

Chapter 12

Differential Object Marking in Heritage and Homeland Italian

Margherita Di Salvo, Università degli Studi di Napoli Federico II

Naomi Nagy, University of Toronto

Abstract

We examine variable patterns of use of differential object marking (DOM) in conversational Italian recorded in Toronto, Canada, and Calabria, Italy. An exhaustive sample of 366 direct objects, produced by Homeland and three generations of Heritage speakers, shows retention of the DOM system. Successive generations have lower rates of DOM, but this is because they don't produce enough tokens of certain syntactic and semantic types (e.g., left-dislocated or indefinite pronouns). Thus, they have less opportunity to use DOM: token distributions account for their lower rates. In contexts with sufficient tokens, significant contrasts emerge, indicating that all generations retain the conditioning of relevant factors (Definiteness, Referent of Object, Verb Type, Dislocation). No effects of social network or linguistic practices emerged.

Keywords comparative variationist sociolinguistics, Calabria, Differential Object Marking, heritage language, Italian, morphosyntactic variation, Toronto.

Introduction

This paper examines differential object marking (DOM) in a corpus of spoken Italian gathered in Toronto¹ and in Calabria, Italy for the Heritage Languages Variation and Change (HLVC) project (Nagy 2009, 2011, 2015). DOM is used in many languages around the world (cf. Bossong 1991). Among these are many European languages where, per Sornicola (1998: 66), DOM is discontinuous in space and, in some language varieties, still incipient (Nocentini 1985: 303). The following examples from Mardale (2009) show DOM in Romanian, Spanish, and Sardinian. In these languages the object noun phrase (NP), despite sharing semantic-pragmatic properties, is introduced by different prepositions, *pe* in Romanian but *a* in Spanish and Sardinian:

(1) *l-am întâlnit *(pe) Ion (Romanian)*

him had-met DOM John

‘I met John.’

(2) *vi *(a) Juan (Spanish)*

saw DOM John

‘I-saw John.’

(3) *an furatu *(a) Ercole (Sardinian)*

have stolen DOM Hercules

‘They have stolen Hercules’

Marking of object NPs that are characterized by animacy, definiteness and particular syntactic structures (see §2), is documented in Ibero-Romance languages; in Rhaeto-Romanic

¹ On multilingualism in Canada see Rehner, Mougeon and Mougeon (this volume).

languages; in Spanish; in the Bern, Fribourg, Carcassonne and Narbonne dialects of French; in some varieties of Corsican and in Italian. Relevant to this study, DOM is found in Calabrian varieties and in the regional Italian spoken in Calabria. Examples (4)-(6) are from the HLVC corpus of Calabrese Italian:²

- (4) Ø *aiutavano a noi*
 they help.3pl DOM us

‘They used to help us.’ (I1F73A, 13:03)

- (5) *facevamo, "vedi se Lo puoi convincere tu A Nicola."*
 make.2pl see.2sg if Pro. can convince.INF you DOM Nicola

‘We were like, “See if you can convince Nicola.” (I1F61A, 36:50)

- (6) *e penso che Aiuta a i bambini*
 and think.1sg that help.3sg DOM the children

‘And I think that he helps children.’ (I2F44, 09:57)

We must consider the possibility that DOM-marking works differently in Calabrese Italian vs. the Calabrese dialect, and that speakers could differ in how and if they combine these distinct varieties during the recordings we analyze. However, we are aware of no previous studies that distinguish DOM patterns in these varieties. The native-speaking author of this paper has the impression, from listening to the HLVC recordings, that DOM is the same in both varieties, although it differs from Standard Italian, where DOM is prescriptively absent. Guardiano (2010: 102) reaches a similar conclusion in her study of Regional Sicilian Italian vs. Sicilian dialects.

² Except where otherwise noted, all remaining examples are from the HLVC Corpus (Nagy, 2009). Speaker codes identify the language, generation, sex and age of the speaker.

DOM has been the topic of synchronic and diachronic studies in Italian, but not previously subjected to variationist analysis. To formalize slightly, DO-marking is expected when the direct object is specific and human, whether or not it is definite. No DO-marking preposition is expected if the direct object is non-specific or inanimate (Irizarri van Suchtelen 2016: 99, for Spanish). Thus, there are some unclear areas, such as animals (animate but not human), as well as certain verbs which reject DOM (Irizarri van Suchtelen 2016: 100). Prescriptively, the prepositional object is mandatory with definite nominals such as 1st person pronouns, as in (4) (Loporcaro 2009: 131) or proper nouns as in (5), while is optional with +human common nouns as in (6) (Guardiano 2010).

We hypothesized that the above-mentioned parameters (animacy, definiteness of objects and syntactic structure of the sentence) will therefore affect DOM usage in the Calabrian Italian spoken as a heritage language in Toronto. We also hypothesized that DOM could vary according to external factors, such as the speaker's generation, sex and ethnic orientation. These factors help us understand how variation and change operate in heritage language situations, where it has been proposed that what is sometimes described as attrition may in fact be faithful replication by later generations of the input they receive from earlier generations (cf. Aalberse, Backus and Muysken 2019). This is investigated by comparing two groups of speakers who were born and raised in Calabria, but one of which emigrated to Canada, and two successive generations born in Canada.

The Status of DOM in Romance Languages

The Italian ecology is characterized by multilingualism consisting of Italian and multiple dialects (D'Agostino 2012, De Blasi 2014, Palermo 2015). It has been described in terms of *dilalia*,

meaning that Italian and its dialects are used in the same communicative domains. It is important to note that here the term ‘dialect’ refers to the many varieties of Italian which come from spoken Latin, and not to diatopic varieties of Standard Italian.

In the continuum between Standard Italian and dialects, DOM emerges variably across Italy: there are sporadic traces of DOM in varieties in Trieste, Genoa, on the island of Elba and in Sardinia (Pittau 1972, Iemmolo 2009, Boeddu 2017). Relevant to this study, its use is more frequent in all the varieties spoken in Central and Southern Italy (Rohlf 1966, Loporcaro 2009, Fiorentino 2003a, 2003b, Maiden and Parry 1997) than in the north. Traces of DOM are also found in colloquial Italian spoken in northern regions though it is subject to diaphasic variation: there it appears only in oral, colloquial varieties and informal written Italian (Berretta 1989: 224). For spoken Italian see also Cortelazzo (1972), Berretta (1989), Telmon (1993) and Berruto (2006).

The use of DOM is therefore non-standard, appearing only in colloquial and informal Italian and Southern dialects. There are no studies showing whether DOM is perceived as non-standard or popular. However, it seems to be stigmatized as it is used to mark the discourse of a fool in the song *Mio cugino* ‘My Cousin’ by Elio e Le Storie Tese,

In spoken Italian and Italian dialects and in the other languages which differentially mark direct objects, the presence or absence of the prepositional marker is related to three parameters:

- a. the semantic, syntactic and pragmatic characteristics of the object NP
- b. the characteristics of the verb
- c. the order of the constituents

With regards to the characteristics of the object NP, earlier studies of several languages claimed that animacy, definiteness, and topicality influence DOM. From these studies, it

emerged that grammaticality is determined by the interaction of the [+definite] and [+human] features which trigger DOM and by factors related to the intrinsic semantic content of the nominal element since [-animate], [+mass], and [+abstract] features proscribe the possibility of using this marker.

Aissen (2003) and Laca (2006) built an implicational model that takes into account the syntactic and semantic-pragmatic categories of the Object (O), quoted in (7). In this model, DOM is compulsory in the categories at the top of the scale, and progressively less obligatory further down the scale. At the bottom, DOM is not possible.

(7) *Implicational scale of the contexts favouring DOM (Mardale 2008: 451-452)*

human pronouns > human proper nouns / animate pronouns

human definite NPs / animate proper nouns / inanimate pronouns

human indefinite specific NPs > animate definite NPs > inanimate proper nouns

human non-specific NPs > animate indefinite specific NPs > inanimate definite NPs

animate non-specific NPs > inanimate indefinite specific NPs > inanimate non-specific

NPs.

For the Italo-Romance varieties, the scale was further refined in studies carried out by Guardiano (2000, 2010), starting with Sicilian varieties. Guardiano (2010: 12) developed a scale of ten types of nominal arguments distributed according to the likelihood of object marking (with a preposition) on items in each category, which we reproduce in Table 1. This closely follows animacy hierarchies proposed in Comrie (1989) and Silverstein (1976) and has been used to describe Italian varieties similar to what we investigate.

Table 1. Hierarchy of types of objects proposed in Guardiano (2010)

1.	first and second person personal pronouns
2.	third person singular pronouns with human referent
3.	proper nouns (person or animal)
4.	kinship nouns preceded by an expression of possession
5.	third person pronouns with non-human animate referent
6.	common nouns of people
7.	common nouns of animals
8.	common nouns of objects
9.	mass nouns
10.	abstract nouns

Based on other Romance languages such as Romanian, Spanish (Mardale 2008), and Sardinian (Boeddu 2017), DOM is grammatical and compulsory in categories 1 to 4; it is grammatical but optional in categories 5 to 7; and it is ungrammatical in categories 8 to 10. Diachronically, it has spread farther down the hierarchy in some Spanish varieties than others (Aissen 2003: 463). Descriptions of Italo-Romance varieties indicate that a preposition is compulsory with 1st, 2nd, and 3rd person singular personal pronouns (Loporcaro 2009: 131). However, recent studies on the accessibility of categories 8 to 10 to preposition marking highlighted a rather unstable scenario. For example, with regards to Spanish, Tippetts (2010: 205) writes:

inanimate [direct objects] that had the other features associated with highly individuated objects, especially definiteness and specificity, were marked far more frequently and constituted more than half of all marked inanimates.

In the Calabrian varieties, a case of DOM with not only an inanimate but also indefinite object is reported by Marchese (2016: x) who nonetheless attributes the presence of the prepositional marker to the preceding context:

an isolated case of prepositional accusative with inanimate was recorded in a Calabrian native speaker who had been living in Rome for many years. This could be influenced by the presence of an inanimate antecedent: *'ntruzzava quandu unu 'ntruzzava na cosa [...] na machina, na cosa che si ntruzzava [...] ntruzzài a cchià perzuna, pe' ddire, a nna machina.*³

Similar counterexamples with highly specific, definite and topical objects are also found in South American Spanish (Lopez 2012). With non-human Os, therefore, it is definiteness which determines the presence of the preposition. In all the studies mentioned here, this type of marking is extremely rare: one single occurrence in Marchese's corpus, and five occurrences out of 699 possible ones in the Spanish spoken in Buenos Aires and Madrid (Balasch 2011: 119). These rare outcomes show how, alongside animacy, the parameters of definiteness and topicality play a role. So, while in some nominal categories (personal pronouns) DOM can be easily seen as compulsory, in others there is alternation which does not categorically relate to general parameters. Balasch (2011) pointed this out and employed statistical analysis to illustrate the multiple factors at play.

NP-features are not the only crucial factors, a role is also played by the type of verb, something which has not been incorporated in many studies of Spanish DOM of a quantitative nature (but cf. von Heusinger's 2008 corpus study and Irizarri van Suchtelen's 2016 dissertation, which show that verbs which more frequently take human direct objects are more frequently DO-marked than other verbs, even when they don't have an animate object). Diachronic studies conducted on multiple languages demonstrated the importance of the diachrony of the verb

³ Translation: 'ntruzzava' is when one hit something [...] a car, something that collided [...] I bumped into that person, for example, a car.

(henceforth V) since DOM appears most often with those verbs that showed oscillation between the dative and the accusative construction as far back as Late Latin (for Italian see Sornicola 1997, 1998 and Fiorentino 2003b; for Romanian see Hill 2015). Regarding Italian, Berretta (1989) also noted the occurrence of DOM with psychological verbs and constructs with the causative *fare* ('do'/'make'). Starting with Pottier (1968), many studies showed that the presence of the preposition is due to the properties of action and aspect of the verb. Specifically, DOM is triggered by less prototypical objects in transitive structures (mainly telic verbs, cf. Iemmolo 2009) as per Hopper and Thompson's (1980) hierarchy of transitivity.

Finally, the order of the constituents is also crucial. According to Renzi (1988) DOM is favoured under three conditions (which do not make it compulsory) (Berretta 1989):

- a. when the object consists of a deictic pronoun as in: *chiamo a lui ogni sera* ('I call TO him every evening')
- b. when it is left-dislocated and, therefore, separated from the rest of the verb phrase as in: *a lui ho incontrato ieri* ('TO him I met yesterday')
- c. when it is referred to by an unstressed pronoun in the body of the phrase as in *a me mi ha chiamato mamma* ('TO me my mum called').

Based on qualitative observations, Leonetti (2008) claims that DOM preferably occurs in structures with a marked order of the constituents, and in particular with left dislocation.

Regional variation has been studied mainly in polycentric languages such as Spanish, where we find conflicting reports. There are those who do not report any variation, such as, for example, Balasch (2011), who after comparing the *Corpus de Mérida/Venezuela* and the *Corpus del habla culta de Madrid/Spain* concluded that "contextual conditioning is identical in Mérida and

Madrid; though the overall overt *a* rate is much higher in the latter”, while Tippetts (2010) proposes dialect variation as the way forward in future research.

Some studies examine Spanish in contact situations in countries of immigration whose language does not use DOM (usually English). These could contribute to an understanding of the ways in which heritage speakers recognize, lose or acquire syntactic, semantic and pragmatic categories. The few *heritage language* studies carried out on this subject showed that age of acquisition is crucial. Both Rodríguez-Mondoñedo (2008) and Ticio and Avram (2015) showed that children learn the rules which regulate DOM very early and that the critical age for this structure is three years. However, experimental studies comparing children living in homeland vs. heritage language contexts show important differences in rates of DOM production, with an average of 30% wrongful omissions of the prepositional marker found in the heritage group (Montrul and Sanchez-Walker 2013; several others cited in Irizarri van Suchtelen 2016: 102). Divergence between homeland and heritage varieties have been attributed to incomplete acquisition (cf. Montrul and Bowles 2009:381); as properties of a bilingual variety (Di Venanzio et al. 2012); or as due to contact with English, which lacks DOM (cf. Montrul and Bowles 2009:368; Montrul and Sánchez-Walker 2013).

Earlier studies of first-generation adult immigrants highlighted a considerable tendency to omit the prepositional marker (up to 50% of the possible contexts) by Spanish speakers living in the US, leading to the conclusion that “even advanced heritage speakers are very inaccurate with DOM” (Montrul et al. 2015: 576). In contrast, most studies report high accuracy for HL speakers in the zero-marking contexts (Irizarri van Suchtelen 2016: 103).

Comparative studies on multiple languages within the same immigration context are also rare, with the exception of Montrul et al. (2015) who adopted an experimental approach. These

authors report acceptability judgments to analyze the structure in Spanish, Hindi, and Romanian as spoken in the homelands and in the US by 1st and 2nd generation speakers of different ages, level of bilingualism and gender, and found that in Spanish the erosion of the structure is more advanced than in Romanian and Hindi, despite the greater visibility of the Spanish language in American public life. They propose that the structural properties of the DO-markers together with the syntax of definiteness and specificity in each language account for the degree of DOM erosion in each language (more than the external factors). Because DOM is affected in the judgments of adult Mexican Spanish immigrants, whereas Hindi and Romanian immigrants are not affected, the degree of erosion of DOM observed in the Spanish heritage speakers is more extensive (in terms of number of individual speakers) than the extent observed in Hindi and Romanian. Therefore, for SOME linguistic structures and for SOME populations, incomplete acquisition in heritage speakers is also related to qualitatively different input provided by the parental generation (Montrul et al. 2015: 604).

Italian has been investigated much less with respect to DOM. From the very few studies carried out on Italian spoken abroad we observe a completely different scenario: Di Salvo (2017) compared 50 native speakers of Italo-Romance Southern dialects who immigrated to Bedford, England in their adulthood and a control group of speakers who remained in Italy. The objective was to verify if, and up to what point, the limited exposure to Italian in England could influence DOM erosion. The study showed that, in this immigrant variety, DOM is quantitatively and qualitatively present in the same way as in Italy. Subsequent studies of the Italian spoken in England confirmed this tendency, showing the extension of the prepositional marker to contexts where, according to descriptions of spoken homeland Italian, the preposition was not expected according to the features of the object (Di Salvo 2019). DOM examples from “Transnational

migrations: the case of the Italian communities in the UK” Corpus have been found even when O is:

(8) animate and indefinite:

Nonna qua ha Portato a na ragazza

grandma here have.3sg take.Past DOM a girl

‘Grandmother took a girl here’ (BED1M_AnB)

(9) inanimate and indefinite:

sul ie nu Avesse truat a niente

only I not have.Conditional find.Past DOM nothing

‘Only I wouldn’t have found anything’ (CAM1F_Sta)

(10) inanimate but definite:

Ø faccio a O *apple crumble*

(I) do.1sin DOM The *apple crumble*

‘I do the apple crumble’ (MON1F_F)

(8) and (9) were produced by first generation immigrants living in Bedford and (10) by a speaker who returned to Italy after a long time living in Bedford. They suggest that emigration and exposure to a language where DOM is not present, such as English, may lead to erosion of the structure but also to an extension to less canonical cases such as (8) or even to cases that are ungrammatical in Italian like (9) and (10). The current study allows us to understand how widespread this trend is and whether it is caused by a loss of awareness of the features of

definiteness and animacy, which influence the use of DOM, a hypothesis which is contradicted by canonical use in the remaining contexts, as highlighted in Di Salvo (2017), or are indicative of the contact between grammars and the innovation which such contact can produce.

Methods

Our analysis is situated in the comparative variationist framework. We compare patterns of variation (use of DOM vs. its omission, or, more explicitly *a* + O vs. O) in samples of spontaneous utterances from four groups of speakers: Homeland, Gen1 Heritage, Gen2 Heritage, and Gen3 Heritage. The Homeland speakers have always lived in Calabria, Italy, and were recorded in conversation with other Calabrese speakers in 2013.⁴ The Gen1 speakers were also born and raised in Calabria until at least age 18 but subsequently have lived for at least 20 years in Toronto. The Gen2 speakers were born in Toronto (or arrived before age 6), and their parents qualify as Gen1. The Gen3 speakers were all born in Toronto and their parents qualify as Gen2. These three generations of heritage speakers were recorded between 2009 and 2016 in Toronto. Each of the four groups has eight speakers.⁵

All data was collected and analyzed following the standard Labovian sociolinguistic interview protocol (cf. Labov 1984). All interactions were initiated and recorded in Calabrese Italian. Further methodological details for the [Heritage Language Variation and Change \(HLVC\)](#) Project are available in Nagy (2009, 2011, 2015).

⁴ We are grateful to the fieldworkers who worked hard to recruit speakers, interview them and transcribe their speech. They are listed at http://ngn.artsci.utoronto.ca/HLVC/3_2_active_ra.php and http://ngn.artsci.utoronto.ca/HLVC/3_3_former_ra.php. We thank SSHRC and the University of Toronto's Faculty of Arts and Science for funding to the second author (Standard Research Grant 410-2009-2330, Insight Grant 435-2016-1430).

⁵ The same methodology was used to examine Heritage Cantonese by Tse (this volume.)

From the conversational speech transcribed in ELAN (Wittenburg et al. 2006, which allows time-alignment to the recordings), all utterances with verbs that require a direct object, and whose object was animate, were selected.⁶ Each was coded for the binary dependent variable: whether the direct object was preceded by the preposition *a* or not. In subsequent tiers in ELAN (see Nagy and Meyerhoff 2015), each independent linguistic variable was coded, as in Table 2. This allows us to operationalize as probabilistic constraints the theoretical claims that certain contexts require/prefer/proscribe DOM more than others.

Table 2. Coding of variables with illustrative examples from the HLVC Corpus

Dependent Variable: Differential object marker					
Present:	<i>come</i>	<i>la</i>	<i>Conosca</i>	<i>a</i>	<i>Padre Amedeo?</i>
	how	him	Know	DOM	Father Amedeo?
	‘How do you know Father Amedeo?’ (I1F71A, 24:04)				
Absent:	Ø	<i>abbiamo</i>	<i>dovuto</i>	<i>chiamare</i>	Ø
	(we)	have	needed	to-call	DOM
	“We had to call the Father’ (I1M62A, 31:36)				
	<i>il</i>	Prete			
	the	priest			

Independent linguistic variables
Definiteness

⁶ As in many studies of minority and lesser-studied languages, constraints on resources make it difficult to collect and analyze token sets as large as those often encountered in studies of majority languages. We exhaustively extracted all relevant tokens from available recorded interviews that are, on average, one hour long.

Definite: Ø *abbiamo incontrato a queste due*
 (we) have.1pl met DOM these two
 ‘we met these two my friends’ (IXF14A, 09:9)
amiche
 friends

Indefinite: *lei aiuta a parecchie persone*
 she help.3sing DOM many people
 ‘she helps many people’ (I1M61A, 09:59)

Type of Object

1st and 2nd Ø *aiutavano a noi*
 person
 personal (they) help.Past DOM us
 pronouns: ‘(they) used to help us’ (I1F73A, 07:8)

3rd person Ø *so chiamato a loro*
 personal
 personal (I) am called DOM them
 pronouns: ‘I called them’ (I1F71A, 31:80)

Proper *come la conosca a Padre Amedeo?*
 nouns: how Pronoun know DOM Father Amedeo?
 (Object)
 ‘How do you know Father Amedeo?’ (I1F71A, 24:04)

Other Ø *non ho conosciuto a*
 pronouns: (I) not have.1sing recognize.Past DOM
 ‘I did not recognize anybody’ (I1F73A, 25:27)
nessuno
 nobody

Kinship preceded term + possessor: *mille occhi che guardano al*
 thousand eyes that take care DOM+the
 '1,000 eyes that take care of your son' (IXM35A, 06:43)
tuo figlio
 your son

Common human nouns: *interrogare allo Studente*
 Ask.INF DOM + the student'
 'to ask the student' (I1M60A, 37:09)

Type of Verb

Psych: Ø *non ho conosciuto a Nessuno*
 (I) not have.1sing recognize.Pas DOM Nobody
 'I did not recognize anybody' (I1F73A, 25:27)

Accusative/ Dative Alternating: Ø *aiuta ai Bambini*
 (he) help.Pres.3sing DOM + the Children
 'he helps children' (I2F44A, 59:7)

Telic: Ø *hanno ammazzato a essa*
 (they) have.Past.3pl killed DOM her
 'they killed her' (I1F71A, 26: 13)

Other: Ø *lasciava a Sonia con lei*
 (I left.2Past. DOM Sonia with her
) **1sin**
 'I left Sonia with her' (1F61A, 08: 02)
)

Position of object

Left Dislocated: *a sto F[...]* Ø *non*
 DOM this name I non

‘This *name*, I have never heard him’ (IXM63A,
21:11)

ho mai sentito

have.1st never heard

SVO: *non ho conosciuto a Nessuno*

not have.1sg recognize.PAST DOM Nobody

‘I did not recognize anybody’ (I1F73A, 25:27)

Additionally, the speaker's Generation and Sex was coded for each token. This allows us to determine which contexts favour or disfavour the use of DOM and, in turn, whether each speaker group operates in the same manner. This comparison of the strength and direction of conditioning effects, rather than just comparison of overall rates, provides a more nuanced understanding of the patterns of variation and can show more clearly whether heritage and homeland speakers share a grammar. These comparisons are made via Mixed Effect Models that indicate which factors have a significant effect (and how strong that effect is as well as its direction) when all linguistic and social factors are considered simultaneously. First a mixed effects model (MEM) is fit to all the data combined to see the general picture: which of the factors hypothesized to condition DOM variation actually do, and whether the direction of the effect is as expected. Then models are constructed for each of the four speaker groups separately. Once the best-fitting model is selected via comparison of Akaike Information Criteria (AICc) scores and iterative testing of different combinations of factors, the factor weights (or log odds) assigned to each factor in the model of each speaker group are compared.

We excluded tokens from categories 6-10 of Guardiano's hierarchy as, in one case (7. Animal referents), there were only four tokens and, in the others, there were very large numbers of Os, but only 1-4 tokens, per category, marked with DOM. Thus, they would cloud the results while only illustrating vanishingly low rates of DO-marking, as illustrated in Balasch (2011). After the conversational interview, interviewers asked the Heritage speakers a series of questions constituting the Ethnic Orientation Questionnaire (EOQ). We excluded three who didn't respond to the EO and two who produced no DOM tokens, leaving 19. Each question enquires about preferences between Italian or Canadian language and culture. Each response is scored: 0 for English; 1 for both/mixed; 2 for Italian. We looked at four sets of responses that we thought

might meaningfully correlate to the rate of use of DOM. Scores are averaged over related questions. The sets of questions are summarized in Table 3.

Table 3. *Ethnic orientation questionnaire sections analyzed*

Ethnic Orientation (A1): Whether the speaker reports themselves as more Canadian or more Italian
School (B2): Whether the speaker reports learning it at school or acquiring Italian at home
Network (A2 to A5): Whether the speaker reports many Italians in their social network (friends, neighborhood, at work and in their childhood social network)
Language Preferences (B3, C1 to C5): Whether the speaker reports a preference for speaking English or Italian (with family, with friends, when talking about emotional issues, with (grand)parents and with (grand)children.

We calculated Spearman's rho, a correlation measure based on rank values, between each of the four EOQ scores and the individuals' factor weights from the MEM model in which all speakers were examined together (Table 4). (Using the FW instead of the raw percentage accounts for any effects that might be due to a skewed distribution of tokens.) For the speakers with 0% DOM, we replaced the FW with 0. Positive correlations would reveal that speakers who prefer to speak Italian, engage in Italian-dominant social networks, etc., use DOM more often.

Results

We first look at the use of DOM in the contexts where the syntactic and semantic literature leads us to expect it. We turn first to contexts where DOM is proscribed, noting that it enjoys occasional use, by both Homeland and Heritage speakers there. Then we consider the effects of Ethnic Orientation.

Modeling DOM Where It Is Expected

Table 4 presents the best-fitting model for the dataset as a whole. The overall rate of DOM is 17% for these 309 tokens. All tokens for Gen3 females were excluded because that group categorically omitted DOM. Speaker was included in this model as a random effect but is not reported since the same factor weight (0.5) was reported for all speakers, indicating that none are particularly outliers. Levels of the factors have been reduced from those shown in Table 4 to produce the best-fitting model. Significance levels for each factor are given in parentheses.

Table 4. Model of DOM for all speakers combined

Factor	logodds	n	% DOM	Weight
TYPE OF OBJECT (1.68e-10)				
personal pronouns and proper nouns	1.65	51	59%	0.84
other pronouns	0.42	51	22%	0.60
kinship terms	-0.67	25	12%	0.34
human nouns	-1.40	182	5%	0.20
TYPE OF VERB (0.003)				
dative/accusative	1.42	38	37%	0.81
Telic	-0.18	162	17%	0.46
Psych	-0.17	29	14%	0.46
Other	-1.07	80	10%	0.26
DISLOCATION (0.005)				
dislocated	0.99	17	53%	0.73
SVO	-0.99	292	15%	0.27
DEFINITENESS (0.015)⁷				
definite	0.85	265	19%	0.70
indefinite	-0.85	44	5%	0.30
GENERATION (not significant)				
homeland	[0.30]	58	24%	[0.58]
gen2	[0.10]	64	13%	[0.52]
gen1	[-0.01]	149	17%	[0.50]
gen3	[-0.39]	38	13%	[0.40]
SEX (not significant)				
Male	[0.21]	136	21%	[0.55]
Female	[-0.21]	173	15%	[0.45]

⁷ Neither Definiteness nor Type of Object alone fully account for the distribution of DOM. Definiteness can trigger different DOM rates within one type of object, and types of object differ

Notes: N=309, centered input probability 0.22). n is the token count for that level, %DOM is the rate of DO-marking for that level, FW is the corresponding centered factor weight.

Each linguistic factor tested has a significant effect in the direction predicted by previous literature. However, none come close to exhibiting the categorical effects suggested in the syntactic literature.

Several models we tested suggest some interaction, likely related to generation, as indicated by a different ranking for the levels by percentage of DO-marked tokens vs. by factor weights. For example, in the model presented in Table 4, DO-marked percentages for the generation groups are in the predicted order (most for Homeland to least for Gen3), but not quite in that order by factor weight. However, neither Generation nor Sex has a significant effect here: we lack strong evidence for generational change. This leads us to further analysis in which we consider the generational groups separately.

Table 5 presents the best-fitting models for each of the four speaker groups. A regression model cannot be presented for Gen3 for two reasons. First, the females in that group behave categorically, never producing DOM (and they are excluded from this analysis). Second, the

in rates even when all are definite. It is true, of course, that some object types (e.g., personal pronouns and kinship terms) appear only in definite NPs. However, a model that included an Interaction factor for the two factors fit the data worse than the model provided (according to the AICc criteria). The model provided in Appendix B, with an interaction factor combining Definiteness and Type of Object, illustrated that being Indefinite decreases the likelihood of DOM within the Type of Object categories where both definite and indefinite tokens were produced, confirming that both factors are needed.

. A similar comparison of models was conducted to see if fit could be improved by including the interaction Verb Type * Definiteness, but it improves neither explanatory power nor goodness of fit. It does reveal that Telic (n=21) and Psych verbs (n=8) with Indefinite objects are categorically unmarked for DOM, and that all verb types have tokens with both Definite and Indefinite objects (except the “Other” category).

Gen3 males have categorical behaviour (0% or 100% DOM) for some levels of every factor.

Thus, there are strong effects in Gen3, which we discuss below.

Table 5. Comparison of models for four speaker groups (headings labeled as in Table 2)

	Homeland (n=58, 24% DOM) %			Gen1 (n=149, 17% DOM) %			Gen2 (n=64, 12% DOM) %			Gen3 Males (n=85, 6% DOM) %	
	n	DOM	FW	n	DOM	FW	n	DOM	FW	n	DOM
TYPE OF OBJECT											
other pronouns	10	40	.72	25	20	.99	9	22	.36	7	0
kinship terms	6	33	.63	28	61	1.00	5	60	.88	4	100
personal pronouns and proper nouns	14	43	.60	9	0	[0]	5	20	.73	5	0
human nouns	28	7	.14	87	5	.95	45	4	.08	11	9
DISLOCATION											
dislocated	5	60	.75	8	50	.83	3	33		1	100
SVO	53	21	.25	141	16	.17	61	12		37	11
TYPE OF VERB											
dative/accusative	6	33		13	46	.88	5	60	.91	9	33
Telic	33	30		84	18	.48	36	6	.18	9	0
Psych	0			14	14	.43	6	33	.61	9	0
Other	19	11		38	8	.16	17	6	.22	6	33
DEFINITENESS											
Definite	53	25		130	19		53	15		29	17
indefinite	5	20		19	5		11	0		9	0
SEX											
Male	45	22	.49	68	12	.47	22	9	.34	38	13
Female	13	31	.51	81	22	.53	42	14	.66	47	0

We see again that the rate of DOM decreases from one group to the next (from Homeland at 24% to Gen3 at 6% (for males, and 0% for Gen3 Females, as noted above). We will show that this apparent change is accounted for by different distributions of tokens in the speaker groups and is *not* evidence of a change in progress.

We next discuss the consistency of effects across groups. In Table 5 the factors are listed from strongest to weakest effect, overall (based on consideration of both ranges, canonically calculated by subtracting the smallest FW from the largest, and p-values). Within each factor, the levels are listed from most to least favouring, according to Homeland (centered) factor weights. Using a cut-off of $p < 0.1$ because of small token counts, factor weights are shown for significant factors only.

Type of Object has a significant effect in each group: Impersonal pronoun objects favour DOM the most for Homeland and Gen1 speakers. Gen 2 and 3 speakers use very few impersonal pronouns. While they show lower rates of DOM for impersonal pronouns, those rates may not be representative with so few tokens produced. Kinship terms also strongly favour DOM in all speaker groups, including a categorical effect in Gen3 males. In all cases, nouns referring to humans have very low rates of DOM. Thus, this important factor plays a consistent role from generation to generation.

Dislocation has a strong effect in Homeland and Gen1, with dislocated tokens much more likely to have DOM than SVO tokens. This trend carries across to Gen2 and to Gen3 (males), but doesn't reach significance because there are so few dislocated tokens (3 and 1, respectively). Type of verb also has a consistent effect, which emerges as significant when there are enough tokens produced in the Dative/Accusative category (the most favouring context) and the Other category (the most disfavouring context). That is, Dative/Accusative always most strongly favours DOM, but the factor is not significant in Homeland ($n=6$) and Gen3 ($n=9$). "Other" types of verbs disfavour DOM the most. Some differences in the order of levels is observed between groups but can be attributed to a dearth of tokens.

Definiteness also plays a consistent role: Definite noun objects are more likely to have DOM than indefinite. However, in these four models of smaller datasets, Definiteness never emerges as significant because of the interaction with Type of Object mentioned in footnote 5. This factor, recall, did emerge as significant in the model with all tokens, where we have more tokens to represent each level of each factor.

Finally, we tested the effect of the Sex of the speaker because of the categorical pattern in Gen3: no Gen3 female produces any DOM. However, this is not part of any obvious trend: females produce *more* DOM than males, though never significantly more, in the other three groups.

Speaker was included as a random effect in each model. We report a standard deviation of 1.48 for Homeland and 0.23 for Gen2, but 0 for Gen1 (and not applicable for Gen3). Rates for individuals are listed in Appendix A.

To sum up, rather than successive generations losing the effect of particular factors conditioning the rates of DOM, they simply do not produce enough tokens of certain syntactic and semantic types to produce a significant effect for those factors. Because later generations of speakers have fewer tokens of “Other” pronouns, Dislocated subjects, and Indefinite subjects, they have less opportunity to use DOM, and this accounts for their lower rates. However, in each context where there are sufficient tokens for a contrast to emerge, it does so in a similar way across groups. These differences in distribution also account for the apparent differences in rates in the analysis with all speakers together. Thus, Table 4 shows that Generation is not a significant predictor of DOM rate, while Table 5 shows why this is the case.

The one exception is the Gen3 females. We are at a loss to account for why they never use DOM, when females in other generations use it more than males, and Gen3 males continue to use it.

Non-canonical Use of DOM

We now turn to the cases where our speakers produced DOM in contexts where the theoretical literature does not predict it, that is, the types of objects that appear in the lower part of Guardiano's (2010) hierarchy. Recall that in most published studies of DOM in Spanish, heritage and homeland speakers performed similarly in these contexts (Irizarri van Suchtelen 2016: 102). The speakers involved in the Heritage Language Variation and Change (HLVC) project produced only eight such tokens, across some 30 hours of conversations, distributed as in Table 4. Two of these eight were from a Homeland speaker. Due to the rate of DO-marking for these frequent types of objects being so low, we did not extract non-DO-marked tokens in these categories.

There are two further findings of interest: as we saw above, speakers produced a robust number of tokens (197) from one category in the lower half of the hierarchy that is, 6. Common nouns referring to people, and marked these with DOM only 5% of the time. Speakers produced a total of four tokens that were nouns with reference to an animal: two with DOM (in Table 6 and two without.

Table 6. Distribution of objects with DOM from the bottom of the Guardiano hierarchy

Type of Object	Guardiano	
	category	n (Source)
Pronoun with a non-human referent	5	1 (Gen 3)
Noun with reference to an animal	7	2 (Homeland, Gen 1)
Noun with reference to a thing	8	4 (Homeland, Gen 1, Gen 3)
Abstract noun	10	1 (Gen 1)

These scarce tokens may suggest the beginning of innovative practices in DOM, and not simplification, in this heritage variety. Alternatively, the low rates of DO-marking in categories 5-10 are simply the probabilistic realization of the hierarchy. (11-15) are examples of these unusually DOM-marked Os.

- (11) *sì, **allu** sud l' abbiamo sempre girato*
 yes, **DOM+the** South Pronoun have.1pl always visit.Past
 (Object)

‘Yes, the South we always visited’ (I1F71A, 18: 43)

- (12) *guardo **alla** televisione*
 I watch **DOM+the** TV
 ‘I watch TV’ (I1M61A, 22: 40)

- (13) *abbiamo svuotato **a** mezzo congelatore **alla** roba*
 have. emptied **DOM** half freezer **DOM+the** the thing
 ‘We emptied half the freezer from the thing’ (I1F71A, 3019)

These three non-canonical uses are from 1st generation speakers and show that the use of the preposition can be influenced by other syntactic parameters: in the first example, we see that dislocation triggers the use of the preposition, even in absence of the expected semantics. According to the literature on DOM in spoken Italian and our results (see Table 4, Table 5), word order (with respect to left-dislocation) is a crucial parameter influencing DOM. In (12) and

(14), the verb *guardare* is of interest since it is one of those verbs which in Late Latin could be used both with the dative and the accusative, another parameter which is reported to favour DOM (and the most-favouring type of verb in our analysis, see Table 4). In (11-14), the Os are not human but they are definite: this is symptomatic of the importance of definiteness in DOM.

- (14) Ø *guardano* *al* *libro*
 they Watch.Pres.3pl. **DOM + Article** book
 ‘(they) watch the book’ ((I2M28A, 22:02)

- (15) Ø *leggevo* *a* *qualcosa*
 I Read.Past.1sing **DOM** something
 ‘I used to read something’ (I3F33A, 29:44)

In (15), there are no other syntactic settings that should trigger DOM and it is possible to presume this to be an idiosyncratic usage. This token is particularly interesting in light of the fact that, otherwise, Gen3 females do not exhibit the use of DOM in our sample. This solitary example suggests that at least one member of that group is aware of the DOM construction.

(Non)-Effects of Ethnic Orientation

We considered several measures of language use, preference and ethnic orientation. In no case was there any effect that came close to significance: there is no relationship between use of DOM and whether a Heritage Italian speaker thinks of themselves as more Canadian or more Italian ($\rho = -0.26$, $p = 0.29$); whether they report learning Italian at school vs. acquiring it at home ($\rho = -0.08$, $p = 0.75$); whether they report a larger or smaller number of Italians among their friends, in their neighborhood, at their place of work, nor in their childhood social network

($\rho = -0.1$, $p = 0.69$); nor whether they report speaking more English or more Italian in a range of contexts ($\rho = 0.27$, $p = 0.26$). Additionally, the slight slope of the correlation goes in an unexpected direction in three of the four cases: higher EO scores correspond to (slightly) lower factor weights for DOM in all cases except Language Preferences.

Summary of Findings

Our analyses suggest that the Type of Object is the factor which most strongly predicts the presence or absence of DOM. The rates of DOM for the different Types of Object descend according to Guardiano's hierarchy: the highest rate of DOM is found for her levels 1-3, then 4, then 6. (Level 5 is excluded due to low token count.) Virtually nil rates are found for levels 7-10. DOM is used more when the object is human and known or present.

The Type of Verb is the second most important predictor. To the best of our knowledge, this is the first quantitative analysis of DOM, in any language, to quantify the effect of contrasting verb types and to rank its effect in relation to other factors. Montrul 2004 examined lexical aspect but found no effect; Montrul and Sánchez Walker 2013 report more DOM with verbs that *can* take animate objects while verbs that *must* take an animate object exhibit less DOM, according to Irizarri van Suchtelen (2016: 104). That author considered the effect of animacy of the typical object of each verb, but not in concert with other factors. We show that verbs that were Dative/Accusative-alternating in Latin, Psych Verbs and Telic verbs (to a lesser degree), favour DOM independently of the type of object they co-occur with, as both these factors emerged as significant.

DOM is more frequent with dislocated objects than *in situ* objects.

Finally, Definiteness plays a lesser role (at least in our token set which included only objects with +human reference). This replicates Tippetts' (2010, reported in Schwenter 2014: 242) analysis of three varieties of Spanish, where the animacy of the object has the strongest effect, followed by Definiteness. However, to some extent the explanatory power of definiteness is decreased in our models because of the categorically-definite nature of a few noun types, as noted in Appendix B.

There is no significant effect of the external factors examined: generation, gender and ethnic orientation (except for the categorical lack of DOM for Gen3 females). The lack of correspondence between rates of DOM use and speakers' reported ethnic orientation, social networks, schooling and language practices counts against the possibility that speakers are doing identity work with this variable⁸. Although this is a variable that exhibits variation among homeland speakers (and thus the variation cannot be chalked up to incomplete acquisition, simplification, attrition, etc.), it does not appear to be used to indicate any affinity to Italianness or Canadianness. With only 19 speakers to compare, we must also recognize that this may be a problem of low numbers.

Due to the lack of effect of any of the Ethnic orientation and language preference factors, whose importance has been shown in studies in SLA such as the one by Terry (in this volume), we remain at a loss to account for the distinctive behaviour of the Gen3 females, that is, the categorical absence of DOM in their speech. Recall that, prescriptively, DOM is not part of standard Italian. However, none of the speakers spoke standard Italian during these recordings.

⁸ Li, Bayley and Zangh (in this volume) also found that internal factors in Chinese marker *le* are more important than external ones.

We consider it possible that their speech shows influence of standard Italian. However, no Gen3 females report acquiring Italian entirely at school. With this exception, we can conclude, as did Leonetti (2008: 60, quoted in Balasch 2011: 116), “DOM in Romance is sensitive to a series of dimensions that make up a multi-dimensional bundle of factors.” This is true for all four generations we compared.

Discussion and Conclusion

The analyses, as summarized just above, show that heritage speakers maintain the relevant semantic (e.g., animacy, definiteness, telicity), syntactic (position of subject) and morphological (pronoun vs. noun vs. proper noun) distinctions that homeland speakers have, and apply them to the distribution of DOM in a very similar fashion. Furthermore, the presence and direction of these effects all support the suggestions from grammatical descriptions about where DOM is used or expected, vs. where it is proscribed. However, none come close to exhibiting the categorical effects suggested in the syntactic literature. As has been frequently shown, stochastic trends in performance align with categorical claims about competence.

The majority of quantitative, analytic studies of DOM, with which we can compare our findings, have examined Spanish (though see Dufter and Stark 2008 on DOM with indirect objects in Italian). Such studies have highlighted three needs: (1) to compare multiple generations of heritage speakers, (2) to quantitatively analyze homeland (baseline) data, and (3) to examine additional languages.

Following these suggestions, we examined Italian speakers, and showed that heritage and homeland speakers retain the same grammar in a remarkably robust fashion. We have seen this consistency previously among Heritage Italian speakers: no cross-generational differences were

found in analyses of VOT, both word-initial (where English has longer VOT than Italian, cf. Nagy and Kochetov 2013) and in particular word-internal contexts (where Italian has longer VOT than English, Nodari et al. 2019) and in variable null subject use (cf. Nagy 2015). We can now add DOM to the list of variables that have been examined by comparing spontaneous speech samples from Homeland and Heritage Italian generations and do not show attrition-like cross-generational effects in how the variable is conditioned. In the current study we again also see little difference in rates of use, and these differences are accounted for by the different distribution of tokens produced by each group. This is similar to an outcome reported for young Spanish teenagers, in an experimental study which also compared homeland and heritage speakers, by Guijarro-Fuentes and Marinis (2011:227):

the two groups showed a similar pattern of errors regardless of their grade of bilingualism.

More importantly, their linguistic performance does not seem to be related to any of the external factors included in this study.

Additional evidence that speakers in all generations are still aware the syntactic and semantic constraints on DOM is the virtually categorical lack of DO-marking of objects in the bottom half of the animacy-related hierarchy (see Guardiano 2010), similar to the 0.7% rate of DOM reported for two corpora of Homeland Spanish speakers (Balasch 2011: 116).

The consistent effects of the linguistic factors, and the lack of effect for the social factors, supports the trend reported in previous studies that the role of internal factors is predominant. Montrul et al. (2015: 604), for example, compared DOM in three heritage groups (Hindi, Romanian and Spanish migrants in the US), showing that:

While both internal linguistic factors and external sociolinguistic factors play a role in the structural changes observed, we support the conclusion that the structural properties of the

DOM markers together with the syntax of definiteness and specificity in each language (more than external factors) seem to account for the degree of DOM erosion in each language.

While Montrul et al.'s (2015) experimental study reports erosion, our corpus-based findings for Italian instead robustly show our speakers' stochastic sensitivity to the syntactic and semantic factors described in the theoretical literature, consistently across generations. As discussed in (Nagy, 2015), we may attribute the difference in outcomes to methodological differences. These may begin with the selection of different types of speakers: HLVC participants are volunteers who say they are comfortable to talk for an hour in their heritage language while some experimental heritage-language studies exclude speakers who are "too fluent." The context of data collection also differs: HLVC speakers are recorded by a fellow heritage-language speaker in a context (such as their home or neighborhood café) where they are accustomed to speaking their heritage language, while many experimental studies are conducted in institutional settings where the heritage language has been proscribed throughout their school and/or work experience. Finally, the difference between living in a highly multilingual city like Toronto in an officially bilingual country vs. many of the American cities where experimental tasks were conducted may influence how heritage languages are maintained. So, while some of the studies of heritage Spanish report findings similar to ours (cf. Guijarro-Fuentes and Marinis, 2011; Irizarri van Suchtelen, 2016), these methodological differences may account for the outcomes that differ between previous studies of DOM and this first variationist study of DOM in Italian.

References

- Aalberse, Suzanne, Ad Backus, and Pieter Muysken. 2019. *Heritage Languages: A Language Contact Approach*. Amsterdam and Philadelphia: John Benjamins.
- Aissen, Judith. 2003. "Differential Object Marking: Iconicity vs. Economy." *Natural Language and Linguistic Theory* 21: 435–483.
- Balasch, Sonia. 2011. "Factors Determining Spanish Differential Object Marking within Its Domain of Variation." In *Selected Proceedings of the 5th Workshop on Spanish Sociolinguistics*, edited by Jim Michnowicz and Robin Dodsworth, 113–124. Somerville, Mass.: Cascadilla Proceedings Project.
- Berretta, Monica. 1989. "Sulla presenza dell'oggetto preposizionale in italiano: note tipologiche." *Vox Romanica* 48: 13–37.
- Berruto, Gaetano. 2006. *Sociolinguistica dell'Italiano Contemporaneo*. Rome: Carocci.
- Boeddu, Daniela. 2017. "Estudio diacrónico del acusativo preposicional sardo." Doctoral dissertation, Universidad del País Vasco.
- Bossong, Georg. 1991. "Differential Object Marking in Romance and Beyond." In *New Analyses in Romance Linguistics*, edited by Georg Bossong, Dieter Wanner, and Douglas Kibbee, 143–171. Amsterdam: John Benjamins.
- Comrie, Bernard. 1989. *Language Universals and Linguistic Typology*, 2nd ed. Chicago: University of Chicago Press.
- Cortelazzo, Manlio. 1972. *Avviamento Critico allo Studio della Dialettologia Italiana, III, Lineamenti di Italiano Popolare*. Pisa: Pacini.
- D'Agostino, Mari. 2012. *Sociolinguistica dell'Italia Contemporanea*. Bologna: Il Mulino.
- De Blasi, Nicola. 2014. *Geografia e Storia dell'Italiano Regionale*. Bologna: Il Mulino.

- Di Salvo, Margherita. 2017. "L'oggetto preposizionale nell'italiano parlato in contesto dell'extraterritorialità." *L'Italia Dialettale* 78: 93–124.
- Di Salvo, Margherita. 2019. *Repertori Linguistici degli Italiani all'Esteri*. Pisa: Pacini.
- Di Venanzio, Laura, Katrin Schmitz, and Anna-Lena Rumpf. 2012. "Objektrealisierungen und -auslassungen bei transitiven Verben im Spanischen von Herkunftssprechern in Deutschland." *Linguistische Berichte* 232: 437–461.
- Dufter, Andreas, and Elisabeth Stark. 2008. "Double Indirect Object Marking in Spanish and Italian." In *Theoretical and Empirical Issues in Grammaticalization*, edited by Elena Seoane and María José López-Couso, 111–129. Amsterdam: John Benjamins.
- Fiorentino, Giuliana (ed.). 2003a. *Romance Objects. Transitivity in Romance Languages*. Berlin: Mouton de Gruyter.
- Fiorentino, Giuliana. 2003b. "Prepositional Objects in Neapolitan." In *Romance Objects: Transitivity in Romance Languages*, edited by Giuliana Fiorentino, 117–151. Berlin: Mouton de Gruyter.
- Guardiano, Cristina. 2000. "Note sull'oggetto diretto preposizionale in siciliano." *L'Italia Dialettale* LXI: 7–41.
- Guardiano, Cristina. 2010. "L'oggetto diretto preposizionale in siciliano. Una breve rassegna e qualche domanda." In *Quaderni di Lavoro ASIIt 2010: Studi sui dialetti della Sicilia*, edited by Jacopo Garzonio, 95–115. Padova: Unipress.
- Guijarro-Fuentes, Pedro, and Theodoros Marinis. 2011. "Voicing Language Dominance: Acquiring Spanish by British English/Spanish Bilingual Children." In *Bilingual Youth: Spanish in English-Speaking Societies*, edited by Kim Potowski and Jason Rothman, 227–248. Amsterdam and Philadelphia: John Benjamins. <https://doi.org/10.1075/sibil.42.14gui>

- Hill, Virginia. 2015. *Formal Approaches to DPs in Old Romanian*. Leiden: Koninklijke Brill NV.
- Hopper, Paul, and Sandra Thompson. 1980. "Transitivity in Grammar and Discourse: The Transitivity Hypothesis." *Language* 56: 251–299.
- Iemmolo, Giorgio. 2009. "La marcatura differenziale dell'oggetto in siciliano antico." *Archivio Glottologico Italiano* 94 (2): 185–225.
- Irizarri van Suchtelen, Pablo. 2016. "Spanish as a Heritage Language in the Netherlands: A Cognitive Linguistic Exploration." Doctoral dissertation, Radboud University.
<https://repository.ubn.ru.nl/bitstream/handle/2066/159312/159312.pdf>.
- Labov, William. 1984. "Field Methods of the Project on Linguistic Change and Variation." In *Language in Use: Readings in Sociolinguistics*, edited by John Baugh and Joel Sherzer, 28–53. Englewood Cliffs, N.J.: Prentice Hall.
- Laca, Brenda. 2006. "El objeto directo." In *Sintaxis historica del español. Vol 1: La frase verbal*, edited by Concepción Company, 197–204. México City: Universidad Nacional de México.
- Leonetti, Manuel. 2008. "Specificity in Clitic Doubling and in Differential Object Marking." *Probus* 20: 33–66.
- Li, Wei, Bayley, Robert, Zhang, Xinye. 2022. "An Investigation of the Use of the Multifunctional Particle *-le* by Second Language Learners of Mandarin Chinese", this volume.
- Lopez, Luis. 2012. *Indefinite Objects: Scrambling, Choice Functions, and Differential Marking*. Cambridge, Mass.: MIT Press.
- Loporcaro, Michele. 2009. *Profilo Linguistico dei Dialetti Italiani*. Rome-Bari: Laterza.
- Maiden, Martin, and Mair Parry. 1997. *The Dialects of Italy*. London: Routledge.

- Marchese, Floriana. 2016. "Il lessico del dialetto di Polia (VV)." Doctoral dissertation, Università di Firenze.
- Mardale, Alexandru. 2008. "Microvariation within Differential Object Marking: Data from Romance." *Revue Romaine de Linguistique* LIII (4): 448-467.
- Mardale, Alexandru-Daniel. 2009. *Les prépositions fonctionnelles du roumain: études comparatives sur le marquage casuel*. Paris: L'Harmattan.
- Montrul, Silvina. 2004. "Subject and Object Expression in Spanish Heritage Speakers: A Case of Morphosyntactic Convergence." *Bilingualism: Language and Cognition* 7 (2): 125–142.
- Montrul, Silvina, Rakesh Bhatt, and Roxana Girju. 2015. "Differential Object Marking in Spanish, Hindi and Romanian as Heritage Languages." *Language* 91: 564–610.
- Montrul, Silvina, and Melissa Bowles. 2009. "Back to Basics: Incomplete Knowledge of Differential Object Marking in Spanish Heritage Speakers." *Bilingualism: Language and Cognition* 12(3): 363–383.
- Montrul, Silvina, and Noelia Sánchez-Walker. 2013. "Differential Object Marking in Child and Adult Spanish Heritage Speakers." *Language Acquisition* 20 (2). 109-132.
- Nagy, Naomi. 2009. "Heritage Language Variation and Change."
<http://ngn.artsci.utoronto.ca/HLVC/>. Accessed 23 January 2020.
- Nagy, Naomi 2011. A Multilingual Corpus to Explore Geographic Variation. *Rassegna Italiana di Linguistica Applicata* 43 (1-2): 65–84.
- Nagy, Naomi. 2015. "A Sociolinguistic View of Null Subjects and VOT in Toronto Heritage Language." *Lingua* 164B: 309–327.
- Nagy, Naomi and Alexei Kochetov. 2013. "VOT Across the Generations: A Cross-Linguistic Study of Contact-Induced Change." In *Multilingualism and Language Contact in Urban*

- Areas: Acquisition – Development – Teaching – Communication*, edited by Peter Siemund, Ingrid Gogolin, Monika Schulz, and Julia Davydova, 19–38. Amsterdam and Philadelphia: John Benjamins.
- Nagy, Naomi and Miriam Meyerhoff. 2015. “Extending ELAN into Variationist Sociolinguistics.” *Linguistic Vanguard* 1 (1): 271–281.
- Nocentini, Alberto. 1985. “Sulla genesi dell’oggetto preposizionale nelle lingue romanze.” In *Studi linguistici e filologici per Carlo Alberto Mastrelli*, edited by Agostiniani, Luciano; Vittoria Grazi and Alberto Nocentini, 299–311. Pisa: Pacini.
- Nodari, Rosalba, Chiara Celata, and Naomi Nagy. 2019. “Socio-indexical Phonetic Features in the Heritage Language Context: Voiceless Stop Aspiration in the Calabrian Community in Toronto.” *Journal of Phonetics* 73: 91–112.
- Palermo, Massimo. 2015. *Linguistica Italiana*. Bologna: Il Mulino.
- Pittau, Massimo. 1972. *Grammatica del Sardo-nuorese. Il Più Conservativo dei Parlari Neolatini*. Bologna: Pàtron.
- Pottier, Bernard. 1968. “L’emploi de la préposition ‘a’ devant l’objet in espagnol.” *Bulletin de la Société de Linguistique de Paris* 63: 63–85.
- Rehner, Kathrine, Mougeon, Raymon, Mougeon, Françoise. 2022. Variation in Choice of Prepositions with Place Names on the French L1–L2 Continuum in Ontario, Canada. This volume.
- Rodríguez-Mondoñedo, Miguel. 2008. “The Acquisition of Differential Object Marking in Spanish.” *Probus* 20: 111–145.
- Rohlfs, Gerard. 1966. *Grammatica Storica della Lingua Italiana e dei suoi Dialetti*. Torino: Einaudi.

- Silverstein, Michael. 1976. "Hierarchy of Features and Ergativity." In *Grammatical Categories in Australian Languages*, edited by Robert M. W. Dixon, 112–171. New Jersey: Humanities Press.
- Sornicola, Rosanna. 1997. "L'oggetto preposizionale in siciliano antico e in napoletano antico." *Italienische Studien* 18: 66–80.
- Sornicola, Rosanna. 1998. "Processi di convergenza nella formazione di un tipo sintattico: la genesi ibrida dell'oggetto preposizionale." In *Les nouvelles ambitions de la linguistique diachronique, Actes du XXIIe Congrès International de Linguistique et de Philologie Romanes (Bruxelles 23-29 Juillet 1998) II*, edited by Annick Englebert, 419–427. Brussels: Max Niemeyer Verlag.
- Telmon, Tullio. 1993. "Varietà regionali." In *Introduzione all'Italiano Contemporaneo: La Variazione e gli Usi*, edited by Alberto Sobrero. 93–149. Rome-Bari: Laterza.
- Terry, Kristen Kennedy. 2022. Sociostylistic Variation in L2 French: What schwa Elision Patterns Reveal about Language Acquisition during Study Abroad, this volume.
- Ticio, Emma and Luisa Avram. 2015. "The Acquisition of Differential Object Marking in Spanish and Romanian: Semantic Scales or Semantic Features?" *Revue roumaine de linguistique* 4: 383–402.
- Tippets, Ian Robert. 2010. "Differential Object Marking in Spanish: A Quantitative Variationist Study." Doctoral dissertation, The Ohio State University.
- Von Heusinger, Klaus. 2008. "Verbal Semantics and the Diachronic Development of DOM in Spanish." *Probus* 20 (1): 1–31.
- Wittenburg, Peter, Hennie Brugman, Albert Russel, Alex Klassmann, and Han Sloetjes. 2006. "ELAN: A Professional Framework for Multimodality Research." In: *Proceedings of LREC*

2006, Fifth International Conference on Language Resources and Evaluation, edited by Nicoletta Calzolari, Khalid Choukri, Aldo Gangemi, Bente Maegaard, Joseph Mariani, Jan Odijk, and Daniel Tapias, 1556-1559. Paris: European Language Resources Association.

Appendix A: Individual speaker information

Speaker	N	% DOM	EO	Net- work	School	Speaking Prefs
IXF14A	3	100				
IXF18A	1	0				
IXF22A	6	0				
IXF43A	3	33				
IXM35A	24	25				
IXM47A	6	0				
IXM61A	5	40				
IXM63A	10	20				
I1F61A	25	12	2	1.5	0	1.8
I1F65A	0	n.a.	2	1.5	0	2.0
I1F71A	41	22	1	1.5	n.d.	0.5
I1F73A	15	40	1	1.0	1	1.0
I1M60A	18	6	0	1.5	0	1.2
I1M61A	17	29				
I1M62A	13	0	2	1.5	1	1.2
I1M75A	20	10	1	1.3	1	0.5
I2F32A	1	0				
I2F34A	16	13	1	0.0	1	0.7
I2F44A	10	30	1	0.3	2	1.2
I2F45A	15	7	2	2.0	2	2.0
I2M14A	0	n.a.	1	1.5	2	1.2
I2M19A	6	17				
I2M28A	7	0	1	1.5	2	0.3
I2M30A	9	11	1	0.8	2	0.5

I3F20A	5	0	1	1.0	2	1.2
I3F21A	8	0	1	1.5	1	0.8
I3F31A	14	0	2	1.3	1	0.2
I3F33A	20	0	1	0.8	1	0.7
I3M18A	9	11	1	1.3	1	0.7
I3M22A	7	0	1	1.0	2	0.2
I3M27A	9	0	1	1.0	2	0.0
I3M28A	13	31	1	1.5	2	0.7
<i>mean</i>	11.1	15	1.2	1.2	1.3	0.9
<i>median</i>	9.0	11	1.0	1.3	1.0	0.7
<i>std. dev.</i>	8.7	21	0.51	0.48	0.73	0.57

Appendix B: Kind of Object in a model with an interaction between Definiteness and Type of Object

Factor	n	% DOM	FW
definite pers. pronouns and proper names	49	61%	0.91
definite impers. pronouns	33	30%	0.75
definite kinship term	25	12%	0.49
definite human nouns	158	5%	0.30
indefinite impers. pronouns	18	6%	0.26
indefinite human nouns	24	4%	0.20
