**Leísmo or fake leísmo?**

**New insights into Catalan Contact Spanish from the FEC Corpus**

¿Leísmo o falso leísmo? Nuevas miradas sobre el catalán en contacto con el español a partir del corpus FEC

**Monja Burkard**

*University of Vienna*

**ABSTRACT.** The aim of this study is to describe the use of the pronouns *le* and *les* in the Spanish spoken in the province of Barcelona. The paper analyzes a sample of 12 reading tasks and 22 oral interviews taken from two corpora: the FEC (Fonología del Español Contemporáneo, Pustka *et alii* 2018) corpus (with mainly Catalan-dominant bilingual speakers) and the Corpus oral de profesionales de la lengua castellana en Barcelona (Sinner 2001, with mainly Spanish-dominant bilinguals). Several linguistic variables as well as the language dominance of the speakers are taken into account in order to find out 1) if there is *leísmo* in Catalan Contact Spanish (although there is no equivalent of *leísmo* in Catalan) and if so, 2) which features of the referent and of the verb trigger *leísmo* in this variety, and 3) if Catalan-dominant bilinguals produce *leísmo* to a lesser extent than Spanish-dominant speakers. Whereas the results of the reading task suggest that *leísmo* is not absent in the Spanish of Barcelona, there are only a few cases of *leísmo* in spontaneous speech in both corpora. Regarding the linguistic variables, it turns out, on the one hand, that *leísmo* is not restricted to *leísmo correcto* in the corpora, and on the other, that the majority consists of fake *leísmo* cases. Thus, the data seem to suggest that the Spanish of Barcelona is only a fake *leísmo* variety. With respect to the language dominance, however, the result is that Catalan-dominant
bilinguals do not produce fewer cases of *leísmo*, since the *leísmo* rate is higher in the FEC corpus than in the Sinner corpus.

**Keywords:** Clitic pronouns, *leísmo*, fake *leísmo*, Spanish, Catalan Contact Spanish, bilingualism

**Resumen.** El objetivo de este estudio es describir el uso de los pronombres *le* y *les* en el español hablado en la provincia de Barcelona. El artículo analiza una muestra de 12 tareas de lectura y de 22 entrevistas orales extraídas de dos corpus: el FEC (*Fonología del Español Contemporáneo, Pustka et alii* 2018, con hablantes bilingües cuya lengua dominante es el catalán) y el *Corpus oral de profesionales de la lengua castellana en Barcelona* (Sinner 2001, con hablantes bilingües cuya lengua dominante es el español). Se toman en consideración algunas variables lingüísticas, así como la lengua dominante de los hablantes, para descubrir 1) si hay *leísmo* en el catalán en contacto con el español (aunque no hay un *leísmo* equivalente en catalán) y si es así, 2) qué rasgos del referente y del verbo activan el *leísmo* en esa variedad, y 3) si los hablantes bilingües con predominio de catalán presentan *leísmo* en menor medida que los que tienen como lengua dominante el español. Mientras que los resultados de la tarea de lectura sugieren que el *leísmo* no está ausente en el español de Barcelona, hay muy pocos casos de *leísmo* en el habla espontánea de ambos corpus. Con respecto a las variables lingüísticas, por una parte, resulta que el *leísmo* no está restringido en los corpus al *leísmo correcto* y, por otra, la mayoría son casos de falso *leísmo*. Por tanto, los datos parecen sugerir que el español de Barcelona ofrece solo una variedad de falso *leísmo*. Con respecto a la lengua dominante, sin embargo, el resultado es que los hablantes bilingües con predominio del catalán no presentan menos casos de *leísmo*, puesto que la ratio de *leísmo* es más alta en el corpus FEC que en el corpus de Sinner.

**Palabras clave:** pronombres clíticos, *leísmo*, falso *leísmo*, español, español en contacto con el catalán, bilingüismo

**1. INTRODUCTION**

Over the last 20 years, there has been an increasing interest in investigating the features of Catalan Contact Spanish, the variety of Spanish spoken in Catalonia, Spain (cf. Boix-Fuster & Sanz 2008: 102). While most of the studies on this topic have concentrated on contact in general (cf. Galindo Solé 2003, Borràs Castanyer *et alii* 1997; Wesch 1997; Blas Arroyo 2011), lexicon (cf. Casanovas Catalá 2000, 2008), phonetics (cf. Serrano Vázquez 1996-1997; Wesch 1997; Davidson 2012), and pragmatics (cf. Vann 2007), less attention has been paid to morphosyntax (cf. Sinner 2004; Illamola & Vila 2015) and in particular to the production of clitics (cf. Jiménez-Gaspar, Pires & Guijarro-Fuentes 2017). The present study seeks to contribute to this issue by examining the phenomenon of *leísmo* in Catalan Contact Spanish in the
province of Barcelona. As opposed to some older work that restricts *leísmo* to Northern and Central Spain (cf. Quilis et alii 1985), we suggest that *leïsmo* does indeed exist in this variety. Yet, in line with more recent research (cf. Fernández-Ordóñez 2012, Roselló Verdeguer 2017), it is argued that real *leísmo* must be distinguished from only apparent (fake) *leïsmo*, and it is shown that Barcelona Catalan Contact Spanish only exhibits *leïsmo* of the latter type.

The paper is structured as follows. Section 2 offers a brief description of the phenomenon of *leïsmo*, including different types of *leïsmo* and its geographical distribution. Section 3 consists of a review of prior studies on *leïsmo* in the Catalan-speaking area. Section 4 presents my research questions. Section 5 introduces the corpora the data were taken from, as well as the principles of the analysis. Section 6 presents the results, which are discussed in Section 7. I conclude in Section 8 by providing an outline of some open questions.

2. **LEÍSMO**

2.1 **Definition**

The term *leïsmo* refers to the Spanish phenomenon in which the dative pronoun *le(s)* is used instead of the accusative pronoun *lo(s)/ la(s)*. As can be seen in Table 1, in Standard Spanish, the third person unstressed pronoun system derives from the Latin system, which distinguished between case, gender, and number. While the Spanish accusative pronouns *lo, la,* and *lo* derive from the Latin accusative forms *illum, illam, illud,* the dative pronoun *le* stems from the dative form *illi* (cf. Fernández-Ordóñez 1999). Therefore, the case-distinguishing usage is also known as the “etymological” system.

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th></th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>masculine</td>
<td>feminine</td>
<td>neuter</td>
</tr>
<tr>
<td>accusative</td>
<td><em>lo</em></td>
<td><em>la</em></td>
<td><em>lo</em></td>
</tr>
<tr>
<td>dative</td>
<td><em>le</em></td>
<td></td>
<td><em>les</em></td>
</tr>
</tbody>
</table>

**Table 1. Case-distinguishing system**

---

1 We use the term *standard* here, since the DPD (*Diccionario panhispánico de dudas* 2005) regards the case-distinguishing system as the norm of Standard Spanish and refers to *leïsmo* as the “improper use of *le(s)* in the function of a direct object” (DPD 2005: 392; my translation from Spanish). However, due to its frequency, Klein-Andreu (1992: 170) suggests that *leïsmo* with animate masculine referents can be understood as the de facto standard.
In contrast, *leismo*-varieties simplify the pronouns system by collapsing not only case but also gender distinctions, as we can see in example 1 (clitics are in bold print and the referents are underlined here and in the subsequent examples):

(1) a) **Lo** vi ayer. (Standard Spanish)  
3.M.ACC.SG I saw yesterday  
‘I saw him/it yesterday.’

b) **La** vi ayer. (Standard Spanish)  
3.F.ACC.SG I saw yesterday  
‘I saw her yesterday.’

c) **Le** vi ayer. (*Leismo* varieties)  
3.DAT.SG I saw yesterday  
‘I saw him/her yesterday.’

2.2 Types

The literature on *leismo* distinguishes several types of leismo regarding the features of the referent. The most frequent and widespread type of *leismo* is *leismo* with [+human], [–feminine], [+singular] referents (see 2a) below. Its use has been accepted by the RAE (1974: 425)² and is therefore often referred to as *leismo correcto*. With any other features of the referent, *leismo* is far less widespread and far less accepted (and therefore called *leismo incorrecto*). The latter type comprises *leismo* with inanimate masculine referents in the singular (2b) as well as *leismo* with plural forms, where however it is more frequent if the referent is [+human] (2c) than it is with inanimate referents (2d). The most infrequent but still attested type is *leismo* with feminine referents (2e, f), whereas *leismo* is not attested with neuter referents (Fernández-Ordóñez 1999: 1).

(2) a) ¿Conoces a **Juan**? Sí, le conozco hace tiempo. ‘Do you know Juan? Yes, I’ve known him for a long time.’

b) ¿Sabes dónde está mi libro? No, no le he visto por aquí. ‘Do you know where my book is? No, I haven’t seen it around here.’

c) Esta tarde voy a recoger a los niños del colegio y les llevaré al parque. ‘This evening I’m going to pick up the children from school and take them to the park.’

d) Fui a buscar los discos que querías y les encontré en la tienda de abajo. ‘I went to get the records that you wanted and I found them in the shop downstairs.’

---

² The *Real Academia Española* even regarded it as the preferred form until the fourth edition of its *Gramática* from 1796.
Leísmo or fake leísmo? New insights into Catalan Contact Spanish from the FEC Corpus

Leísmo or fake leísmo? New insights into Catalan Contact Spanish from the FEC Corpus

In addition to this division into different subtypes of leísmo regarding the features of the referent, there is a basic distinction made by Fernández-Ordóñez (1999) in the Gramática descriptiva de la Lengua Española. She argues that many cases traditionally regarded as leísmo are not true instances of leísmo, and that specific verbs and verb constructions have to be taken into account as well. She calls this leísmo aparente, ‘apparent’ or fake leísmo, i.e. cases in which the use of le is not due to a loss of the case distinction, but instead either to true case alternations that come along with a semantic change or to dialectal variation in case assignment. This is the case with the following verbs or verb constructions:

a) Experiencer verbs such as aburrir ‘to bore’, asustar ‘to shock’, decepcionar ‘to disappoint’, divertir ‘to amuse’, impresionar ‘to impress’, molestar ‘to disturb’. With these verbs, there are two structures possible: an agentive one in which the object is pronominalized in the accusative, and a non-agentive one in which the object is pronominalized in the dative. Thus, the use of le basically depends on the animacy of the subject: if it is inanimate, the object tends to pronominalize in the dative, whereas with animate subjects, the object pronominalizes mostly in the accusative, e.g.:

i. A mi hijo lo asustó aquel perro. ‘My son was frightened by the dog.’
ii. A mi hijo le asustan los truenos. ‘My son is afraid of thunder.’ (Fernández-Ordóñez 1999: 1324)

b) The subject of infinitive clauses with verbs such as obligar ‘to oblige’, autorizar ‘to authorize’, or ordenar ‘to order’: commonly, these verbs pronominalize their objects in the accusative. Some regions (such as La Mancha and Andalusia), however, preserve the ancient pattern, according to which all types of causative verbs require the dative case: Le ordené ejecutar la sentencia ‘I ordered him to execute the sentence’ (DPD 2005: 394).

c) Verbs that can omit their direct object, such as atender ‘to attend to’, servir ‘to serve’, pagar ‘to pay’, robar ‘to rob’, seguir ‘to follow’, tocar ‘to touch’: according to Fernández-Ordóñez (1999: 1328-1329) these verbs can omit their direct object. In that case, the pronominalized object in the dative is not a direct but an indirect object, and it goes along with a change of meaning. Thus, for example, atender can also mean ‘to listen to’ when used without the direct object, e.g.
**Cuando el profesor hable en la clase, debéis atenderle** ‘When the teacher talks in class, you must listen to him’ (Fernández-Ordóñez 1999: 1328).

d) Verbs whose case government has been reinterpreted, e.g. *ayudar* ‘to help’, *aconsejar* ‘to advise’, *avisar* ‘to advise’, *obedecer* ‘to obey’: in medieval Spanish, these verbs used to govern the dative case, whereas nowadays they are more often used with the accusative\(^3\), but still exhibit some case variation, e.g.

i. *María está muy atareada, pero Juan le está ayudando.* ‘Maria is very busy, but Juan is helping her’

ii. *María está muy atareada, pero Juan la está ayudando.* (Fernández-Ordóñez 1999: 1332-1333)

e) Verbs with objects that require a predicative complement, such as *llamar*: when used in the sense of ‘to address as’ this verb required a double accusative construction in Latin (DPD 2005: 401), but was reinterpreted in Romance as a transitive verb with an obligatory predicative complement. Yet, variation is attested in Old Spanish already and has been maintained up to the present (Fernández-Ordóñez 1999: 1335-1336):

i. *En España a la papa le llaman patata.* ‘In Spain the potatoe is called *patata*’

ii. *La llamaré Juana.* ‘I’ll call her Juana’ (DPD, 2005: 401)

f) Impersonal *se* followed by a clitic pronoun: constructions with impersonal *se* combined with the pronoun *le* referring to a direct object have often been regarded as a form of *leísmo*. However, according to Fernández-Ordóñez (1999) these constructions originally required *le*, regardless of the gender of the object, e.g. *Se le considera el mejor actor de su tiempo* ‘He is considered the best actor of his time’ (DPD 2005: 395).

As already mentioned, none of these verbs or verb constructions employed with *le(s)* count as real *leísmo* for Fernández-Ordóñez (1999). Whereas many older studies do not consider this distinction, this paper will incorporate this aspect. Moreover, it is claimed that many of the existing studies on *leísmo* need to be reanalyzed with regard to specific verbs and verb constructions.

### 2.3 Geographical distribution

*Leísmo* is attested in the western part of Castile as well as in adjacent regions such as the Asturian border area with Cantabria, Cantabria itself, and the Basque Country (cf. Fernández-Ordóñez 2012). However, the pronoun systems are not at

---

\(^3\) According to Pineda (2013) these verbs have undergone a process of monotransitivitation.
all uniform in the leísmo varieties, but rather can be divided into several subsystems (cf. Fernández-Ordóñez 1994, Gómez Seibane 2012). The map in Figure 1 shows the geographical distribution of these subsystems. According to Klein-Andreu (1981) and Fernández-Ordóñez (1994), the referential system (from north-western Castile to La Mancha), as well as the system in Cantabria and the transitional systems in the West and Northeast (sistema de transición occidental/ nororiental), exhibit - apart from gender, number, and animacy - the feature [±countable] as an additional factor that triggers - to different degrees - the use of le in these regions. In the Basque Country, speakers whose L1 is Basque most often omit accusative clitics, but with the increase of Spanish dominance, they tend to use le(s) for animate referents and lo(s)/ la(s) (or omission) for inanimate referents in the accusative (cf. Fernández-Ordóñez 1994).

The literature on leísmo in Catalan Contact Spanish is scarce. Therefore, the Catalan-speaking area is missing on the map. The few existing studies will be presented in the following section.

---

4 Roughly speaking, in the way that countable referents tend to trigger the leísmo-form.

5 In some studies, the Catalan-speaking area is deliberately excluded, e.g. in León Zurdo 2017, who compares the use of leísmo in two corpora: ALPI (Navarro Tomás 1962) and COSER (2005). However, she fails to take the Catalan-speaking area into account, although both corpora include data from this region.
3. **LEÍSMO IN THE CATALAN-SPEAKING AREAS**

According to Llorente (1980: 22), systematic investigation of the geographical distribution of *leísmo* was still lacking up to the 1980s. One the one hand, he admits that there are no convincing data for the Catalan-speaking area:

> Los datos de que yo dispongo son los referentes a Andalucía […], a Canarias […], a Aragón, Navarra y Rioja, a Extremadura, Salamanca, Zamora y parte occidental de Ávila. Del resto del dominio, sólo los datos dispersos e imprecisos que se encuentran en las monografías dialectales, en los manuales y tratados y en algunos trabajos que tocan el tema de manera más o menos directa, y no resultan del todo convincentes. (Llorente 1980: 22-23)

On the other hand, however, he assumes *leísmo* to be absent in the Catalan-speaking area:

> Confiemos en que los andaluces, murcianos, aragoneses, hispanoamericanos y los hispanohablantes de Cataluña, Levante y Baleares resistan denotadamente, se muestren inmunes a la infección y sigan, en este aspecto concreto, hablando y escribiendo con corrección y elegancia nuestra lengua, como lo han hecho hasta ahora (Llorente 1980: 26-27)

A similar claim was made by Casanovas Català (1996) for Lleida, who did not detect any instance of *leísmo*. However, this result is also not based on quantitative data.

One of the first to include the Catalan-speaking area in the empirical research of *leísmo* was Klein-Andreu (1992). She investigated the acceptance of *le* in accusative contexts with respect to the animacy of the referent, comparing monolingual speakers with bilingual speakers (from Catalonia/ Valencia, Basque Country and Galicia). According to her results, Catalan/ Valencian bilinguals accept 48% of all occurrences of *le* when the referent is animate. In a second analysis of written language, she comes to the conclusion that the Catalan writers Vázquez Montalbán and Porcel use *le* in accusative contexts in 91% of all occurrences, as long as the referent is human and masculine.

A second study that suggests that at least *leísmo correcto* is not absent in the Catalan-speaking area was realized by Blas Arroyo (1994) for Valencian Spanish. He conducted complementation tasks and acceptability judgement tasks, as well as semi-structured interviews, with 88 young bilinguals from Valencia. The analysis of all the three datasets shows that *leísmo* (with human masculine referents) is in fact the
preferred option (in the interviews and complementation task leísmo rates are around 80%), whereas leísmo incorrecto is very infrequent.

Based on all of these results, Gómez Seibane (2012) proposes the following (provisional) system for the Catalan-speaking area:

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th></th>
<th></th>
<th>masculine</th>
<th>feminine</th>
<th></th>
<th>masculine</th>
<th>feminine</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>masculine</td>
<td>feminine</td>
<td>neuter</td>
<td>masculine</td>
<td>feminine</td>
<td></td>
<td>masculine</td>
<td>feminine</td>
</tr>
<tr>
<td></td>
<td>animate</td>
<td>inanimate</td>
<td></td>
<td>animate</td>
<td>inanimate</td>
<td></td>
<td>animate</td>
<td>inanimate</td>
</tr>
<tr>
<td>accusative</td>
<td>le – lo</td>
<td>lo</td>
<td>la (&gt; le)</td>
<td>lo</td>
<td>les – los</td>
<td>los</td>
<td>las</td>
<td></td>
</tr>
<tr>
<td>dative</td>
<td>le</td>
<td>le</td>
<td>le (&gt; la)</td>
<td>le</td>
<td>les</td>
<td>les</td>
<td>les (&gt; las)</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Provisional system for the Catalan-speaking area according to Gómez-Seibane (2012: 30)

Ordóñez & Roca (2017) treat leísmo in Catalan Contact Spanish from a theoretical point of view. Concentrating their analysis on causative structures (e.g., *Le hicimos reír mucho* ‘We made him/ her laugh a lot’), they conclude that the collapse of gender distinction is due to the placement of the argument of the causative structure in an external position with respect to the infinitive verb phrase, which triggers the expression of this argument in the dative. Yet, the authors do not specify if their analysis is based on new empirical data or on the description by Gómez Seibane (2012).

However, more recent work challenges Gómez Seibane’s (2012) description, as does the study by Roselló Verdeguer (2017) on Valencian Spanish. He works with the corpus PRESEEA-Valencia (PRESEVAL), which consists of 72 interviews with both Spanish monolinguals from Valencia and Spanish-Valencian bilinguals. Crucially, his results are not consistent with Blas Arroyo’s (1994) findings. Firstly, there are only a few occurrences of leísmo in his corpus: of all instances of *le* (955), only 109 function as direct objects, which corresponds to a percentage of 11.4% (for plural *les* the rate is 50/390, i.e. 12.8%). When we consider all instances of pronominalized direct objects (2781), the percentage of *le* is 3.9% (109/2781) and of *les* 1.8% (50/2781). Secondly, he also includes the verbs and verb constructions introduced by Fernández-Ordóñez (1999) and finds that 71.6% of the occurrences of *le* as a direct object (i.e. 78/109) are in fact cases of fake leísmo (for the plural *les* it is 33/50, i.e. 66%). Hence, in total, out of the 159 cases of leísmo only 48 (30.18%) constitute instances of ‘real’ leísmo. In sum, Roselló Verdeguer (2017) concludes that Valencian Spanish is not a leísmo

---

6 A great majority of these are instances of leísmo correcto (human, masculine singular referents).
variety, even though *le(s)* is common with specific verbs and verb constructions (fake *leísmo* contexts).

In summary, while older work does not consider Catalan Contact Spanish as a *leísmo* variety, some studies conducted in the 1990s come to the result that *leísmo* does exist in the Catalan-speaking area. However, Roselló Verdeguer’s (2017) results suggest that these studies may have overgeneralized the concept of *leísmo*. Since they were published before the introduction of the phenomenon of fake *leísmo* by Fernández-Ordóñez (1999), and do not consider specific verb constructions, we do not know if what they identify as *leísmo* does in fact correspond to real *leísmo* or not.

4. RESEARCH QUESTIONS

In the light of results mentioned in the previous section, we should re-raise the question whether Catalan Contact Spanish is a *leísmo* variety or not, taking into account the distinction between real and fake *leísmo* made by Fernández-Ordóñez (1999). This study aims to contribute to this question. While since most of the previous studies focused on other regions such as Valencia or Lleida, I chose the province of Barcelona as the location of research.

First of all, this paper seeks to find out if *leísmo* exists in the spoken variety of the province of Barcelona and, secondly, which type of *leísmo* we find there. Gómez Seibane’s (2012) hypothesis (see Table 2) suggests that we will most likely find *leísmo correcto*, but it is also possible to detect cases of *leísmo* with feminine or inanimate referents. Apart from the features of the referent, the present analysis also includes the type of verbs and verb constructions in order to decide if the instances of *leísmo* are cases of true *leísmo* or of fake *leísmo*. The third question concerns linguistic dominance. Since there is no equivalent of *leísmo* in Catalan (which is a case-distinguishing system), the question arises whether (if *leísmo* exists at all in this area) Catalan-dominant bilinguals produce less *leísmo*.

Summing up, the research questions are as follows:

1) Is there *leísmo* in the Catalan Contact Spanish of Barcelona?
2) If so, which features of the referent and of the verb trigger *leísmo* in this variety?
3) Do Catalan-dominant bilinguals produce fewer cases of *leísmo*?
5. METHOD
5.1 The province of Barcelona as the research location

The province of Barcelona is characterized by its extensive contact between Catalan and Spanish. While in the rest of Catalonia, Catalan speakers are in the majority, in the bigger cities like Barcelona (and its metropolitan area), as well as in Lleida and Tarragona, the use of Spanish has increased during the last 50 years, and in some areas, Spanish speakers now constitute over 50% of the population (cf. Vila-Pujol 2007: 64). This is due, on the one hand, to the repression\(^7\) of the Catalan language during the Franco regime (Vallverdú 1984) and, on the other hand, to the mass immigration to Catalonia and other industrial regions of Spain in the second half of the 20\(^{th}\) century (cf. Blas Arroyo 2011: 374; Recolons 1987). Immigration to Catalonia still increases, and the latest census of linguistic use in Catalonia (IEC 2015) shows that only 27.8% of the population of Barcelona and its metropolitan area regards Catalan as their habitual language (compared to 60% who consider Spanish to be their habitual language). Nevertheless, Catalan is still highly present in the province of Barcelona. According to the same survey, 97.7% of the population indicate to understand and 77.6% to speak Catalan. Furthermore, under a law of language politics passed in 1983, Catalan has to be the main language of instruction in all primary and secondary schools in Catalonia\(^8\).

In sum, as pointed out by Davidson (2012), Barcelona is a promising location for conducting research on Catalan Contact Spanish due to the prolonged contact between Catalan and monolingual Spanish speakers.

5.2 Corpora

The data for this analysis were taken from two corpora. Both of them are orthographically transcribed and are available in searchable documents. The first one is FEC Fonología del Español Contemporáneo (Pustka et alii 2018), which actually aims to document the pronunciation of Spanish around the world. In addition to sociodemographic information collected via a questionnaire, it contains recordings of two reading tasks (word list and text) and a semi-structured interview with questions about culture, leisure activities, and linguistic awareness. Following the FEC protocol, we recorded

---

\(^7\) Catalan was banned from public sphere, including education, and publications in Catalan were very limited (Vallverdú 1984: 24). However, according to González Ollé (1995: 54) books in the regional languages began to be reauthorized already around 1950.

\(^8\) Although every child has the right to receive its first instruction in its “habitual” language, be it Catalan or Spanish (cf. art. 21).
12 local speakers. These recordings (around 20 min/speaker totaling 6 hours) now form part of the FEC corpus.

From this subcorpus, the transcriptions of the interviews and the text as well as the sociodemographic information were included for analysis. The text of the reading task consists of 391 words and contains the sentence *algo lo pica* ‘something pricks him’. Since the reading task aims to detect phonological phenomena, it is actually not suitable for morphosyntactic analyses. However, the recordings contain several deviations from the text. In cases where speakers deviate from the text, producing *algo le pica* instead of *algo lo pica*, the analysis of the reading task might indeed give some insights into the speakers’ use of the pronouns.

The second corpus is the *Corpus oral de profesionales de la lengua castellana en Barcelona* (Sinner 2001) which consists of interviews with 11 speakers (100-120min/speaker, totalling 18h 15min) from Barcelona. Among the interviews, Sinner (2001) also conducted conjugation tasks, acceptability judgement tasks, denomination tasks (color denomination tasks; tasks with pictures), which were also recorded and added to the corpus. In addition, a sociodemographic questionnaire was used, which also included questions about the informants’ linguistic biography and their linguistic dominance.

### 5.3 Sociodemographic and sociolinguistic data

The 12 speakers of the FEC corpus are distributed evenly with respect to gender and belong to three different age groups (see Table 3). All of the speakers were raised in the province of Barcelona (Barcelona, Sant Cugat, Sant Joan Despí, Terrassa, or Vilanova i la Geltrú) and have not lived elsewhere for a period longer than one year. Some of the speakers consider themselves balanced bilinguals and some are Catalan-dominant bilinguals (see Table 3). Importantly, none of the informants of the FEC corpus has a special formation in linguistics.

In contrast, the informants of the Sinner corpus are professionals, such as translators, proofreaders, or language teachers, who are either Spanish-dominant or balanced\(^9\) (see Table 3).

---

\(^9\) Data from one Catalan-dominant speaker was excluded from the analysis.
**5.4 Analysis**

In a first step, we collected all the sentences containing a pronominalized third person object. Then, we coded for the dependent variable, the form of the pronoun (*le/ lo/ la/ les/ los/ las*), and the following independent variables: syntactic function (indirect complement/ accusative, direct complement/ dative, attribute), number of the referent (singular/plural), gender of the referent (feminine, masculine, and neuter), animacy (animate, inanimate), and verb features (types a-f, see Section 2.2).

### 6. RESULTS

#### 6.1 Reading task

In the FEC Barcelona subcorpus, 4 out of 12 informants read *algo le pica* instead of *algo lo pica*. This is in so far surprising, as it is the highest leísmo rate compared to other FEC subcorpora (see Table 4).

<table>
<thead>
<tr>
<th>City</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seville</td>
<td>1/12</td>
</tr>
<tr>
<td>Bilbao</td>
<td>3/12</td>
</tr>
<tr>
<td>Madrid</td>
<td>3/12</td>
</tr>
<tr>
<td>Barcelona</td>
<td>4/12</td>
</tr>
<tr>
<td>Bogotá</td>
<td>0/12</td>
</tr>
</tbody>
</table>

**Table 4.** Occurrences of *algo le pica* in the reading task

Even in Madrid and Bilbao, which are prototypical leísmo varieties, we find fewer substitutions. Therefore, these results suggest that leísmo is not completely absent in Barcelona.
6.2 Semi-spontaneous speech
6.2.1 Analysis of the general distribution of pronouns

From both corpora, we extracted a total of 1531 (1195\textsuperscript{10} in the Sinner corpus and 222 in the FEC In cases where corpus) pronominalized objects, which are distributed as follows:

In both corpora, the most frequently used pronoun is \textit{lo}, which appears 737 times (61.6\%) in the Sinner (2001) corpus and 138 times (62.2\%) in the FEC corpus. This is not surprising since it does not only refer to masculine and neuter direct objects, but also to attributes. The second most frequently used pronouns are \textit{la} in the Sinner corpus (19.8\%) and \textit{le} (18.5\%) in the FEC corpus. The relatively high number of

\textsuperscript{10} Five sentences with object pronouns were excluded from the corpus, since they were incomplete and it was therefore impossible to determine their syntactic function.
feminine direct object pronouns in the Sinner corpus might be due to the fact that it also contains the recordings of an acceptability judgement task. In the course of that task, the participants first had to read a sentence and, before judging it, say whether they understood it. As a consequence, we find a relatively high number of (127) instances of the answer (no) la entiendo ‘I (don’t) understand it’. The percentage of le is 11.9% in the Sinner corpus. In both corpora, we find few cases of plural pronouns, namely 7% in the Sinner corpus and 10.8% in the FEC corpus.

6.2.2 Analysis of the distribution of pronouns with respect to syntactic function and gender

The analysis of the data regarding the first independent variable, i.e. the syntactic function of the pronoun, yields the following distributions:

<table>
<thead>
<tr>
<th>Direct complement</th>
<th>Indirect complement</th>
<th>Attribute</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>lo</td>
<td>735 (99.7%)</td>
<td>0</td>
<td>2 (0.3%)</td>
</tr>
<tr>
<td>la</td>
<td>237 (100%)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>le</td>
<td>14 (9.9%)</td>
<td>128 (90.1%)</td>
<td>0</td>
</tr>
<tr>
<td>los</td>
<td>12 (100%)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>las</td>
<td>31 (100%)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>les</td>
<td>8 (22.2%)</td>
<td>28 (77.8%)</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1037 (86.8%)</strong></td>
<td><strong>156 (13.1%)</strong></td>
<td><strong>2 (0.2%)</strong></td>
</tr>
</tbody>
</table>

Table 5. Distribution of pronouns regarding syntactic function in the Sinner (2001) corpus

<table>
<thead>
<tr>
<th>Direct complement</th>
<th>Indirect complement</th>
<th>Attribute</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>lo</td>
<td>137 (99.3%)</td>
<td>0</td>
<td>1 (0.7%)</td>
</tr>
<tr>
<td>la</td>
<td>18 (100%)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>le</td>
<td>12 (29.3%)</td>
<td>29 (70.7%)</td>
<td>0</td>
</tr>
<tr>
<td>los</td>
<td>13 (100%)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>las</td>
<td>3 (100%)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>les</td>
<td>4 (44.4%)</td>
<td>5 (55.6%)</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>187 (84.2%)</strong></td>
<td><strong>34 (15.3%)</strong></td>
<td><strong>1 (0.5%)</strong></td>
</tr>
</tbody>
</table>

Table 6. Distribution of pronouns regarding syntactic function in the FEC Barcelona corpus

As can be seen in Tables 5 and 6, both corpora predominantly contain pronouns that function as direct objects (87.9% in the Sinner corpus and 84.7% in the FEC corpus). The percentage of indirect object pronouns are 12.0% in the Sinner corpus.
and 15.3% in the FEC Barcelona corpus, whereas there are almost no pronouns substituting attributes. If we consider the corresponding forms of the pronouns, we see that lo, la, los and las only appear as direct objects\(^{11}\), whereas le and les appear both as indirect objects (Sinner corpus: 87.6%; FEC Barcelona corpus: 68.0%) and as direct objects (Sinner corpus: 12.4%; FEC corpus: 32.0%). The latter case includes all the occurrences of leísmo. Hence, in the Sinner corpus, 22 out of 1037 (2.1%) third person direct object pronouns appear as le(s) and thus constitute instances of leísmo. In the FEC Barcelona corpus, the leísmo rate is higher, i.e. 8.6% (16/187). Nevertheless, both rates seem to be rather low. However, if we consider that leísmo occurs mainly with masculine referents and that it is not possible with neuter referents, it becomes clear the latter case has to be excluded in the computation of the leísmo rate. The following tables show the distribution of third person direct object pronouns with respect to the gender of the referents:

<table>
<thead>
<tr>
<th></th>
<th>lo</th>
<th>la</th>
<th>le</th>
<th>los</th>
<th>las</th>
<th>les</th>
</tr>
</thead>
<tbody>
<tr>
<td>m</td>
<td>141</td>
<td>0</td>
<td>8</td>
<td>12</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>f</td>
<td>2</td>
<td>228</td>
<td>3</td>
<td>0</td>
<td>27</td>
<td>1</td>
</tr>
<tr>
<td>n</td>
<td>581</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>unclear(^{12})</td>
<td>11</td>
<td>9</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>735</td>
<td>237</td>
<td>14</td>
<td>12</td>
<td>31</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 7. Distribution of direct object pronouns regarding the gender (feminine (f), masculine (m), neuter (n)) of the referent in the Sinner corpus

<table>
<thead>
<tr>
<th></th>
<th>lo</th>
<th>la</th>
<th>le</th>
<th>los</th>
<th>las</th>
<th>les</th>
</tr>
</thead>
<tbody>
<tr>
<td>m</td>
<td>58</td>
<td>0</td>
<td>9</td>
<td>11</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>f</td>
<td>1</td>
<td>16</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>n</td>
<td>77</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>unclear</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>137</td>
<td>18</td>
<td>12</td>
<td>13</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 8. Distribution of direct object pronouns regarding the gender (feminine (f), masculine (m), neuter (n)) of the referent in the FEC corpus

The tables show that in both corpora la refers exclusively to feminine referents, whereas le/les appear with both masculine (Sinner corpus: 13/22 (59%); FEC corpus 12/26 (46.2%)) and feminine (Sinner corpus: 4/22 (18.2%); FEC corpus: 3/16 (18.8%))

\(^{11}\) In other words, there is no loísmo (the employment of lo instead of le) or laísmo (the employment of la instead of le) in the data.

\(^{12}\) In some cases, the referent or some of its features could not be clearly identified.
Leísmo or fake leísmo? New insights into Catalan Contact Spanish from the FEC Corpus

Leísmo or fake leísmo? New insights into Catalan Contact Spanish from the FEC Corpus

referrers. Crucially, in both corpora, lo refers mostly to neuter referents (79% in the Sinner corpus and 56.2% in the FEC corpus), where leísmo is not possible. Therefore, if the neuter pronouns are excluded, the leísmo rates are as follows: 22/456 (4.8%) in the Sinner corpus and 16/108 (14.8%) in the FEC corpus:

<table>
<thead>
<tr>
<th></th>
<th>le(s)</th>
<th>lo(s)/la(s)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sinner (2001)</td>
<td>16</td>
<td>92</td>
<td>108</td>
</tr>
<tr>
<td>FEC Barcelona</td>
<td>22</td>
<td>434</td>
<td>456</td>
</tr>
<tr>
<td>Total</td>
<td>38</td>
<td>526</td>
<td>564</td>
</tr>
</tbody>
</table>

Table 9. Distribution of pronouns in the function as direct objects; neuter referents are excluded

A Chi-squared test of homogeneity on this dataset with the null hypothesis (which means that the proportion of leísmo is the same in Sinner (2001) and FEC Barcelona) yields a X-squared value of 12.325 with p < 0.001. Thus, the null hypothesis has to be rejected, given that there is a significant difference regarding the proportion of leísmo between the two corpora.

6.2.3 Analysis of the features of the referents and the verbs

Regarding the second question, which concerns the different types of leísmo, we first consider the animacy of the referent as an additional variable (see Tables 10 and 11). In both corpora, all of the [+animate] referents are also [+human].

<table>
<thead>
<tr>
<th></th>
<th>le</th>
<th>les</th>
</tr>
</thead>
<tbody>
<tr>
<td>m</td>
<td>f</td>
<td>unclear</td>
</tr>
<tr>
<td>animate</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>inanimate</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>unclear</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 10. Distribution of leísmo regarding gender and animacy in the Sinner corpus

<table>
<thead>
<tr>
<th></th>
<th>le</th>
<th>les</th>
</tr>
</thead>
<tbody>
<tr>
<td>m</td>
<td>f</td>
<td>unclear</td>
</tr>
<tr>
<td>animate</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>inanimate</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>unclear</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 11. Distribution of leísmo regarding gender and animacy in the FEC corpus
In the Sinner corpus, only 3 out of 22 occurrences of *leísmo* are instances of *leísmo correcto* (masculine human referents in singular), which is considered to be the most widespread *leísmo* type. Yet, this relatively low number is not surprising if we consider the fact that both corpora contain very few animate pronominalized direct objects in general, which might be due to the content of the interviews (which, for example, do not contain many anecdotes involving other people).

(3) a) […] si le lees cuentos a tu hijo, pues también él coge el gusto por el cuento. ¿No? y *le* enseñas, juegos con él.
   ‘[…] if you read stories to your son, he also gets to like the story. Right? And you teach him and play with him.’

b) […] a [nombre de un filósofo alemán], cuando, *le* nombraron doctor de honoris causa, en la/ en la universidad [nombre de una universidad alemana].
   ‘[…] to [name of a German philosopher], when they made him an honorary doctor, at the University [name of a German university].’

All of the other cases are instances of *leísmo incorrecto*. Among these, we find *leísmo* with masculine but inanimate referents:

(4) Al batido de cacao, *le* llamamos *cacaolat*.
   ‘The chocolate milkshake, we call it *cacaolat*.’

Examples like the one in (4) constitute the majority of all occurrences in the Sinner corpus (5/22). We also find, although with a lower frequency, *leísmo* with plural forms, mostly with human referents (5a) but also with inanimate referents (5b).

(5) a) A: Y qué crees que se podría hacer para que *los alumnos*
tuvieran/
   B06: Pues, *hacerles* escribir y leer.
   ‘A: And what do you think one could do to achieve that students have/
   B06: Well, make them write and read.’

b) pienso que, sobre *los pronombres* a veces, que están bien, pero
   siempre hace falta marcarles un poco más.
   ‘I think that, sometimes about the pronouns, which are fine, but it’s always necessary
to mark them a little bit more.’

Although it is considered the most infrequent type, we find a few cases of *leísmo* with feminine referents.
Leísmo or fake leísmo? New insights into Catalan Contact Spanish from the FEC Corpus

(6) a) Le pondremos a la abuela en el balancín.
   ‘We’ll put the grandmother on the rocking chair.’

b) Vas a ayudar, a esa persona, que es aquí la productora, entonces, vas a ayudarle.
   ‘You are going to help this person, who is the producer here, so, you are going to help her.’

In the FEC Barcelona corpus, we find a similar distribution. Yet, there are even less instances of leísmo correcto (1/16):

(7) O sea, a un gallego se le reconoce mucho.
   ‘That is, you can easily recognize a Galician.’

Crucially, half of all instances of leísmo (8/16) appear with masculine inanimate referents (8), whereas there are very few instances of plural leísmo (9a) and leísmo with feminine referents (9b)

(8) a) Grandes exportadores de champán, aquí le llamamos cava.
   ‘Big champagne exporters, here we call it cava.

b) A este pastel se le llama la mona.
   ‘This cake is called la mona’

(9) a) A sudamericanos también cuesta entenderles.
   ‘It is also difficult to understand South Americans.’

b) ... a asentar una población, por lo tanto le llamaron villas nuevas.
   ‘... to found a village, therefore, they called it villas nuevas.’

The last independent variables to be analyzed are the verbs and verb constructions introduced by Fernández-Ordóñez (1999). Table 12 contains all instances of leísmo in both corpora of the construction types a)-f) (see Section 2.2).

<table>
<thead>
<tr>
<th></th>
<th>Sinner</th>
<th>FEC</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Experiencer verbs</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>b) Infinitive constructions</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>c) Verbs that can omit their direct objects</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>d) Verbs whose case government has been reinterpreted</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>e) Verbs with objects that require predicative complements</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>f) Constructions with impersonal <em>se</em></td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Others</td>
<td>5</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 12. Occurrences of leísmo regarding verb types
It can be seen that the majority of the occurrences of *leísmo* belongs to type e), i.e. verbs whose objects require predicative complements. This group comprises all the cases in which *leísmo* appears with the verbs *llamar* and *nombrar* (e.g. examples 3b; 4; 8a, b; 9b). In the FEC corpus, 12 out of the 16 occurrences of *leísmo* are of this type, which corresponds to a percentage of 75%, whereas in the Sinner corpus it is a percentage of 54.55% (12/22). The Sinner corpus additionally contains 3 cases of *leísmo* with the verb *ayudar* ‘to help’, i.e. instances of type d) (see example 6b). Finally, each corpus contains one case of *leísmo* with an infinitive construction: a construction with *hacer* (see example 5a) in the Sinner corpus and one with *obligar* in the FEC corpus:

(10) En los bares, en las terrazas, si los bares, si tienen una terraza les obligan a cerrar la terraza.
    ‘In the bars, on the pavements, if the bars, if they have a pavement café, they oblige them to close that outdoor café.’

Of the 38 cases of *leísmo* in the corpora, 30 appear with verbs or verb constructions of the types a)-f). According to Fernández-Ordóñez (1999), these are not instances of true *leísmo* but only cases of fake *leísmo*. Therefore, if we follow this approach, only the remaining 8, subsumed under the category *Others* are real instances of *leísmo* (see examples 5b, 6a, 7, 9 and 11).

(11) Pero no había manera de retenerle.
    ‘But there was no way to hold him back.’

7. DISCUSSION

While some previous studies and the results of the reading task suggest that *leísmo* exists in Catalan Contact Spanish, we see that there are very few cases of *leísmo* in both corpora of spontaneous speech (4.6% in the Sinner corpus and 12.1% in the FEC corpus). This relatively low number might be traced back to the fact that both corpora contain very few animate pronominalized direct objects (which commonly trigger *leísmo*) which is, in turn, due to the content of the interviews that the corpora are built of. The interviews of the Sinner focus on the features of Spanish in Catalonia and also contain experimental data such as an acceptability judgement task. As described in Section 5.2, the FEC corpus consists of interviews on cultural issues and linguistic awareness. Thus, none of the corpora includes descriptions of interactions between
people, which would involve more animate pronominalized objects. Therefore, it is also not surprising that leísmo is not restricted to leísmo correcto (human masculine referents in singular) in our data, although it is considered the most common type of leísmo. However, it mainly occurs with other combinations of features of the referents. In fact, leísmo most frequently occurs with masculine, inanimate referents with the verb llamar. Considering the involved verbs, it becomes clear that the majority (about 80%) of all instances of leísmo are in fact cases of fake leísmo. Therefore, regarding the second question, our data suggest that Catalan Contact Spanish might be rather a fake leísmo variety. In this aspect, our results are in line with Roselló Verdeguer’s (2017) results for Valencian Spanish. With respect to the third question, although leísmo occurs infrequently in both corpora, we see that the relative number of leísmo is significantly higher in the FEC corpus. This is in so far surprising, as most of the informants of the FEC corpus are Catalan-dominant. Since leísmo is not existent in Catalan, we would expect a higher number of case-distinguishing use of third person pronouns than in the Sinner corpus. A possible explanation for that could be found in the extralinguistic information about the informants. While none of the informants of the FEC corpus has a special formation in linguistics, the informants of the Sinner corpus are professionals, including translators, proofreaders, and language teachers. Consequently, the more “standard” use of third person pronouns could be due to the high normative awareness of the informants. In his analysis of the corpus, Sinner states that the informants’ acceptability judgements as well as their use of some linguistic elements are “related to what they know or believe to know about the analyzed phenomenon, relating to the norms of usage” (Sinner 2004: 566). Yet, more investigation, such as acceptability judgements of the informants on this specific phenomenon, would be needed to test this hypothesis.

8. CONCLUSION

The present study has sought to investigate the phenomenon of leísmo in Catalan Contact Spanish. Therefore, we analyzed a reading task and oral interviews taken from two corpora: the Sinner corpus and the FEC corpus. While the results of the reading task suggest that leísmo is not absent in the Spanish of Barcelona, there are only a few cases of leísmo in spontaneous speech in both corpora. A closer analysis of the pronoun’s referents and the involved verb constructions showed one the one hand that leísmo incorrecto appears more frequently in our data than leísmo correcto. However, most of the instances of leísmo are, in fact, examples of fake leísmo. This means that Catalan Contact Spanish might rather be regarded as a fake leísmo variety.
Regarding the linguistic dominance of the informants, we can say that for our data it is not true that Catalan-dominant bilinguals produce fewer cases of *leísmo*, since the *leísmo* rate is higher in the FEC corpus.

A desideratum for future work on *leísmo* is to include the verb constructions introduced by Fernández-Ordóñez (1999) into the study of *leísmo* in varieties traditionally regarded as *leísmo* varieties. Since very few studies contain this variable, the rate of fake *leísmo* in these varieties is unknown. Thus, the inclusion of this factor is necessary in order to draw a relevant comparison with Catalan Contact Spanish.

Additionally, future work on *leísmo* in Catalan Contact Spanish could benefit from including experimental data into the analysis. Since the results of previous experimental work, such as Blas Arroyo’s (1994) study, differ from those presented in the present paper, it may be fruitful to conduct experimental studies that additionally include verb constructions as a variable. It also could be worthwhile to incorporate more speaker-dependent variables such as age, sex, and education to gain insights into extralinguistic factors that might trigger *leísmo*.

Finally, it may be beneficial to analyze the Menorcan FEC subcorpus with regard to *leísmo*, as the language contact situation is different from that of the province of Barcelona. Since Catalan is much more present in rural areas of the Balearic Islands, a comparison with the Menorcan subcorpus could provide more insights on the effects of linguistic dominance on *leísmo*.

Although there are still some open questions regarding *leísmo* in Catalan Contact Spanish, this paper has tried to contribute not only to the investigation of the phenomenon of *leísmo* and its variation but also to the recently growing body of research on the features of Catalan Contact Spanish.

**BIBLIOGRAPHY**


Leísmo or fake leísmo? New insights into Catalan Contact Spanish from the FEC Corpus

crossroads with other languages. Amsterdam/Philadelphia: Benjamins, pp. 87-106. https://doi.org/10.1075/sibil.37.07boi


Leísmo or fake leísmo? New insights into Catalan Contact Spanish from the FEC Corpus


