883,618 words for the 19<sup>th</sup> century and 1,499,345 words for the 20<sup>th</sup> century.<sup>4</sup> Unfortunately, the CORDE is not annotated. Therefore, it is not possible to search for specific syntactic patterns such as DOM. Since verbal factors such as affectedness have proved to have an impact on the occurrence of DOM (cf. von Heusinger/Kaiser 2011; Romero Heredero in press), we manually searched for transitive verbs with high and low affectedness, i.e. predicates with a higher and a lower preference for DOM, in order to obtain a balanced data set with respect to this verbal factor. Note, however, that affectedness will not be treated in this study.

The verbs with high affectedness are the following (the number of tokens are given in brackets): *cuidar* ‘to take care of’ (6), *ejecutar* ‘to execute’ (2), *golpear* ‘to hit’ (7), *herir* ‘to hurt’ (3), *humillar* ‘to humiliate’ (7), *matar* ‘to kill’ (38), *violar* ‘to rape’ (1), and *violentar* ‘to force’ (1). The verbs with low (or no) affectedness are *buscar* ‘to look for’ (38 tokens), *conocer* ‘to know’ (10), *contemplar* ‘to contemplate’ (14), *entender* ‘to understand’ (4), *escuchar* ‘to listen to’ (7), *esperar* ‘to wait for’ (10), *mirar* ‘to look at’ (87), *observar* ‘to observe’ (12), *oír* ‘to hear’ (14), and *ver* ‘to see’ (35).<sup>5</sup> The number of tokens involving verbs with high and low affectedness amounts to 65 and 231, respectively. The search was conducted by means of regular expressions.<sup>6</sup>

Altogether, we found 296 instances of full definite and indefinite human direct object NPs. For comparison, Alfaraz (2011) has 270 tokens (excluding non-specific NPs) while von Heusinger/Kaiser (2011) have 423 tokens. By contrast, Laca (2006)

---

4 In contrast to other Caribbean Spanish varieties, Cuban Spanish is well represented in the corpus. For comparison, the corpus contains a total of 393,119 and 134,287 words for Puerto Rico and the Dominican Republic respectively for the 19<sup>th</sup> and 20<sup>th</sup> centuries.

5 With regard to the perception verbs *escuchar* ‘to listen’, *mirar* ‘to look at’, *oír* ‘to hear’, and *ver* ‘to see’, we also looked at ACI constructions, which had a frequency of 1, 3, 5, and 6 tokens respectively. These cases were always differentially marked (cf. García García 2018, 235–237 for DOM with ACI constructions).

6 Regular expressions allow for the substitution of one or more characters. For example, the string *mat\** results in 1,483 tokens from 72 different documents in the time span between 1800 and 1975. The tokens include inflected forms such *mataba* ‘s/he was killing’, *mató* ‘s/he killed’, *matando* ‘killing’, etc. However, they also contain other forms such as *matanza* ‘carnage’, *materia* ‘matter’, *matrimonio* ‘marriage’, etc. which had to be excluded manually. Of the 1,483 tokens, only 38 involved instances of *matar* ‘to kill’ with definite and indefinite human NPs as their direct objects.

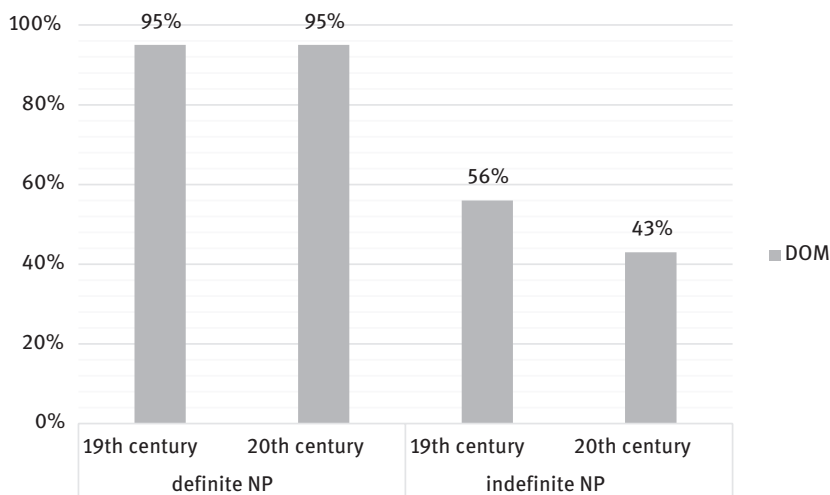
has a total of 775 tokens, which are distributed along the centuries as follows: 42 (12<sup>th</sup> c.), 97 (14<sup>th</sup> c.), 76 (15<sup>th</sup> c.), 181 (16<sup>th</sup> c.), 189 (17<sup>th</sup> c.), 85 (18<sup>th</sup> c.), and 105 (19<sup>th</sup> c.). The instances of human direct objects found in the corpus were subsequently classified according to verb, affectedness (high vs. low), century (19<sup>th</sup> vs. 20<sup>th</sup> century), year, DOM (presence vs. absence), definiteness (definite vs. indefinite NP), author, and work. Appendix 1 gives an overview of the sources that contained instances of DOM with human direct objects in combination with the verbs selected. The table in Appendix 1 is arranged according to century, author, and record. The examples taken from the corpus are cited according to year, author, and a shortened name of the record (e.g. 1966/Lezama/Paradiso). We did not distinguish specific from non-specific indefinite direct objects. Our search was restricted to direct object NPs with human referents in SVO sentences. That is, we excluded collective nouns (*gente* ‘people’, *multitud* ‘crowd’, etc.),<sup>7</sup> animate non-human NPs,<sup>8</sup> inanimate NPs, proper names, bare nouns, left dislocations, and impersonal constructions.

### 4.3 Results

The results of the corpus analysis of Cuban Spanish are arranged in Figure 6 according to definiteness (definite vs. indefinite NP) and century (19<sup>th</sup> vs. 20<sup>th</sup> century). With regard to definite human direct objects, DOM has a relative frequency of 95%, both in the 19<sup>th</sup> (120/126) and 20<sup>th</sup> (118/124) centuries. With regard to indefinite human direct objects, DOM has a relative frequency of 56% (9/16) and 43% (13/30) in the 19<sup>th</sup> and 20<sup>th</sup> centuries, respectively. That is, the occurrence of DOM has remained stable with definite human direct objects (95%) while it has experienced a slight decrease with indefinite human direct objects (56% > 43%). Examples of lack of DOM with definite human direct objects are given in (4), where the NPs *los doce prisioneros* ‘the twelve prisoners’ and *el padrino de la boda* ‘the best man’ are not differentially marked.

<sup>7</sup> Compare *aguantar a la gente* ‘to put up with people’ (1938/Serpa/Contrabando) to *esperar la gente de mister Bourton* ‘to wait for Mr Bourton’s people’ (1938/Serpa/Contrabando).

<sup>8</sup> For example, *matar un gallo blanco* ‘to kill a white cock’ (1906/Ortiz/Brujos) vs. *matar a un cerdo* ‘to kill a pig’ (1906/Ortiz/Brujos).



**Figure 6:** DOM in Cuban Spanish according to definiteness and century (CORDE).

(4) Lack of DOM with definite human direct objects

- a. *Aquiles mató con su mano Ø los doce prisioneros* (1889/Martí/Edad).  
'Achilles killed the twelve prisoners with his own hands.'
- b. *Sonriéndose marchó hacia la sala para buscar Ø el padrino de la boda* (1966/Lezema/Paradiso).  
'Smiling he went to the hall to look for the best man.'

Examples of lack of DOM with indefinite human direct objects are shown in (5), where the NPs *otra madre* 'another mother' and *un hombre* 'a man' are not differentially marked.

(5) Lack of DOM with indefinite human direct objects

- a. *Conque ya sabes... a buscar Ø otra madre* (1884/Ortega/Cleopatra).  
'You already know... Go and look for another mother.'
- b. *Yo maté Ø un hombre* (1938/Serpa/Contrabando).  
'I killed a man.'

In order to examine inter-speaker variation, we further looked at the instances of DOM with human direct objects according to author (cf. Table 1). For example, Gómez de Avellaneda consistently uses DOM with both definite and indefinite human direct objects. However, in Insúa, Lezama Lima, and Serpa, the presence of DOM is more frequent with definite human direct objects while the absence of DOM

is more frequent with indefinite human direct objects. Notably, none of the authors generally avoided DOM with both definite and indefinite human direct objects.

**Table 1:** DOM with human direct objects according to author and definiteness.

	Definite NP		Indefinite NP	
	DOM	no DOM	DOM	no DOM
Anónimo	5	0	0	0
Auber Noya	15	1	0	1
Augier	4	0	0	0
Bobadilla	9	0	0	0
Cabrera Infante	7	2	1	2
Carlos Felipe	14	0	2	1
Carpentier	9	0	2	1
El Cucalambé	6	0	0	1
Fernández Retamar	0	0	1	0
Gómez de Avellaneda	56	0	3	0
Guillén	1	0	0	0
Heredia	5	1	0	1
Hostos	6	2	2	1
Insúa	11	0	1	3
Lachatañaré	4	1	1	1
Lezama Lima	39	1	2	6
Loynaz	3	0	1	0
Martí	22	1	3	1
Ortega Munilla	13	1	1	2
Ortiz	3	2	1	0
Serpa	6	1	1	3

The results gained from Table 1 point to the existence of variation within some authors. Examples of variation involving definite and indefinite human direct objects are given in (6) and (7), respectively.

- (6) Inter-speaker variation with definite human direct objects (taken from 1938/Serpa/Contrabando)
- a. *na más que mató a su mujer.*  
'Just for killing his wife.'

- b. *¿busca Ø los guardacostas?*  
‘Is he looking for the coastguards?’
- (7) Inter-speaker variation with indefinite human direct objects (taken from 1966/Lezama Lima/Paradiso)
- a. *se sentó en un café para esperar a un amigo, que le soportaba sus crisis*  
‘He sat down in a café to wait for a friend, who tolerated his crisis.’
- b. *conozco Ø un profesor de estética que nos visitó hace pocos meses.*  
‘I know an aesthetics teacher who visited us a couple of months ago.’

Let us turn to the hypotheses postulated in the previous section. The corpus analysis provides evidence that in 19<sup>th</sup>-century Cuban Spanish, DOM is much more frequent than in 16<sup>th</sup>-century European Spanish. Recall from Section 2 that in 16<sup>th</sup>-century European Spanish, DOM occurs with definite and indefinite NPs with a relative frequency of 74%/65% and 17%/35%, respectively (cf. Figure 5). By contrast, in 19<sup>th</sup>-century Cuban Spanish, DOM occurs with definite and indefinite NPs with a relative frequency of 95% and 56%, respectively. Thus, H1 is not borne out for 19<sup>th</sup>-century Cuban Spanish. In other words, 19<sup>th</sup>-century Cuban Spanish has not retained a prior language stage. This result therefore challenges the assumption held by scholars such as Pérez Guerra (1992, 489) that Caribbean Spanish preserves the patterns of DOM as found in prior stages of European Spanish. This issue will be discussed in more detail in the ensuing section.

#### 4.4 Discussion

In this section, we will first address DOM retraction and then critically discuss the implications derived from the type of language data (written vs. spoken). The corpus search in texts written by Cuban authors in the 19<sup>th</sup> and 20<sup>th</sup> centuries has revealed a slight decrease of DOM with indefinite human direct objects. Table 2 summarizes the occurrence of DOM with definite and indefinite human direct objects in 16<sup>th</sup>-century European Spanish (Laca 2006; Romero Heredero, this volume), 19<sup>th</sup>- and 20<sup>th</sup>-century written Cuban Spanish (CORDE), and spoken Modern Cuban Spanish (Alfaraz 2011). Assuming that the patterns of DOM in 16<sup>th</sup>-century European Spanish also applied for 16<sup>th</sup>-century Cuban Spanish, it follows that there is a rise of DOM between the 16<sup>th</sup> and 20<sup>th</sup> centuries (e.g. 74%/65% > 95% in the case of definite NPs). In this respect, Cuban Spanish resembles European Spanish, which has also experienced DOM expansion (cf. Figure 2). However, we could detect a slight decrease of DOM with indefinite human direct objects from the 19<sup>th</sup> to the 20<sup>th</sup> century (56% > 43%). Considering Modern

Cuban Spanish on the basis of Alfaraz's spoken data, this tendency continues to develop with indefinite human direct objects (43% > 33%) while DOM also begins to decrease with definite human direct objects (95% > 70%), which suggests an ongoing change involving DOM retraction. In this respect, Cuban Spanish differs from Modern European Spanish. The diachronic picture that emerges from Table 2 further suggests that in Cuban Spanish, DOM retraction began with indefinite human direct objects and subsequently expanded to definite human direct objects. These findings support evidence for the patterns of DOM retraction, which affects less prominent categories (indefinite NPs) prior to more prominent ones (definite NPs). In this sense, Cuban Spanish seems to be another instance of DOM retraction within the Romance language group, as has been reported for Portuguese (Delille 1970) and Catalan (Dalrymple/Nikolaeva 2011, 212).

**Table 2:** Diachronic overview of DOM with definite and indefinite human direct objects.

	Definite NP	Indefinite NP
16 <sup>th</sup> -century European Spanish (Laca 2006)	74% (26/35)	17% (1/6)
16 <sup>th</sup> -century European Spanish (Romero Heredero, this volume)	65% (468/720)	35% (85/240)
19 <sup>th</sup> -century Cuban Spanish (CORDE)	95% (120/126)	56% (9/16)
20 <sup>th</sup> -century Cuban Spanish (CORDE)	95% (118/124)	43% (13/30)
Modern Cuban Spanish (Alfaraz 2011)	70% (168/240)	33% (10/30)

A word of caution, however, should be that we are comparing two different types of sources. On the one hand, we have written corpora, which are associated with a formal (or standard) style. On the other, we have spontaneous speech, which is associated with a casual (or informal) style. As a consequence, written language can be assumed to be more averse to language innovations such as DOM retraction whereas spoken language is probably more progressive in this respect. This might be the reason for the higher frequency of DOM found in our corpus search based on written language (CORDE) compared to the lower frequency of DOM attested in the spoken data used by Alfaraz (2011).

In order to deepen our understanding of this kind of variation, we complemented the production studies (corpus, interviews) with a grammaticality judgment task. While production data only provides information about the more acceptable form in a given context, acceptability judgments also offer insights into less commonly used forms in that context. We think that such information will help us to observe the diachronic development of this variation in much clearer light.



## 5 Judgment tasks

We designed a questionnaire in order to assess the acceptability grade regarding the presence and absence of DOM with human definite and indefinite direct objects both in Cuban and European Spanish. Section 5.1 formulates the hypotheses according to the patterns of DOM in Modern Cuban and Modern European Spanish previously described in Section 2 and Section 3, respectively. Section 5.2 gives a detailed account of the study design. Section 5.3 presents the results of the acceptability judgments of DOM and lack of DOM according to animacy and definiteness. Section 5.4 discusses the results.

### 5.1 Hypotheses

In Section 4, we studied the use of DOM in Spanish records written by Cuban authors in the 19<sup>th</sup> and 20<sup>th</sup> centuries. We then compared the results of our corpus analysis to the findings of Alfaraz (2011), Laca (2006), and Romero Heredero (this volume). We found that the Cuban corpus of written language behaved very much like the European corpus of written language, but quite differently from the Cuban corpus of spoken language of Alfaraz (2011). This provided, on the one hand, evidence for H1 that in Cuban Spanish, DOM underwent retraction and, on the other hand, evidence against H2 that Cuban Spanish has retained a prior language stage. In this section, we provide additional empirical evidence for H1, namely data from a grammaticality judgment task. In this task, participants had to decide how acceptable they found a human direct object with or without DOM. Grammaticality judgment tests offer a different empirical perspective than corpus analyses. As for specific grammatical constructions such as DOM, in a corpus we usually find only one form in a given sentence, let's say the more acceptable form, which in the case of constructions with human definite direct objects will be those with DOM. By contrast, in grammatical judgment tasks we obtain a graded evaluation between competing forms, i.e. between standard and less standard forms, which in the case of sentences with definite human direct objects correspond to those with and without DOM, respectively.

According to our hypothesis H1, we predict the following results from the grammaticality judgments. If there is DOM retraction, we expect Cuban Spanish speakers to rate the absence of DOM (noDOM) with human definite and indefinite direct objects better, i.e. with higher acceptability values, than their European counterparts (P1a). In addition, we also predict that Cuban Spanish speakers will rate the presence of DOM worse, i.e. with lower acceptability values, than their European counterparts (P1b). Since we quantify over judgments, we thus expect

a higher average acceptability value for cases of noDOM and a lower average acceptability value for DOM instances from Cuban Spanish speakers in comparison with their European Spanish counterparts.

(8) Predictions for the grammaticality judgments tasks based on H1

H1: In Cuban Spanish, DOM underwent retraction.

P1a: Speakers of Cuban Spanish will show higher acceptability values for noDOM than speakers of European Spanish.

P1b: Speakers of Cuban Spanish will show lower acceptability values for DOM than speakers of European Spanish.

## 5.2 Study design

The questionnaire employed for Cuban and European Spanish was comprised of general information on the sociolinguistic background of the participants (age, gender, education level, first and second language), instructions with four examples illustrating how to fulfil the judgment task, the judgment task itself, and final comments. The judgment tasks had an approximate duration of 15 minutes. The questionnaires for European and Cuban Spanish differed slightly from each other since the vocabulary had to be adapted to lexical variation (e.g. *celular* and *móvil* for ‘mobile phone’ in Cuban and European Spanish, respectively).

The grammaticality judgment task consisted of a Likert scale ranging from 1 (unacceptable) to 6 (totally acceptable). The questionnaire included 16 test items, 8 with definite and 8 with indefinite human direct objects in SVO sentences.<sup>9</sup> In addition, we provided 16 fillers which served as control items. The test items displayed a direct object, which was employed once with and once without DOM. The two different versions appeared in different item lists (questionnaire A and questionnaire B) such that the participants could only see a single version of the same direct object. The two experimental item lists were pseudo-randomized in different orders for questionnaires A and B before being distributed to the participants. Examples of items with and without DOM both with definite and indefinite human direct objects are shown in (9) and (10), respectively (cf. Appendix 2 for the complete list of test items employed for definite and indefinite human direct objects). The verbs were selected according to affectedness. More specifically, the verbs with high affectedness are *cuidar* ‘to take care of’, *golpear* ‘to hit’,

---

<sup>9</sup> This questionnaire further contained 32 items which tested inanimate direct objects in SVO sentences as well as animate and inanimate direct objects in clitic-doubling constructions. For the present study, these conditions have not been considered.

*matar* ‘to kill’, and *lesionar* ‘to injure’. The verbs with low affectedness are *acusar* ‘to accuse’, *denunciar* ‘to report’, *oír* ‘to hear’, and *ver* ‘to see’. Finally, the filler sentences were comprised of 8 grammatical (e.g. *Francisco renunció al puesto de trabajo* ‘Francisco refused the job’) and 8 ungrammatical control sentences (e.g. \**José le llevó Juan al libro* ‘José brought Juan to the book’).

- (9) DOM condition with definite and indefinite human direct objects
- a. *Patricio lesionó al portero en la discoteca.*  
‘Patricio injured the doorman at the nightclub.’
  - b. *Alberto lesionó a un defensa en el partido de la semana pasada.*  
‘Alberto injured a defender during last week’s match.’
- (10) noDOM condition with definite and indefinite human direct objects
- a. *Patricio lesionó Ø el portero en la discoteca.*  
‘Patricio injured the doorman at the nightclub.’
  - b. *Alberto lesionó Ø un defensa en el partido de la semana pasada.*  
‘Alberto injured a defender during last week’s match.’

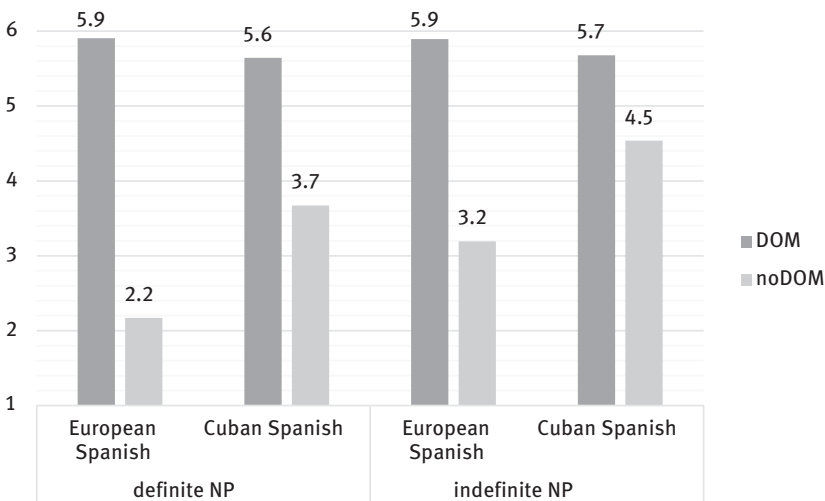
With regard to the distribution of the questionnaire, we employed two different methods depending on the country in question. For Cuban Spanish, the questionnaires were handed out to the participants by a student from the University of Cologne in a university classroom of Havana. For European Spanish, the questionnaires were distributed electronically by means of the platform Google Forms. The access link was made available on the websites of universities and social networks. We obtained a total of 214 filled-out questionnaires, of which 75 were from Cuba (38 for questionnaire A and 37 for questionnaire B) and 139 from Spain (82 for questionnaire A and 57 for questionnaire B). After revising the filled-out questionnaires, we had to remove 16 participants since the answers to the control fillers (both grammatical and ungrammatical) deviated considerably from the expected values in more than 20% of the answers. Thus, the number of valid questionnaires amounted to 62 for Cuba (33 for questionnaire A and 29 for questionnaire B) and 136 for Spain (79 for questionnaire A and 57 for questionnaire B).

### 5.3 Results

Figure 7 summarizes the results of the questionnaire for European and Cuban Spanish. The DOM condition with definite human direct objects (cf. ex. 9a above) is highly acceptable in both varieties (European Spanish: 5.9 vs. Cuban Spanish 5.6). Interestingly, the DOM condition for indefinite human direct objects (cf. ex.

9b) is also very acceptable in both varieties (5.9 vs. 5.7). Thus, prediction P1b, according to which speakers of Cuban Spanish should show lower acceptability values for DOM than speakers of European Spanish, is not supported by the questionnaire study.

As for the noDOM condition, we observe more variation. Definite human direct objects without DOM (cf. ex. 10a) are not acceptable in European Spanish (2.2) while they are much more acceptable in Cuban Spanish (3.7). We observe a very similar pattern for the lack of DOM with human indefinite direct objects (cf. ex. 10b): In European Spanish it is not as acceptable as in Cuban Spanish (3.2 vs. 4.5). Hence, prediction P1a, according to which speakers of Cuban Spanish should exhibit higher acceptability values for noDOM than speakers of European Spanish, is clearly confirmed by our questionnaire study.



**Figure 7:** Acceptability values for DOM and noDOM with definite and indefinite human direct objects in European and Cuban Spanish (1 = unacceptable, 6 = totally acceptable).

## 5.4 Discussion

As for the DOM condition, the results of the judgment task experiment confirm the assumptions in the literature and the observations from the corpus studies in Section 4. Definite human direct objects with DOM are always rated as perfect forms, both in European and Cuban Spanish. This supports the assumption that DOM is obligatory with definite human direct objects. Contrary to pre-

diction P1b, however, we found no difference between European and Cuban Spanish, even though Alfaraz (2011) mentioned some examples with definite direct objects without DOM (cf. ex. 2 above). The very high acceptability of DOM with definite direct objects in Cuban Spanish suggests that if there is retraction, it is optional since the forms with DOM are fully acceptable. We observe a very similar pattern for DOM with indefinite direct objects: They are rated as totally acceptable both in European and in Cuban Spanish. At first glance, this is surprising since corpus studies point to a clear difference between the distribution of DOM with definite and indefinite direct objects, the latter showing a much lower frequency of DOM (cf. Figures 1 and 2 as well as Table 2 in Sections 1, 2 and 4.4, respectively). This contrast is not reflected in the acceptability study.

We can account for these different results by assuming that in corpora we usually find the more acceptable form for a certain context whereas the data from the questionnaire study shows whether or not a form is acceptable. Since DOM is optional with indefinites, or more specifically, obligatory with human specific indefinites, but optional with human non-specific indefinites, participants always rated DOM with human indefinite direct objects with very high acceptability values. Thus, the high acceptability of DOM with human indefinite direct objects does not contradict the results from the corpus analysis. Note that, as for definite direct objects, we do not find a difference between European and Cuban Spanish with respect to the acceptability of DOM with indefinite direct objects. This suggests that, if there is retraction in Cuban Spanish, it is also optional for indefinite direct objects.

While the acceptability of DOM is always very high in both varieties, the acceptability of noDOM cases differs strongly and therefore allows for interesting observations. The lack of DOM (noDOM condition) with definite human direct objects is ungrammatical in European Spanish and should therefore be rated very low, which is actually the case (2.2). Speakers of Cuban Spanish, however, rate this construction much higher (3.7). This is consistent with the observation in Alfaraz (2011) that noDOM is found much more often in Cuban Spanish than in European Spanish (cf. Table 2). The relatively high acceptability of noDOM in Cuban Spanish can be viewed as indirect evidence of DOM retraction in this variety.

The same differences between Cuban and European Spanish concerning the acceptability of noDOM with definite direct objects are also attested for noDOM with indefinite direct objects. In European Spanish, indefinite direct objects without DOM receive medium-range grammaticality scores (3.2) while in Cuban Spanish they are rated as quite acceptable (4.5). It is surprising that European speakers of Spanish rate this construction as only halfway acceptable. If DOM

is optional with indefinite direct objects, we would expect the lack of DOM to be much more acceptable, as is the case with speakers of Cuban Spanish. We think, however, that our examples of direct objects in simple transparent sentences clearly provide instances of specific indefinites. We further assume with the literature that direct objects without DOM cannot receive a specific interpretation (Leonetti 2004, 98–99).<sup>10</sup> Under this assumption, the rather low rating seems to reflect the mismatch between a specific interpretation of the direct object and its realization with a form that is restricted to non-specific meanings, at least in European Spanish. For Cuban Spanish, the quite acceptable ratings for noDOM with indefinites suggest that the mentioned requirement of DOM with specific indefinite direct objects does not hold. Be this as it may, we see a clear difference between European and Cuban Spanish. As has been shown, the absence of DOM is much more acceptable in Cuban than in European Spanish. Again, this points to a higher flexibility of DOM in Cuban Spanish and to a first step towards retraction.

In summary, the acceptability study confirms prediction P1a that Cuban Spanish speakers show higher acceptability values for noDOM cases than European Spanish speakers, but not prediction P1b that Cuban Spanish speakers exhibit lower acceptability values for DOM cases than European speakers. We still take this as support for our hypothesis H1 that, in Cuban Spanish, DOM underwent retraction. We would also like to assert that the different empirical methods complement each other. The spoken and written data presented in Sections 3 and 4 show that the distribution of DOM clearly differs between Cuban and European Spanish. This contrast is not mirrored in the questionnaire study for the acceptability of DOM, but rather for the acceptability of noDOM. From a more general point of view, we would like to stress that corpus data and judgment data are different methods which may unveil underlying contrasts. Still, combining both may provide a broader empirical coverage: They involve different types of data (production vs. acceptability) and thus provide different types of evidence that might or might not point into the same direction, as is the case with DOM retraction in Cuban Spanish.

---

**10** As an anonymous reviewer correctly points out, the question of whether an indefinite direct object is interpreted as specific or non-specific can be disambiguated by context. For example, this can be achieved by adding to the test items a further sentence containing a modal operator indicating the epistemic (non-)specificity of the indefinite direct object in question, such as 'I know X/I do not know X'. For our next experiments, we will introduce this modification in order to control for (epistemic) specificity.

## 6 Conclusions and discussion

Drawing on data from spontaneous speech, a diachronic corpus analysis, and grammaticality judgment tasks, this paper has given a synchronic and diachronic account of Differential Object Marking (DOM) in Cuban Spanish. The data from spontaneous speech reveals that DOM is less frequent in Cuban Spanish than in Buenos Aires, Madrid, and Mexico City Spanish. The lower frequency of DOM as compared to other varieties of Spanish raises the question of whether Cuban Spanish has retained a prior language stage or, rather, underwent DOM retraction. In this respect, Caribbean Spanish has previously been assumed to preserve a prior language stage of European Spanish (Pérez Guerra 1992, 489). More specifically, spoken Modern Cuban Spanish resembles written 16<sup>th</sup>-century European Spanish, especially with respect to DOM with definite human direct objects. Our diachronic corpus analysis has conclusively shown that DOM in written Modern Cuban Spanish is not a remnant of a prior language stage since it experienced a clear expansion between the 16<sup>th</sup> and 19<sup>th</sup> centuries (cf. Section 4.4 for details).

The data from spontaneous speech from the 20<sup>th</sup> century (Alfaraz 2011) together with the written corpus data from the 19<sup>th</sup> and 20<sup>th</sup> centuries (CORDE) favour the hypothesis that in Cuban Spanish, DOM underwent retraction. However, DOM retraction seems to be a rather recent phenomenon which began to develop in the 20<sup>th</sup> century. Moreover, it is much more evident in Alfaraz's (2011) spoken data from spontaneous speech than in our written corpus data from CORDE, where we could only detect a slight decrease of DOM with indefinite direct objects (from 56% to 43%), but not with definite direct objects. These differences might be due to the fact that spontaneous speech represents an informal style whereas written texts reflect a rather formal, more conservative language use (cf. Kock/De Mello 1997 for discussion on the *Habla Culta*). Indirect evidence that DOM retraction is a recent development in Cuban Spanish comes from other Caribbean varieties such as Puerto Rican Spanish. For example, López-Morales (1992, 141) indicates that DOM is less frequently found among young speakers.

While spontaneous speech and written language constitute instances of language production, grammaticality judgment tasks allow us to carefully evaluate the rate of acceptability with the presence and absence of DOM. In addition, they are diagnostic tools for detecting language change. In this respect, the judgment tasks conducted for Cuban and European Spanish have provided further evidence for the hypothesis of DOM retraction in Cuban Spanish. Interestingly, we could not find any difference between European and Cuban Spanish for the acceptability of sentences with DOM since both speakers of European and Cuban Spanish rated the test items with DOM as totally acceptable, both with human definite

and with human indefinite direct objects. However, we could observe clear differences with respect to the acceptability of sentences lacking DOM. Speakers of Cuban Spanish rated the absence of DOM with human definite and indefinite direct objects as highly acceptable, as opposed to their European counterparts. Similar observations have been made by Vaquero (1978), who conducted acceptability judgment tasks with university students of Puerto Rico, showing that the absence of DOM received an extremely positive evaluation, at least with human indefinite direct objects.

DOM variation, including both expansion and retraction, remains a promising research field, especially if we extend the empirical focus to more varieties within and beyond the Caribbean area, and if we attempt to model the factors conditioning variation. In order to take on this challenge, we can conclude from the present study that it is crucial to consider both production and acceptability data, and to analyze not only the conditions for the presence of DOM, but also those for the absence of DOM.

## Corpus

CORDE = Real Academia Española, *Banco de datos (CORDE) [en línea]. Corpus diacrónico del español*, <http://www.rae.es>, [last access: 01.02.2019].

## Bibliography

- Aissen, Judith, *Differential Object Marking. Iconicity vs. economy*, *Natural Language and Linguistic Theory* 21:3 (2003), 435–483.
- Alba, Orlando, *Cómo hablamos los dominicanos. Un enfoque sociolingüístico*, Santo Domingo, Grupo León Jimenes, 2004.
- Alfaraz, Gabriela G., *Accusative object marking. A change in progress in Cuban Spanish?*, *Spanish in Context* 8:2 (2011), 213–234.
- Álvarez Nazario, Manuel, *El habla campesina del país. Orígenes y desarrollo del español en Puerto Rico*, Río Piedras, Universidad de Puerto Rico, 1992.
- Balasz, Sonia, *Factors determining Spanish Differential Object Marking within its domain of variation*, in: Michnowicz, Jim/Dodsworth, Robin (edd.), *Selected proceedings of the 5<sup>th</sup> workshop on Spanish sociolinguistics*, Somerville, Mass., Cascadilla, 2011, 113–124.
- Barraza Carbajal, Georgina, *Maricación preposicional de objeto directo inanimado*, in: Company Company, Concepción/Moreno de Alba, José G. (edd.), *Actas del VII Congreso Internacional de Historia de la Lengua Española. Mérida (Yucatán), 4–8 septiembre de 2006*, Madrid, Arco, 2008, 341–352.



- Barrenechea, Ana María/Orecchia, Teresa, *La duplicación de objetos directos e indirectos en el español hablado en Buenos Aires*, in: Lope Blanch, Juan Manuel (ed.), *Estudios sobre el español hablado en las principales ciudades de América*, Ciudad de México, UNAM, 1977, 351–381.
- Barrios, Graciela, *Comportamiento de la preposición “a” con complementos directos en el español de Montevideo*, in: Elizaincín, Adolfo (ed.), *Estudios sobre el español del Uruguay*, Montevideo, Universidad de la República, 1981, 13–25.
- Bosson, Georg, *Empirische Universalienforschung. Differentielle Objektmarkierung in den neuiranischen Sprachen*, Tübingen, Narr, 1985.
- Bullock, Barbara E./Toribio, Almeida Jaqueline, *Reconsidering Dominican Spanish. Data from the rural Cibao*, *Revista Internacional de Lingüística Iberoamericana* 7:2 (2009), 49–73.
- Carter, Phillip M./Lynch, Andrew, *Multilingual Miami. Current trends in sociolinguistic research*, *Language and Linguistics Compass* 9:9 (2015), 369–385.
- Company Company, Concepción, *El avance diacrónico de la marcación prepositiva en objetos directos inanimados*, in: Bernabé Pajares, Alberto, et al. (edd.), *Presente y futuro de la lingüística en España. La Sociedad Lingüística, 30 años después. Actas del II Congreso de la Sociedad Lingüística*, vol. 2, Madrid, CSIC, 2002, 146–154 (= 2002a).
- Company Company, Concepción, *Grammaticalization and category weakness*, in: Wischer, Ilse/Diewald, Gabriele (edd.), *New reflections on grammaticalization*, Amsterdam/Philadelphia, John Benjamins, 2002, 201–215 (= 2002b).
- Cuza, Alejandro, *Introduction. New facts in Cuban Spanish research*, in: Cuza, Alejandro (ed.), *Cuban Spanish dialectology. Variation, contact, and change*, Washington DC, Georgetown University Press, 2017, xix–xxvii.
- Dalrymple, Mary/Nikolaeva, Irina, *Objects and information structure*, Cambridge, Cambridge University Press, 2011.
- Delille, Karl Heinz, *Die geschichtliche Entwicklung des präpositionalen Akkusativs im Portugiesischen*, Bonn, Romanisches Seminar der Universität Bonn, 1970.
- Domínguez, Carmen Luisa, et al., *Observaciones sobre el uso de la preposición “a” en el objeto directo. Un estudio sobre el español de Mérida*, *Letras* 59 (1998), 89–120.
- Dulme, Erika Ehnis, *El uso de la preposición “a” ante objeto directo en el habla popular de la ciudad de México*, in: Moreno de Alba, José G. (ed.), *Actas del II Congreso Internacional sobre el Español de América, Ciudad de México, 27–31 de enero de 1986*, México, UNAM, 1986, 404–407.
- Dumitrescu, Domnita, *El parámetro discursivo en la expresión del objeto directo lexical. Español madrileño vs. español porteño*, *Signo y Seña* 7 (1997), 305–354.
- García García, Marco, *Differentielle Objektmarkierung bei unbelehten Objekten im Spanischen*, Berlin/Boston, de Gruyter, 2014.
- García García, Marco, *Nominal and verbal parameters in the diachrony of Differential Object Marking in Spanish*, in: Seržant, Ilja A./Witzlack-Makarevich, Alena (edd.), *Diachrony of Differential Argument Marking*, Berlin, Language Science Press, 2018, 209–242.
- García García, Marco/Primus, Beatrice/Himmelmann, Nikolaus P., *Shifting from animacy to agentivity*, *Theoretical Linguistics* 44:1–2 (2018), 25–39.
- von Heusinger, Klaus, *Verbal semantics and the diachronic development of DOM in Spanish*, *Probus* 20 (2008), 1–31.
- von Heusinger, Klaus, *The diachronic development of Differential Object Marking in Spanish ditransitive constructions*, in: Seržant, Ilja A./Witzlack-Makarevich, Alena (edd.), *Diachrony of Differential Argument Marking*, Berlin, Language Science Press, 2018, 315–344.

- von Heusinger, Klaus/Kaiser, Georg A., *Affectedness and Differential Object Marking in Spanish*, *Morphology* 21:3–4 (2011), 593–617.
- Hoff, Mark R., *Are Argentines “a”-blind? Acceptability of “a”-marked inanimate direct objects*, in: MacDonald, Jonathan E. (ed.), *Contemporary trends in Hispanic and Lusophone linguistics. Selected papers from the Hispanic Linguistic Symposium 2015*, Amsterdam/Philadelphia, John Benjamins, 2018, 121–142.
- Hoff, Mark R./Díaz-Campos, Manuel, *Conciencia y actitudes. El caso de la “a” acusativa en el español argentino*, *Signo y Seña* 28 (2015), 89–110.
- Jiménez Sabater, Max, *Más datos sobre el español de la República Dominicana*, Santo Domingo, INTEC, 1975.
- Kabatek, Johannes, *Wohin strebt die differentielle Objektmarkierung im Spanischen?*, *Romanistisches Jahrbuch* 67:1 (2016), 211–239.
- Keniston, Hayward, *The syntax of Castilian prose. The sixteenth century*, Chicago, University of Chicago Press, 1937.
- Kock, Josse de/De Mello, George (edd.), *Lengua escrita y habla culta en América y España. Diez casos*, Salamanca, Universidad de Salamanca, 1997.
- Laca, Brenda, *El objeto directo. La marcación preposicional*, in: Company Company, Concepción (ed.), *Sintaxis histórica del español. Primera parte: La frase verbal*, vol. 1, México, UNAM/Fondo de Cultura Económica, 2006, 423–475.
- Leonetti, Manuel, *Specificity and Differential Object Marking in Spanish*, *Catalan Journal of Linguistics* 3 (2004), 75–114.
- Lizárraga Navarro, Glenda Zoé/Mora-Bustos, Armando, *Variación en la marcación diferenciada de objeto en español*, *Forma y Función* 23:1 (2010), 9–38.
- Lope Blanch, Juan Manuel, *El estudio del español hablado culto. Historia de un proyecto*, México, UNAM, 1986.
- López-Morales, Humberto, *El español del Caribe*, Madrid, MAPFRE, 1992.
- Lunn, Patricia, *Tout se tient in Dominican Spanish*, in: Lee, James F./Geeslin, Kimberly L./Clements, J. Clancy (edd.), *Structure, meaning, and acquisition in Spanish. Papers from the 4<sup>th</sup> Hispanic Linguistics Symposium*, Somerville, Mass., Cascadilla, 2002, 65–72.
- Mayer, Elisabeth/Delicado Cantero, Manuel, *Continuity and innovation in Peruvian Spanish. Pragmatics and contact in (Differential) Object Marking*, in: Sessarego, Sandro/González-Rivera, Melvin (edd.), *New perspectives on Hispanic contact linguistics in the Americas*, Madrid, Vervuert, 2015, 99–120.
- Melis, Chantal, *El objeto directo personal en “El Cantar de Mio Cid”. Estudio sintáctico-pragmático*, in: Pensado, Carmen (ed.), *El complemento directo preposicional*, Madrid, Visor, 1995, 133–163.
- Mendoza Quiroga, José G., *Aspectos del castellano hablado en Bolivia*, in: Hernández Alonso, César (ed.), *Historia y presente del español de América*, Ponce, Junta de Castilla y León, 1992, 437–500.
- Montrul, Silvina, *La marcación diferencial del objeto directo en el español de Argentina. Un estudio experimental*, in: Colantoni, Laura/Rodríguez Louro, Celeste (edd.), *Perspectivas teóricas y experimentales sobre el español de la Argentina*, Madrid, Vervuert, 2013, 207–228.
- Montrul, Silvina, *Structural changes in Spanish in the United States. Differential Object Marking in Spanish heritage speakers across generations*, *Lingua* 151, Part B (2014), 177–196.

- Octavio de Toledo y Huerta, Álvaro S., *Aprovechamiento del CORDE para el estudio sintáctico del primer español moderno (ca. 1675–1825)*, in: Kabatek, Johannes (ed.), *Lingüística de corpus y lingüística histórica iberorrománica*, Berlin/Boston, de Gruyter, 2016, 57–89.
- Ordóñez, Francisco/Treviño, Esthela, *Agreement and DOM with impersonal “se”*. A comparative study of Mexican and Peninsular Spanish, in: Kato, Mary/Ordóñez, Francisco (edd.), *The morphosyntax of Portuguese and Spanish in Latin America*, Oxford, Oxford University Press, 2016, 236–258.
- Ortiz Ciscomani, Rosa María, *Los objetos concurrentes y la bitransitividad en el español en perspectiva diacrónica*, in: Eddington, David (ed.), *Selected proceedings of the 7<sup>th</sup> Hispanic Linguistics Symposium*, Somerville, Mass., Cascadilla Press, 2005, 192–202.
- Ortiz Ciscomani, Rosa María, *Construcciones bitransitivas en la historia*, México, UNAM, 2011.
- Pérez Guerra, Irene, *Aportación a un tema en debate en el Caribe hispánico. El arcaísmo del español dominicano*, in: Ariza Viguera, Manuel (ed.), *Actas del II Congreso Internacional de Historia de la Lengua Española*, Madrid, Pabellón de España, 1992, 483–490.
- Régulo Pérez, Juan, *Notas acerca del habla de la isla de la Palma*, *Revista de Historia Canaria* 32 (1970), 12–174.
- Romero Heredero, Diego, *Marcado diferencial de objeto y semántica verbal en español*, Berlin, de Gruyter, in press.
- Tippets, Ian, *Differential Object Marking in Spanish. A quantitative variationist study*, Ohio, Ohio State University PhD thesis, 2010.
- Tippets, Ian, *Differential Object Marking. Quantitative evidence for underlying hierarchical constraints across Spanish dialects*, in: Ortiz-López, Luis A. (ed.), *Selected proceedings of the 13<sup>th</sup> Hispanic Linguistics Symposium*, Somerville, Mass., Cascadilla, 2011, 107–117.
- Vaquero, María, *Enseñar español, pero ¿qué español?*, *Boletín de la Academia Puertorriqueña de la Lengua Española* 6:1 (1978), 127–146.
- Witzlack-Makarevich, Alena/Seržant, Ilja A., *Differential Argument Marking. Patterns of variation*, in: Seržant, Ilja A./Witzlack-Makarevich, Alena (edd.), *Diachrony of Differential Argument Marking*, Berlin, Language Science Press, 2018, 1–40.

## Appendix 1: CORDE sources

This appendix contains the CORDE sources employed for the diachronic corpus analysis (cf. Section 4). The sources are arranged according to century, author, and record. The number of words is given in brackets.

### Sources

Century	Author	Record
19 <sup>th</sup> century	Auber Noya	<i>Ambarina</i> (97,206)
	El Cucalambé	<i>Consecuencias de una falta</i> (9,744); <i>Poesías completas</i> (60,283)
	Gómez de Avellaneda	<i>Dolores</i> (23,973); <i>Dos mujeres</i> (83,824); <i>El artista barquero o los cuatro 5 de junio</i> (68,640); <i>El cacique de Turmequé</i> (17,617); <i>Espatolino</i> (49,301); <i>La baronesa de Joux</i> (12,553); <i>La dama de Amboto</i> (2,262); <i>La flor del ángel</i> (6,935); <i>La montaña maldita</i> (4,165); <i>La ondina del lago azul</i> (10,590); <i>La velada del helecho o el donativo del diablo</i> (19,160); <i>Poesías</i> (43,221); <i>Una anécdota de la vida de Cortés</i> (4,802)
	Heredia	<i>Carta de Boston</i> (621); <i>Poesías</i> (36,951); <i>Teatro 3. Una ópera de Rossini</i> (506)
	Hostos	<i>La peregrinación de Bayoán</i> (85,248)
	Martí	<i>La Edad de Oro</i> (66,806); <i>Lucía Jerez</i> (36,099)
	Ortega Munilla	<i>Cleopatra Pérez</i> (47,098)
20 <sup>th</sup> century	Anónimo	<i>¡Oh, mío Yemayá!</i> (27,527)
	Augier	<i>Prosa varía</i> (126,005)
	Bobadilla	<i>A fuego lento</i> (64,073)
	Cabrera Infante	<i>Tres tristes tigres</i> (137,073)
	Carlos Felipe	<i>De película</i> (8,574); <i>El Chino</i> (18,061); <i>El travieso Jimmy</i> (25,885); <i>Réquiem por Yarini</i> (22,989)
	Carpentier	<i>El siglo de las luces</i> (117,871)
	Fernández Retamar	<i>Fervor de la Argentina</i> (42,768)
	Guillén	<i>El son entero</i> (5,144)
	Insúa	<i>El negro que tenía el alma blanca</i> (70,841)
	Lachatañaré	<i>El sistema religioso de los lucumí y otras influencias africanas en Cuba</i> (75,304); <i>Manual de santería</i> (16,181)
	Lezama Lima	<i>Paradiso</i> (220,057)

(continued)

Loynaz	<i>Jardín</i> (104,393)
Ortega Munilla	<i>Los tres sorianitos</i> (22,669)
Ortiz	<i>Contrapunteo cubano del tabaco y el azúcar</i> (208,650); <i>Los negros brujos</i> (80,325)
Serpa	<i>Contrabando</i> (69,565)

## Appendix 2: Test items of the questionnaire

This appendix lists the test items of the European Spanish version of the questionnaire employed for the judgment tasks (cf. Section 5). The items are arranged according to definite and indefinite human direct objects. Note that they exhibit DOM although the presence and absence of DOM was altered and pseudo-randomized in the questionnaires.

Test items with definite human direct objects:

1. *Gonzalo mató al mafioso en el parque.*  
'Gonzalo killed the mobster in the park.'
2. *Patricio lesionó al portero en la discoteca.*  
'Patricio injured the doorman at the nightclub.'
3. *Carolina golpeó al futbolista en el vestuario.*  
'Carolina hit the soccer player in the locker room.'
4. *Elena cuidó al herido durante dos semanas.*  
'Elena took care of the injured person for two weeks.'
5. *Irene acusó al secretario general en la asamblea.*  
'Irene accused the general secretary at the assembly.'
6. *Cecilia denunció al ladrón en la comisaría de policía.*  
'Cecilia reported the thief at the police station.'
7. *Josefina vio al cura en el parque.*  
'Josefina saw the priest in the park.'
8. *Julio oyó al basurero desde la cocina.*  
'Julio heard the garbage collector from the kitchen.'

Test items with indefinite human direct objects:

9. *Pablo mató a un rehén durante el secuestro.*  
'Pablo killed a hostage during the kidnapping.'

10. *Alberto lesionó a un defensa en el partido de la semana pasada.*  
'Alberto injured a defender during last week's match.'
11. *Andrea golpeó a un fotógrafo a la salida del restaurante.*  
'Andrea hit a photographer at the restaurant's exit.'
12. *Javier cuidó a un enfermo con gran profesionalidad.*  
'Javier took care of an injured person very professionally.'
13. *Laura acusó a un trabajador en la reunión.*  
'Laura accused a worker at the meeting.'
14. *Carlos denunció a un camarero tras una disputa.*  
'Carlos reported a waiter after an argument.'
15. *Ana vio a un político en la tienda del barrio.*  
'Ana saw a politician at the neighbourhood's shop.'
16. *Esther oyó a un bebé en el apartamento de al lado.*  
'Esther heard a baby in the apartment next door.'