

# Quantifying trends of indefiniteness strategies in bilingual speakers of Sicilian and Italian

Vincenzo Nicolò Di Caro

University of Catania

 ORCID ID 0000-0003-2604-2172

vincenzo.dicaro@unict.it

Giuseppe Samo

Beijing Language and Culture University

 ORCID ID 0000-0003-3449-8006

samo@blcu.edu.cn



**How to cite:** Di Caro, Vincenzo Nicolò & Giuseppe Samo. 2025. Quantifying trends of indefiniteness strategies in bilingual speakers of Sicilian and Italian. In *Italian dialects and language contact: Theory, typology, data*, eds. Mario Wild, Stefano Negrinelli, Federica Breimaier, Stefano Cristelli, Michele Loporcaro & Tania Paciaroni. Special issue of *Isogloss. Open Journal of Romance Linguistics* 11(6)/4, 1-23.

**DOI:** <https://doi.org/10.5565/rev/isogloss.488>

## Abstract

In this contribution, we use quantitative methods to account for trends of indefiniteness strategies (e.g. Italian *Bevo del vino* ‘I drink some wine’) in bilingual speakers of central Sicilian and Italian. These varieties differ in the available set (in terms of grammaticality and productivity) of indefiniteness strategies. The usage of these strategies may vary across speakers in this particular linguistic environment, showing tendencies towards one or the other variety. These trends can be easily quantified exploring relevant indices, such as the Index of Language Dominance (ILD; Birdsong et al. 2012). We perform a data analysis on the set of data collected, curated and annotated from two experimental situations by Di Caro (2023). In this respect, we also take into consideration the values annotated for the ILD, a quantitative measure to map grammars of bilinguals, although originally designed for linguistic scenarios where the two varieties are both official languages in the same country or in different countries. The strongest hypothesis we test is a clear correlation between the ILD and sets of syntactic strategies in Deliano. However, our results found no strong correlations

between these two variables. We suggest that new parameters should be created to assess language dominance in bilingual speakers of Italian and other Italo-Romance varieties.

**Keywords:** indefiniteness, Sicilian dialects, weak indefinites, determiners, bilingualism

---

## 1. Introduction

In the present paper we propose a quantitative analysis of the results of a study aiming at investigating the different strategies to mark indefiniteness in the Italo-Romance variety spoken in Delia (cf. Di Caro 2023), taken as a representative of central Sicilian. This variety, as most Italo-Romance varieties spoken throughout Italy, instantiates a case of bilingual community, in the sense of Rowe & Grohmann (2013). Central Sicilian and Italian differ with respect to the strategies used to mark indefiniteness (cf. Brasoveanu & Farkas 2016, Cardinaletti & Giusti 2018 and references therein). As it will be discussed in detail in Section 2, central Sicilian lacks a strategy which is productive in Standard Italian (i.e. the partitive determiner homophonous to the preposition *di* ‘of’ + a definite article; e.g. *del vino* ‘some wine’ lit. ‘of-the wine’). Asymmetries are also found with respect to less-specific quantificational elements or pseudo-partitives (e.g. Italian *un po’ di* ‘a bit of’).

The research questions pursued in our study are the following: i) is it possible to find asymmetries between bilingual speakers in Delia in the selection of the different strategies of indefiniteness marking? ii) is the Index of Language Dominance (cf. Birdsong et al. 2012; see Section 2.3) a good predictor of the bilingual speakers’ selection of the strategies of indefiniteness marking?

We explore quantitative tools, in line with the recent publications on quantitative approaches to generative grammar (Van Craenenbroeck & Van Koppen 2023, Merlo & Samo 2023; see also Pescarini 2022 for a quantitative approach to Romance dialectology). We adopt the frequency of syntactic strategies as a dependent variable, and the Index of Language Dominance to map variation in a representative sample of a bilingual population (Italian and central Sicilian) as an independent variable.

The two linguistic variables (i.e., indefiniteness and the Index of Language Dominance in bilingual populations) call for a brief definition which is provided in Section 2, Section 3 presents the materials and methods of our study. The quantitative and qualitative results are discussed in Section 4. Section 5 draws the conclusions and provides some open questions for future research.

## 2. Indefiniteness, bilingualism and language dominance

This section provides an overview of the three concepts, and the related linguistic areas, that are crucial to the understanding of the present study.

### 2.1. Indefiniteness

According to Brasoveanu & Farkas (2016: 238-39), the best way to define indefiniteness is by first defining the opposite concept, namely definiteness. If

definiteness depicts nominal expressions that refer to an individual already mentioned in the discourse (Milsark 1977), indefiniteness regards nominals that refer to an element that either (i) introduces a new individual in the discourse or (ii) may not refer to any individual at all. The semantic types of indefinites that can be found in Italo-Romance, and, more in general, in Romance are shown in (1):

(1) Uncontroversial or weak indefinites (see Brasoveanu & Farkas 2016):

a. Italian, Di Caro (2023: 2; see also Giusti 2021: 263)

*Bevo*                       $\emptyset$         /*il*        /*del*        *vino*.  
 drink.PRS.1SG                      the        of.the        wine  
 ‘I drink wine.’

b. Piedmontese, Berruto (1974: 57)

*Se-i*                      *fyse*                      *d'*                      *aqua*.  
 if-there                      be.SBJV.3SG        of                      water  
 ‘If there was water.’

Quantificational or strong indefinites (see McNally 2020):

c. Italian

*Alcuni* *ragazzi*.  
 some        boys  
 ‘Some boys.’

d. Italian, adapted from Giusti (2021: 263; see also Brasoveanu & Farkas 2016: 258)

*Alcuni dei* *ragazzi*.  
 some        of.the        boys  
 ‘Some of the boys.’

Pseudo-partitives of the type ‘a bit of’ (see Selkirk 1977):

e. Italian, Cardinaletti & Giusti (2020: 681)

*Vorrei*                      *comprare*                      *un*        *po'*        *di*        *fiori*.  
 want.COND.1SG                      buy.INF                      a                      bit                      of                      flowers  
 ‘I would like to buy some flowers.’

Marked indefinites introduced by the complex determiner ‘a certain’ (see Schwarz 2011):

f. Italian, Di Caro (2023: 22)

*C'*        *è*                      *un*        *certo*                      *timore* *nelle* *sue*        *parole*.  
 there        is                      a                      certain                      fear        in.the        his                      words  
 ‘There’s a certain fear in his words.’

As regards the four strategies concerning weak indefinites (namely ART, bare DI, DI+ART and the ZERO determiner; cf. (1a, b),<sup>1</sup> Cardinaletti & Giusti (2018) provide a syntactic account according to which those strategies are the result of the combination of the overt/covert realization of the two positions available inside the

<sup>1</sup> Italian does not generally display bare DI. It is only found in restricted contexts (with quantified nouns), namely with right dislocated bare nouns, where it is obligatory, and with left dislocated bare nouns, where it is optional (Giusti 2021). On the other hand, it is very productive in northern Italy (hence the example from Piedmontese in (1b)).

DP (i.e. SpecDP and D; Abney 1987).<sup>2</sup> These are shown in Table 1 with the plural noun *libri* ‘books’.

**Table 1.** The combination of SpecDP and D in Italo-Romance

SpecDP	D	Example
∅	∅	<i>libri</i> ‘books’
∅	<i>i</i>	<i>i libri</i> ‘the books’
<i>de</i>	∅	<i>di libri</i> ‘of books’
<i>de</i>	<i>i</i>	<i>dei libri</i> ‘of the books’

**Source:** adapted from Lebani & Giusti (2022: 3)

Cardinaletti & Giusti (2018: 139-140) also take into account the determiner *certo* ‘certain’. In Italo-Romance varieties this determiner is generally used to add a specific indefiniteness nuance, but in some southern varieties with mass and plural nouns it has been documented as a basic indefinite, as in (2) (see also Giammarco 1979: 141, Ledgeway 2009):

(2) Abruzzese, adapted from Rohlfs (1968: 119)

<i>Cə</i>	<i>ʃta</i>	<i>cirtə</i>	<i>pəʀzonə.</i>
there	stay.PRS.3SG	certain	people
‘There are some people.’			

Deliano resorts to the two main strategies shown for Italian to mark indefiniteness, namely ART (cf. (3)) and ZERO (cf. (4)).<sup>3</sup>

(3) Deliano

<i>’Ncapu</i>	<i>la</i>	<i>mènzola</i>	<i>cci=</i>	<i>su</i>	<i>li</i>	<i>libbra.</i>
on	the	shelf	there=	are	the	books
‘There are some books on the shelf.’						

<sup>2</sup> More in detail, Cardinaletti & Giusti (2018: 135) claim that the indefinite operator is in SpecDP, since, according to Giusti (2002, 2015), the definite article in D realizes the functional features of gender and number associated with N, the latter generally not conveying any semantics of definiteness.

<sup>3</sup> The Deliano example in (4) is the natural counterpart of the Italian *Non leggo ∅ libri* ‘I don’t read books’, while It. *∅ Libri non ne leggo* would be syntactically marked. For the syntactic difference of the unmarked sentences in Deliano with respect to Italian in terms of the systematic use of Ethical Dative, Focus Fronting and Syntactic Extraposition we refer the interested reader to Di Caro (2023: 3-4).

## (4) Deliano

Ø *Libbra*            *nun= ni= ljiggiu.*  
 books                NEG= PRTV= read.PRS.1SG  
 ‘I don’t read books.’

However, an interesting point of divergence can be detected, as, according to Giusti (2021: 294), southern Italo-Romance varieties generally lack DI+ART. The data from Deliano confirm that DI+ART is not a productive strategy to express indefiniteness. As a consequence, Deliano resorts to pseudo-partitives more often than Italian. In this respect, a lexical specification emerges according to which *tanticchja di* ‘a bit of’ is used with mass nouns (cf. (5a)) and *нна пуику ди* ‘a bit of’ is used with plural count nouns (cf. (5b)).

## (5) Deliano, adapted from Di Caro (2023: 8)

- a. *Ppi stasira nn’= accattammu tanticchja di murtatella.*  
 for tonight REFL= buy.PRS.1PL a\_bit of mortadella  
 ‘We can buy some mortadella for tonight’s dinner.’
- b. *Ppi stasira nn’= ordinammu нна пуику ди panina.*  
 for tonight REFL= order.PRS.1PL a bit of sandwiches  
 ‘We can order some sandwiches for tonight’s dinner.’

Di Caro (2023: 8) reports that, to a deeper analysis, pseudo-partitive *нна пуику ди* is used by older speakers to refer to greater indefinite quantities and grammaticalized *quattru* ‘four’ is used for the basic indefinite with plural count nouns (cf. *Ppi stasira nn’ordinammu quattru panina* ‘We can have some sandwiches for tonight’s dinner.’). On the other hand, grammaticalized ‘two’ (see Giusti 2021: 280), surfacing in Deliano as *du* or *dui*, is used for very small quantities:

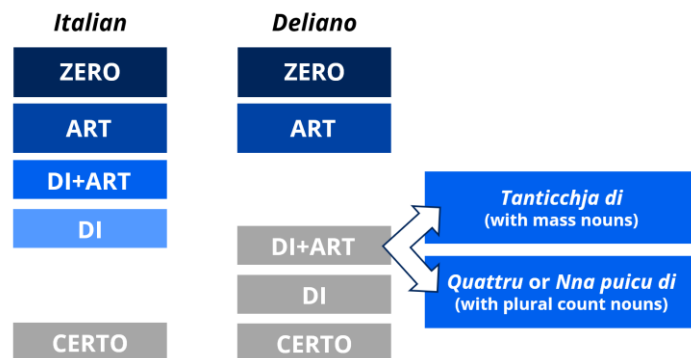
## (6) Deliano

*Cugli du margariti e si= nni= jì.*  
 pick.PST.3SG two daisies and REFL= thence= go.PST.3SG  
 ‘He just picked a couple of daisies and left.’

To conclude the overview on the Deliano strategies of indefiniteness marking, in this variety, and more generally in Sicilian, *certu* ‘certain’ is not used as a basic indefinite. Figure 1 summarizes the differences between Italian and Deliano in this respect, where the blue areas indicate the strategies that are productive in each variety, while the grey areas those that are not.<sup>4</sup>

<sup>4</sup> Note that, in Figures 1 and 2, the Deliano strategies labelled as ‘*Tanticchja di*’ (upper box) and ‘*Quattru* or *Nna puicu di*’ (lower box) are linked to the Italian strategy ‘DI+ART’, because they can be considered as their closest counterparts. Nevertheless, the arrows do not indicate any statistically significant correlation between the use of these strategies and lack of DI+ART in Deliano, which requires further research to be confirmed. In fact, an anonymous reviewer points out that some northern Italian dialects of the Lombard and Venetan areas (cf. Rohlfs 1969: §957), while displaying productive DI+ART indefinite marking, can also resort to pseudo-partitives.

**Figure 1.** Differences between the strategies of indefiniteness marking of Italian and Deliano



We can now turn to the definitions of bilectalism and Index of Language Dominance.

## 2.2. Bilectalism

The linguistic scenario of Italy is known to be particularly rich and varied (cf. Maiden & Parry 1997: 1), with most speakers being bilingual, in the sense that they can speak, to different degrees, Italian (either Standard or Regional) and another Italo-Romance variety, which is usually referred to with the term ‘dialect’.<sup>5</sup> The picture is made more intricate by the presence of other Italo-Romance and non-Italo-Romance languages in some specific areas, such as Alto Adige/Südtirol (where Italian and German are spoken together with an Italian dialect and a Bavaro-Tyrolese dialect), the Aosta Valley (where Italian and French are spoken mainly together with Franco-Provençal), some central and southern areas where Italian and Albanian (Arbëresh) are spoken together with an Italian dialect, or the Sicilian areas where some Gallo-Italic varieties are also spoken (cf. Berruto 1989; De Angelis 2023).

This sociolinguistic setting is sometimes referred to in the relevant Italian literature as ‘dilalia’ (cf. Berruto 1987, 1989, 1993; see also Di Caro 2022: §1.1.4). A sociolinguistic setting is to be considered as ‘dilalic’ when the high variety in the speakers’ repertoire is more and more used in informal contexts too, which are the only contexts where the low variety is spoken. Here lies the main difference between dilalia and diglossia (as described in Ferguson 1959, 1991), the latter regarding a number of sociolinguistic configurations, very common throughout the world, where the high variety is never the mother tongue of any speaker.<sup>6</sup>

The two varieties that mostly characterise dilalia are greatly unbalanced in terms of language usage in written contexts. Italian is, in fact, the only official language and thus the only language allowed in mass media, education, politics and official communication in general. Not only are dialects disallowed in any official

<sup>5</sup> See also §1 of the introduction to this Special Issue.

<sup>6</sup> North African countries are traditionally indicated as cases of diglossic communities (Ferguson 1959). But the Italian peninsula after the Unification in 1861 was also characterized by diglossia, since Italian was only used in official written communication and nobody learnt it as their mother tongue (see De Mauro 1972).

context, but they generally lack any shared written code. This fact, on the one hand, discourages bilingual speakers to promote their own dialects in written contexts and, as a consequence, prevents dialects from being an actual alternative to Italian. On the other hand, it has a negative impact on the language attitudes towards writing in dialect and towards the use of dialects in general, in a sort of vicious cycle that affects all the areas characterised by dilalia (i.e., all of Italy with the exception of Tuscany and Rome; cf. Berruto 1987) and disfavours any political action towards the promotion of dialects.

Nowadays, a more common alternative to the label dilalia, which we use in the present paper, is ‘bilectalism’. This term was first used by Rowe & Grohmann (2013) to indicate the asymmetry between the two varieties spoken by the community living in the Greek speaking part of Cyprus, i.e. a standard variety, namely Standard Modern Greek, and a dialect, namely Cypriot Greek, characterized by structural proximity. Thus, we believe that bilectalism, *mutatis mutandis*, can also be applied to Italy.<sup>7</sup>

The dialect spoken in Delia does instantiate a case of bilectal scenario (see also Kupisch et al. 2023: 2), since this central Sicilian village, with a population of about 4,000, has a very low number of monolingual dialectal speakers left (most of them are over 60). Most monolingual Italian speakers among the younger part of the population have at least a passive competence of Deliano. A study based on a translation task reported in Di Caro (2022: §3.2.2.1) shows that the Deliano lexicon is still vital overall, i.e. it is somehow resisting the convergence towards Italian – which is typical of the smaller communities in Italy, as pointed out by Cerruti & Regis (2020) – but the results of the younger group tested (aged 14-30) show the first signs of the pressure exerted by the Italian lexicon.

### 2.3. Language dominance

This study on Deliano draws from Gertken et al. (2014). Bilingual speakers either use one of the two varieties more often than the other or are better at one of them. In both cases they are said to be dominant in that variety. Language dominance is a continuous and relative construct (see Grosjean 2001, Gertken et al. 2014), so instead of just saying that a bilingual speaker is more dominant in variety A than in B, it would be more accurate to say that that speaker is more dominant in variety A to different degrees than in B. When bilingual speakers are dominant in neither variety, they are said to be perfectly balanced. Birdsong (2014: 372) states that “In the context of bilingualism, dominance refers to observed asymmetries of skill in, or use of, one language over the other”. Note that being balanced does not imply a high degree of proficiency for those speakers. Interestingly, Birdsong (2015) points out that, in terms of measuring, dominance in bilingual speakers is totally comparable to the more general dominance regarding the use of the right or left hand.

The reference tool to calculate the Index of Language Dominance (or ILD) from is the Bilingual Language Profile (or BLP; Birdsong et al. 2012). This is an open

---

<sup>7</sup> Note that the similarities between the Italian and the Cypriot scenarios are of sociolinguistic nature, which are relevant for the present paper. From an internal linguistic perspective, on the other hand, the main difference lies in the fact that Cypriot Greek is a dialect of Greek, while Sicilian is not a dialect of Italian, as both independently evolved from Vulgar Latin (in Coseriu’s 1980 terms, Italian and Sicilian are primary dialects).

access self-assessment questionnaire (both in printable paper or online versions),<sup>8</sup> available in both the languages one wants to measure the dominance of and designed for adult speakers (especially for some items requiring a certain degree of introspection by the participant; Gertken et al. 2014). Four modules make up the questionnaire, each assessing an aspect of the language dominance, namely i) language history, ii) language use, iii) language proficiency (on speaking, listening, reading and writing), iv) language attitudes. To those modules a section is added containing sociolinguistic information about the participant, which does not contribute to the calculation of the ILD and can thus be personalized by the researcher according to the specific purposes of their study (Gertken et al. 2014). The Index is the result of the subtraction of the indexes as individually calculated for the two varieties. It ranges from -218 (total dominance of B) to +218 (total dominance of A). The value 0 indicates a perfectly balanced bilingual.

Di Caro (2023) indicates two shortcomings of the BLP as a tool that is not originally conceived for a dilalic situation like the one under discussion but instead to assess language dominance in those communities speaking two languages that are standard either in the same countries or abroad (e.g. English and Spanish in the United States or English and French in Canada): i) the BLP could only be administered in Italian, Sicilian dialects generally lacking broadly shared written conventions (see Di Caro 2022: Ch. 4);<sup>9</sup> ii) some of the items of specific modules of the BLP forced the final score of the Index towards Italian.<sup>10</sup>

Furthermore, it is important to note that the BLP only provides information within speakers, and not directions of proficiency in one of the two languages between speakers. In other words, we can rely on this measure to quantify how much a speaker is bilingual, but we cannot compare the proficiency of one of two languages between two or among more speakers.

### 3. A quantitative study on bilectal speakers of Sicilian and Italian

#### 3.1. Quantifying the hypotheses

In contemporary linguistic research, there is a discernible shift towards the integration of quantitative and computational methodologies to address theoretical questions. Recently, scholars have observed that an increasing number of analyses have

<sup>8</sup> See Flege et al. (2002), Marian et al. (2007), Gertken et al. (2014), and Gollan et al. (2012) for discussions on the reliability of self-ratings as an assessment method. See, however, Treffers-Daller (2015, 2019) and Solís-Barroso & Stefanich (2019) for some criticisms.

<sup>9</sup> The BLP was generally handed out to participants for them to compile it on their own. The items of the BLP were read aloud by the researcher on request by some older participants. In both cases, participants were not timed.

<sup>10</sup> The most significant one being that regarding the language used when counting. Since counting is an ability that is generally learnt in schools, even those bilectal speakers self-assessed as dominant in Deliano declared to resort to Italian when counting. This is in line with what Poarch et al. (2019: 624-625) highlight for Swabian speakers when asked to report the “years of schooling in Swabian”. Such issues are being addressed by Procentese, Di Caro & Lebani (in prep.), who are working on a new tool, called the ‘Dilalic Language Profile’, specifically designed for the Italo-Romance linguistic landscapes.

embraced innovative approaches, employing statistical tools and computational models to explore the nuances of language (see Van Craenenbroeck & Van Koppen 2023, Merlo & Samo forthcoming for overviews). We operate a study in quantitative computational syntax, where predefined models are defined theoretically, but values are collected by corpus inspection. This paradigm shift underscores a departure from traditional qualitative methods, as researchers increasingly recognize the potential of quantitative techniques in deciphering complex linguistic phenomena. The interplay between linguistic inquiry and computational methods can lead to further research questions and show relevant trends, such as in optionality (Samo & Si 2022). In this work, we explore the methodology for a different dimension of optionality, i.e. the billectal environments.

Two research questions are raised in the introduction, here repeated.

i) can we find asymmetries between billectal speakers in Delia in the selection of the different strategies of indefiniteness marking when speaking Deliano?

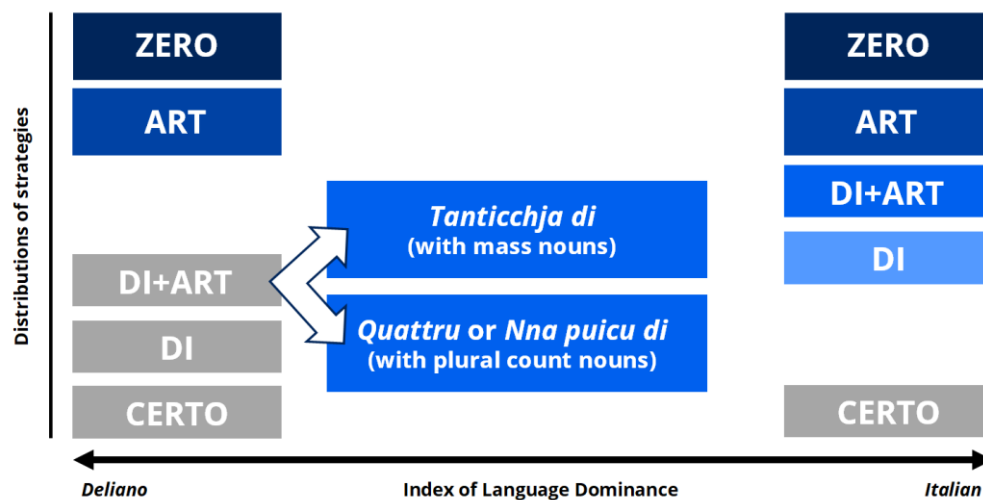
ii) is the Index of Language Dominance (Birdsong et al. 2012) a good predictor of the billectal speakers' selection of the strategies of indefiniteness marking when speaking Deliano?

Both research questions can be explored with quantitative tools. Both questions can therefore be translated into a unique working hypothesis, that should mark clear asymmetries within the billectal population. Let us define  $L_{DEL}$  and  $L_{ITA}$  as the extreme values of the Index of Language Dominance towards, respectively, Deliano and Italian. As a reminder, ZERO/ART are available strategies in both varieties; DI+ART is a productive strategy of Italian, while *Tanticchja* and *Quattru/Nna puicu* are acceptable strategies in Deliano. Let  $S_{DEL}$  and  $S_{ITA}$  be the raw frequencies and distributions in the strategies of Deliano and Italian. If the ILD does represent a good predictor for the billectal speakers' selection of the strategies, then we should expect that the relevant values ( $DEL$  and  $ITA$ ) of the two variables should be correlated. We can then formalize the strongest hypothesis ( $H_1$ ) as follows.

**Table 2.** Hypotheses for Deliano ( $S_{DEL}$ ) and Italian ( $S_{ITA}$ ).

Strategy	$S_{DEL}$	$S_{ITA}$
ZERO/ART	No correlation	No correlation
DI+ART	Negative correlation	Strong positive correlation
<i>Tanticchja</i> , <i>quattru</i> , <i>nna puicu</i>	Positive correlation	Negative correlation

We can visualize such a hypothesis (that we define also as the *strongest* hypothesis) in Figure 2.

**Figure 2.** If language dominance affects indefiniteness strategies

We use standard correlation measures as Pearson's correlation ( $r$ ), Spearman's correlation ( $\rho$ ) and linear regressions. Let us start with the correlations in terms of Pearson's  $r$  and Spearman's  $\rho$ . A value close to 1 indicates a strong positive relationship, meaning that as one variable increases, the other tends to increase as well. Conversely, a value close to  $-1$  indicates a strong negative relationship, implying that as one variable increases, the other tends to decrease. A value of 0 suggests no linear relationship between the variables. We take values higher than 0.5 to be indicators of strong correlations for Pearson's  $r$  and values between 0.40 and 0.60 for moderate Spearman's correlations and higher than 0.60 for stronger correlations. With respect to linear regression, we expect a coefficient of Determination ( $R^2$ ) higher than 0.60. This would represent an alternative hypothesis, which is to be compared to a null hypothesis in which the Index of Language Dominance does not correlate with the syntactic indefiniteness strategies. From a syntactic point of view, our null hypothesis still remains a working hypothesis, suggesting that these strategies are somewhat encoded in syntax, in terms of syntactic transfer or other phenomena, for example of lexical nature. It is worth noting that the null hypothesis might also have an additional outcome. It is possible that the BLP is not a good measure for mapping dialectal data.

Following Samo & Merlo (2019), for many of the statistical tests we use the binomial test. A binomial test is a statistical test used to determine if the proportion of successes in a sample significantly differs from a known or hypothesized proportion. The test assesses whether the observed proportion is significantly different from the expected proportion under a specific hypothesis (in our case, the lack of grammar, meaning chance level). In our case, the binomial is the presence of a given strategy over the others (e.g., 5 strategies equals 20%).

### 3.2. Materials and methods

We based our study on a set of 1005 observations extracted from one of the sub-corpora of the Corpus of the dialect of Delia (or CorDel; Di Caro in prep.), a corpus of recorded collection of linguistic data from interviews, quasi-experiments and

experimental situations. The data of the sub-corpus discussed here were collected between 2020 and 2021. The CorDel aims at documenting the Italo-Romance variety of Delia mainly (but not exclusively) in its oral dimension, as it is actually spoken by the community (of both younger and older speakers), i.e. with all the naturally occurring mixing with Regional and Standard Italian (especially in the lexicon; see Di Caro 2022: §3.2.2.1) one may expect in this scenario (see, e.g., Cerruti & Regis 2020). The nature of this corpus and in particular of this sub-corpus is that we can also retrieve the linguistic profile of the 24 speakers producing the given utterance in two specific experimental settings (elicited and semi-spontaneous production).

Therefore, this subsection is to be split into different subsections. The profile of the participants (§3.2.1.), the two experimental situations (§3.2.2.) and, finally, the methodology of the annotation (§3.2.3.).

### 3.2.1. Participants

A group of 24 billectal speakers was selected from the 56 participants in the CorDel (Di Caro in prep.), who contributed the ‘shopping’ text genre (see Di Caro 2023: §3.1 for details). The group of participants, with an age ranging from 19 to 72, features 12 male and 12 female subjects. Care was taken to balance age groups and gender. Each participant was given a pseudonym.<sup>11</sup> Table 3 shows detailed data, in particular the participants’ Index of Language Dominance, while Figure 3 provides an overview in terms of age and gender distribution.

**Table 3.** An overview of the participants with their sociolinguistic data

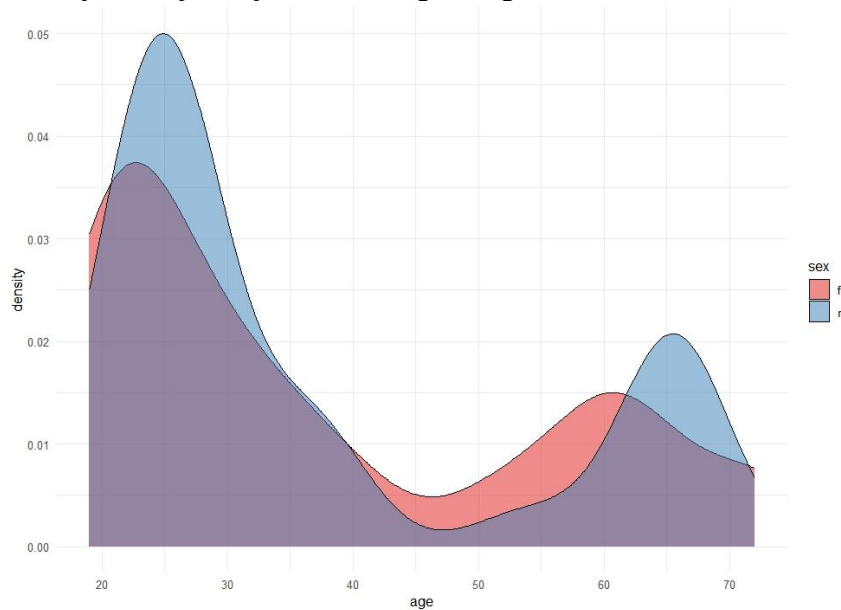
N	Pseudonym	Gender	Age	ILD
1	Ilenia	F	27	+49
2	Calogero	M	26	+38
3	Marianna	F	32	+75
4	Fiorella	F	21	+71
5	Filomena	F	52	+27
6	Rosario	M	23	+22
7	Maria	F	72	+54
8	Stella	F	60	+23
9	Nicoletta	F	39	+74

<sup>11</sup> The participants’ level of education is among the sociolinguistic data collected by the customized BLP for this study (see Di Caro 2023). However, it was not taken into account in the present analyses, since only one participant was below the high school level.

---

10	Gaspare	M	54	+95
11	Palmina	F	62	-10
12	Raffaele	M	33	-28
13	Linda	F	19	+82
14	Gioacchino	M	65	+79
15	Teresa	F	22	+52
16	Pasquale	M	21	+60
17	Geraldina	F	21	+74
18	Davide	M	37	+96
19	Vincenzo	M	26	+103
20	Alberto	M	25	+11
21	Federico	M	28	+64
22	Nicola	M	66	+20
23	Francesco	M	24	-80
24	Paola	F	25	+12

Source: Di Caro (2023: 11)

**Figure 3.** Density of the participants across age and gender.

The ILD was calculated following Birdsong et al.’s (2012) BLP. All details and data are discussed in Di Caro (2023).

### 3.2.2. *Experimental situations: elicited and semi-spontaneous*

The material that allowed the recording of the oral production for the CorDel genre called ‘shopping’, already used in Di Caro’s (2023) preliminary analysis, is a video, shot in the perspective of an Italian speaking young woman during a shopping session at a supermarket. This means that she is generally not shown but the participants can only hear her voice as the products she is purchasing or just commenting upon are shown. Since the excerpt is almost three-minute long, the participants were provided with a list of prompt questions such as “What does the young woman buy in the fruit and vegetable aisle?” or “Does she buy any meat?”, which, however, were not to be considered as compulsory points to answer but rather as a roadmap of the shopping session. Personal comments were also encouraged.

The participants were asked to accomplish two tasks (described in (7)). We re-labelled them here respectively as elicited production (cf. (7a)) and semi-spontaneous production (cf. (7b)):

- (7) a. *Elicited Production*: the participants were asked to describe in Deliano the three-minute videotape excerpt of the young woman’s shopping session.
- b. *Semi-spontaneous production*: the experimenter asks the participants to talk about their own shopping in Deliano in two different contexts: (i) their personal preferences when doing the shopping; (ii) what they bought the last time they went to do the shopping. During this task, participants had the freedom to describe their shopping routines in any order they preferred. They were simply encouraged to provide as much detail as possible and to include, when applicable, the buying of items beyond the main products like water, meat, fruits, and vegetables, such as home and intimate care products.

All the interactions were recorded and manually transcribed.

### 3.2.3. Encoding the linguistic features

Beyond the available data from the CorDel, we semi-automatically annotated the data with the human validation by one of the authors. Following standard practice in quantitative approaches in generative grammar, we encoded each sentence as a vector of features following a theory-counting model (Merlo & Samo forthcoming): each feature of the phenomenon is annotated following a nominal value and the frequency of the phenomenon represents a numeral value. Every feature represents a dimension.

**Figure 4.** A screenshot of the complete data-frame showing all the variables (for 1005 observations)

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
id	pseudonym	age	sex	ucati	job	dominance	delianity	sentence	verb	lex_item	categ	indef	class	list	syntax	polar	type	tense	san_sp	scope	video
9	ilenia	27	f	2	waitress	49	2	Allura, chista parti mischina e va a fa la spisa [...] si	pick	bread	food	art	mass	no	obj	pos	epis	pres	no	no	yes
9	ilenia	27	f	2	waitress	49	2	piglia li mandarina	pick	tangerine	fruit	art	plur	yes	obj	pos	epis	pres	no	no	yes
9	ilenia	27	f	2	waitress	49	2	l'aranci	pick	orange	fruit	art	plur	yes	obj	pos	epis	pres	no	no	yes
9	ilenia	27	f	2	waitress	49	2	piglia quattru mandarina	pick	tangerine	fruit	number	plur	yes	obj	pos	epis	pres	no	no	yes
9	ilenia	27	f	2	waitress	49	2	du aranci	pick	orange	fruit	number	plur	yes	obj	pos	epis	pres	no	no	yes
9	ilenia	27	f	2	waitress	49	2	e quarchi pera	pick	pear	fruit	quarchi	sing	yes	obj	pos	epis	pres	no	no	yes
9	ilenia	27	f	2	waitress	49	2	e quarchi mela	pick	apple	fruit	quarchi	sing	yes	obj	pos	epis	pres	no	no	yes
9	ilenia	27	f	2	waitress	49	2	'Nzumma, poi piglia sei zucchini	pick	courgette	veg	number	plur	yes	obj	pos	epis	pres	no	no	yes
9	ilenia	27	f	2	waitress	49	2	li pipiruna	pick	bell_pep	veg	art	plur	yes	obj	pos	epis	pres	no	no	yes
9	ilenia	27	f	2	waitress	49	2	pipiruna nni piglia tri	pick	bell_pep	veg	number	plur	no	disloc	pos	epis	pres	no	no	yes
9	ilenia	27	f	2	waitress	49	2	un paru di cipuddri	pick	onion	veg	paru	plur	yes	obj	pos	epis	pres	no	no	yes
9	ilenia	27	f	2	waitress	49	2	tanticchia d'agliu	pick	garlic	veg	tanticchia	sing	yes	obj	pos	epis	pres	no	no	yes
9	ilenia	27	f	2	waitress	49	2	li finicichju	pick	fennel	veg	art	sing	yes	obj	pos	epis	pres	no	no	yes
9	ilenia	27	f	2	waitress	49	2	li limuna	pick	lemon	fruit	art	plur	yes	obj	pos	epis	pres	no	no	yes
9	ilenia	27	f	2	waitress	49	2	e li kiwi	pick	kiwi	fruit	art	plur	yes	obj	pos	epis	pres	no	no	yes
9	ilenia	27	f	2	waitress	49	2	limuna nni piglia un sacchetti	pick	lemon	fruit	zero	plur	no	disloc	pos	epis	pres	no	no	yes
9	ilenia	27	f	2	waitress	49	2	Kiwi nni piglia du.	pick	egg	fruit	zero	plur	no	disloc	pos	epis	pres	no	no	yes
9	ilenia	27	f	2	waitress	49	2	Mischina poi chista piglia sei ova	pick	egg	food	number	plur	yes	obj	pos	epis	pres	no	no	yes
9	ilenia	27	f	2	waitress	49	2	piglia lu lattì a la cannella	pick	milk	drink	art	mass	yes	obj	pos	epis	pres	no	no	yes
9	ilenia	27	f	2	waitress	49	2	e un litru di lattì normali	pick	milk	drink	quantif	mass	yes	obj	pos	epis	pres	no	no	yes
9	ilenia	27	f	2	waitress	49	2	Carni 'nveci nun nni piglia	pick	meat	meat	zero	mass	no	disloc	neg	epis	pres	no	no	yes
9	ilenia	27	f	2	waitress	49	2	carni nun nni piglià	pick	meat	meat	zero	mass	no	disloc	neg	epis	past	no	no	yes
9	ilenia	27	f	2	waitress	49	2	Ma l'urtima vota per esempju accattavu li zucchini	buy	courgette	veg	art	plur	no	obj	pos	epis	past	no	no	no
9	ilenia	27	f	2	waitress	49	2	accattavu li cipuddri	buy	onion	veg	art	plur	yes	obj	pos	epis	past	no	no	no
9	ilenia	27	f	2	waitress	49	2	li patati	buy	potato	veg	art	plur	yes	obj	pos	epis	past	no	no	no
9	ilenia	27	f	2	waitress	49	2	poi di carni accattavu la sansizza	buy	sausage	meat	art	mass	yes	obj	pos	epis	past	no	no	no

The sentences extracted from the corpus are represented as vectors of features. The features are listed in (8):

- (8) a. ART, i.e. the definite article used as an indefinite determiner. (n) *pigliavu lu pani* (Ilenia, 27, +49).
- b. GramCard: (i) DU, i.e. the grammaticalized low cardinal ‘two’; (ii) QUATTRU, i.e. the grammaticalized low cardinal ‘four’.
- c. Number, i.e. any enumeration with proper cardinal numbers.
- d. PART, i.e. the partitive article DI+ART.
- e. NNA PUICU, i.e. the pseudo-partitive ‘a bit of’ associated with plural count nouns.
- f. QUANTIF, i.e. any quantification such as ‘a bag of’, ‘a kilogram of’, as well as other configurations. (n) *e di verdura, piglia nna riti di zucchini* (Calogero, 26, +38).
- g. QUARCHI, i.e. the indefinite ‘some’.
- h. TANTICCHJA, i.e. the pseudo-partitive ‘a bit of’ associated with mass nouns.

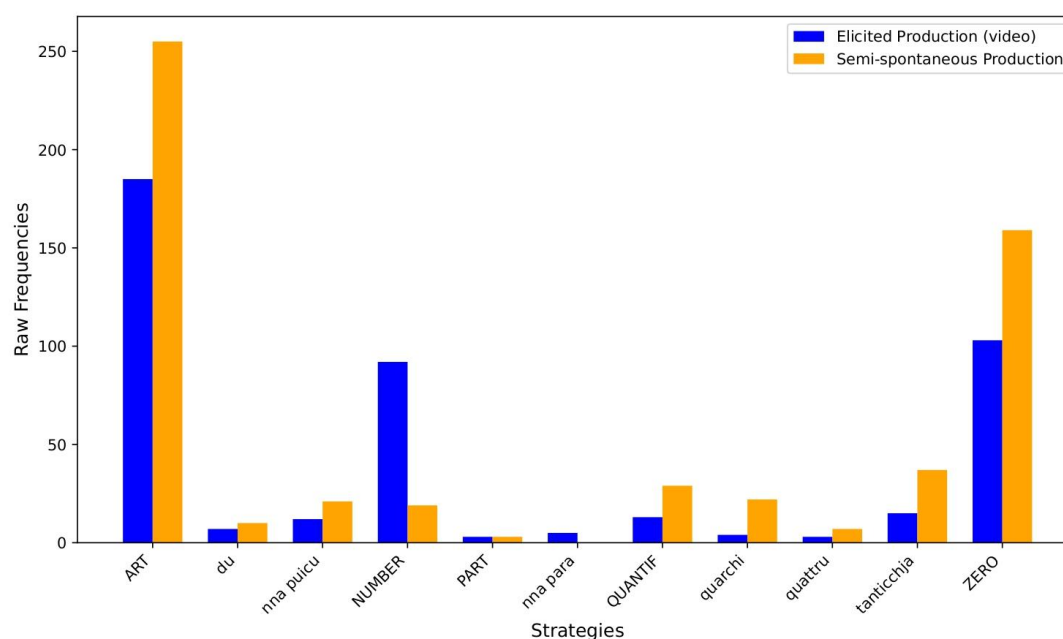
- i. ZERO, i.e. the zero determiner.
- j. CONTEXT, i.e. a higher context that can contain other labels, such as NNA PARA (i.e. any occurrence of ‘a pair of’).

We can now turn to the results of the study.

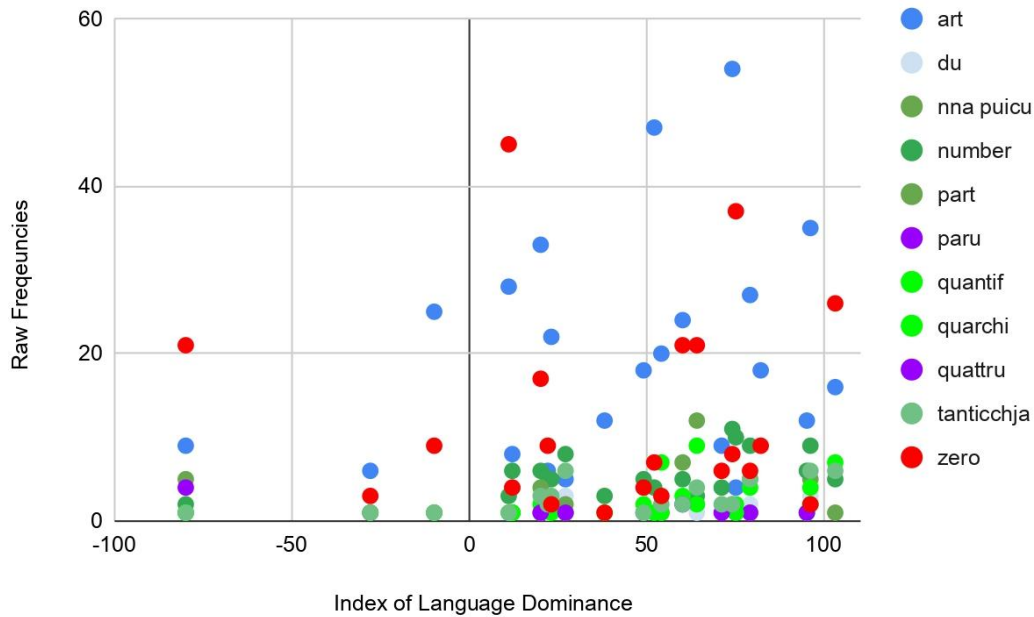
### 3.3. Results

Before focussing our attention on the research questions (Figure 7), we would like to provide basic descriptive statistics on the data and the strategies. The distribution of the strategies across experimental situations are given in Figure 5 and the distributions of the indefiniteness marking strategies according to the ILD are given in Figure 6.

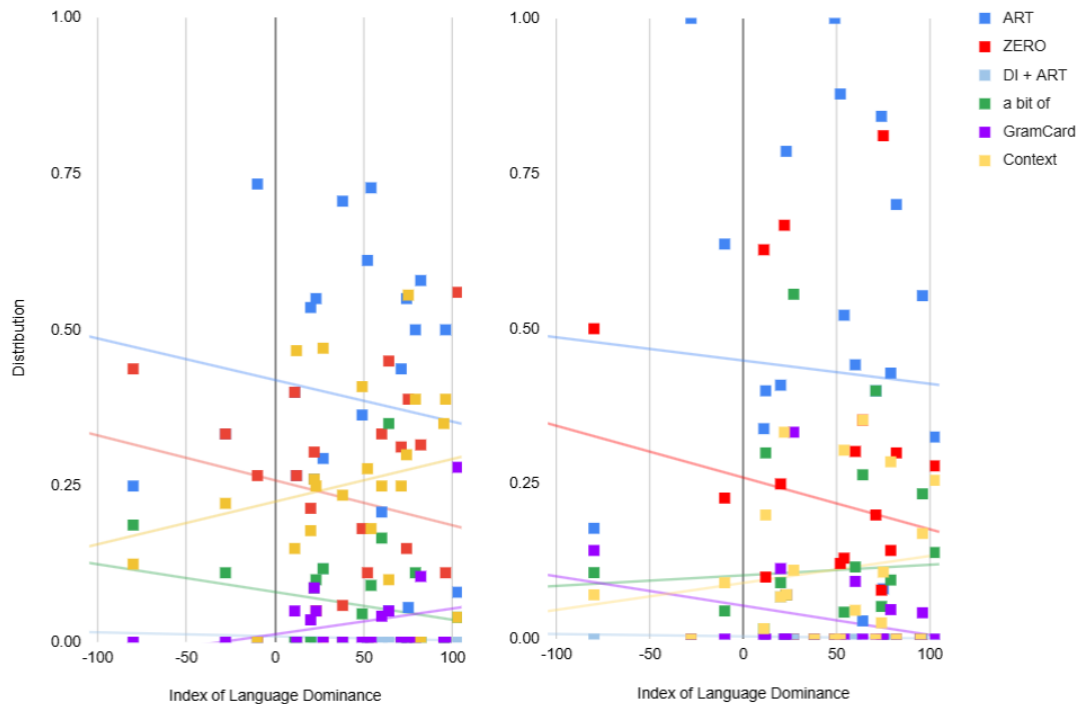
**Figure 5.** Distributions of indefinites/quantifications in the corpus



**Figure 6.** Distributions of the indefiniteness marking strategies according to the Index of Language Dominance



**Figure 7.** The results in the elicited (left) and semi-spontaneous (right) production.



As Figure 7 shows, we observed a minimized usage of DI+ART in both experimental contexts (binomial test  $p = 0.5$ ;  $z = -13.27$ ;  $p < .00000001$ ). DI+ART is a felicitous strategy in Italian, but bilingual speakers do not seem to exploit this strategy. We also detected an overall preference for the production of sentences with ART (43.8%;  $z = 121.34$ ,  $p < .00000001$ ) over ZERO (26.2%;  $z = 162.04$ ,  $p < .00000001$ ),

which is contrasted with a tendency for preferring ZERO (47.7%;  $z = 23.89$ ,  $p < .000001$ ) in our data collected for the extreme (-80) of our ILD. According to our data points, no correlation is found between the ILD and the distributions. Finally, some words with respect to the experimental settings. We tentatively envisage effects of priming of the numerals with respect to grammaticalized cardinals (opposite weak correlations,  $r = -0.27$  in the semi-spontaneous condition;  $r = 0.28$  in the elicited condition).

#### 4. Discussion

It is important to remark that the analyzed data belong to a curated small-sized dataset with all the relevant limitations, such as exhaustiveness and control for additional variables. Despite this, we believe that our quantitative data show clear and statistically relevant (lack of) trends and correlations between an Index of Language Dominance and the relevant syntactic strategies. From a qualitative point of view, our data are novel and thus allow for typical analysis on naturally occurring examples.

General lack of DI+ART in Deliano, as predicted in Giusti (2021), is not surprising. There are only six occurrences of indefinite DI+ART in the corpus, produced by two participants (pseudonymized as Filomena and Nicola). They are both over 50 and display an ILD of, respectively, +27 and +20, which means they are balanced bilinguals. Di Caro (2023: 14) suggests an interference effect caused by Italian, by highlighting that i) some of the six utterances containing the occurrences of DI+ART (shown in Table 3) are mixed with Italian, namely (e) and (f). They display the Italian DI+ART forms *dei* ‘of.the.PL.M’ and *degli* ‘of.the.PL.M’ and they both display some hesitation by the speaker, who in (f) alternates the Deliano DI+ART with its Italian counterpart.<sup>12</sup>

**Table 4.** Naturally occurring examples of DI+ART in the corpus

Participant [Gender, Age, ILD]		Naturally occurring example
Filomena [F, 52, +27]	a.	<i>Sta carusa [...] vidi di li mandarina.</i> ‘This girl sees some tangerines.’
	b.	<i>[...] s'accatta [...] di li peperuna.</i> ‘She buys (herself) some bell peppers.’
Nicola [M, 66, +20]	c.	<i>[...] picchi piglià di la frutta.</i> ‘because she picked some fruits.’
	d.	<i>[...] ppi -gghjiri a -ppigliari di la pasta.</i> ‘to go and fetch some pasta.’

<sup>12</sup> Naturally, the interplay between the two varieties is more intricate and requires further analysis. For instance, an anonymous reviewer points out that the quantified products in (a) and (b) of Table 4 featuring DI+ART still display the typical Sicilian neuter plural ending -a (cf. Sic. *mandarina* vs It. *mandarini* ‘tangerines’ and Sic. *peperuna* vs It. *peperoni*).

	e.	[...] <i>cci sunnu propriu di... dei reparti di l'affettati.</i> 'there are specific aisles for deli meats.'
	f.	<i>Accattu molte delle volte anche di li... degli... di l'hamburger.</i> 'I often buy some hamburgers.'

Source: adapted from Di Caro (2023: 14)

We further add that in (a-d) all the nouns the DI+ART forms refer to are close or identical to their Italian counterparts (cf. Italian *mandarini* 'tangerines', *peperoni* 'bell peppers',<sup>13</sup> *frutta* 'fruits', and *pasta* 'pasta'), which could have fostered the use of this strategy.

#### 4.1. Limitations of this study

The data collection modality for this study presents some advantages. For example, the occurrences and the collocations are not forced by the stimuli of an acceptability judgement task, a richer lexicon can be elicited, it is possible to control for the variables in the speakers' productions, and it is possible to modelize and understand which factors make the relevant indefiniteness marking strategies emerge. Nevertheless, different weaknesses can be identified that can be taken into account for future, more fine-tuned, studies. On the one hand, in such a small corpus, some of the indefiniteness marking strategies identified in previous fieldwork simply never occurred and are likely not to occur even in a wider set of observations. On the other hand, the type of videotape extract the participants were asked to comment on could influence the indefiniteness marking choice. Notably, the use of proper low cardinals 'two' and 'four' in the shopping list favors the use of grammaticalized 'two' and 'four' with respect to other available strategies.

A final note must be dedicated to the fact that the data were annotated mainly manually. The semi-automatic annotation of data still represents a challenging aspect in collecting dialectal data.

#### 5. Conclusions and open questions for future research

In this paper, we have tried to apply quantitative methods (i.e. distributions and correlations) to a carefully curated dataset of biletal speakers in a study on what strategies of indefiniteness marking the participants of the study in Di Caro (2023) resort to in (i) a specific sociolinguistic scenario characterizing most Italo-Romance varieties – which Berruto (1987) proposed to call 'dilalia' (see also Di Caro 2022: Ch. 3); (ii) a specific narrative context, namely the description of a third person's shopping session and of each participant's shopping routine.

Based on the research project started by Cardinaletti & Giusti (2018) on weak indefinites in Italo-Romance, we have confirmed previous research stating that bare DI, DI+ART and CERTO (see Section 2.1) are not productive strategies in Deliano.

According to the sample considered here, the Index of Language Dominance, as calculated from Birdsong et al.'s (2012) Bilingual Language Profile, seems not to

<sup>13</sup> In the case of *peperuna* (Table 4, (b)), note that an older Deliano version of 'pepper bells' is *pipi*. Note also that the contemporary Deliano term for 'pepper bells' is *pipiruna*. So, in this case, the Italian form *peperoni* shows its influence.

affect the indefiniteness strategies chosen. In this spirit, the empirical evidence discussed in this study reinforces the idea behind Procentese, Di Caro and Lebani's (2022, in prep.) idea that a more fine-grained tool specifically dedicated to Italo-Romance dilalic scenarios (i.e., the 'Dilalic Language Profile', or DLP) is needed.

Finally, since the data analyzed in this study are taken from a specific section (i.e. the shopping session) of the corpus of oral and written dialect spoken in Delia (CorDel), other text genres of that corpus not specifically designed to elicit indefinite expressions could reveal other trends in the choice of the indefiniteness strategies or confirm those emerged in our analysis. This could also overcome the lexical limitations discussed in Section 4.3.

## Acknowledgments

Previous versions of this paper were presented in a webinar delivered for the Cambridge Linguistic Society on 01/06/2023, and at the 17th Cambridge Italian Dialect Syntax-Morphology Meeting, organized by the University of Zurich and held in Zurich on 12/09/2023. We would like to thank the organizers, the audiences of the two events, and three anonymous reviewers for their stimulating remarks. Vincenzo Nicolò Di Caro acknowledges the European Union NextGenerationEU PRIN 2022 NRRP project 'Contact-induced change and sociolinguistics: an experimental study on the Gallo-Italic dialects of Sicily' (CUP E53D23019660001 – ID P2022YWS8T; PI: Alessandro De Angelis, University of Messina). While this paper is the outcome of joint work by the authors, for academic purposes, Vincenzo Nicolò Di Caro is responsible for sections 1, 2, 3.2; Giuseppe Samo for sections 3.1, 3.3, 4 and 5. The usual disclaimer applies.

## References

- Abney, Steven. 1987. *The English noun phrase in its sentential aspect*. Ph.D. thesis, Massachusetts Institute of Technology.
- Berruto, Gaetano. 1974. *Piemonte e Valle d'Aosta*. Pisa: Pacini.
- Berruto, Gaetano. 1987. Lingua, dialetto, diglossia, dilalia. In G. Holtus & J. Kramer (eds), *Romania et Slavia adriatica. Festschrift für Žarko Muljačić*, 57-81. Hamburg: Helmut Buske.
- Berruto, Gaetano. 1989. On the Typology of Linguistic Repertoires. In U. Ammon (ed.), *Status and Function of Languages and Language Varieties*, 552-569. Berlin/New York: de Gruyter. <https://doi.org/10.1515/9783110860252.552>
- Berruto, Gaetano. 1993. Le varietà del repertorio. In A. A. Sobrero (ed.), *Introduzione all'italiano contemporaneo. La variazione e gli usi*, 3-36. Roma/Bari: Laterza.
- Birdsong, David. 2014. Dominance and age in bilingualism. *Applied Linguistics* 35: 374-392. <https://doi.org/10.1093/applin/amu031>

- 
- Birdsong, David. 2015. Dominance in Bilingualism: Foundations of Measurement, with Insights from the Study of Handedness. In C. Silva-Corvalán & J. Treffers-Daller (eds), *Language Dominance in Bilinguals: Issues of Measurement and Operationalization*, 85-105. Cambridge: Cambridge University Press. <https://doi.org/10.1017/cbo9781107375345.005>
- Birdsong, David, Gertken, Libby M., & Mark Amengual. 2012. *Bilingual Language Profile. An Easy-to-Use Instrument to Assess Bilingualism*. COERLL, University of Texas at Austin. <https://sites.la.utexas.edu/bilingual/>
- Brasoveanu, Adrian, & Donka F. Farkas. 2016. Indefinites. In M. Aloni & P. Dekker (eds), *The Cambridge Handbook of Formal Semantics*, 238-266. Cambridge: Cambridge University Press.
- Cardinaletti, Anna, & Giuliana Giusti. 2018. Indefinite Determiners: Variation and Optionality in Italo-Romance. In R. D'Alessandro & D. Pescarini (eds), *Advances in Italian Dialectology*, 135-161. Leiden/Boston: Brill.
- Cardinaletti, Anna, & Giuliana Giusti. 2020. Quantified expressions and quantitative clitics. *Linguistics* 58(3): 679-712. <https://doi.org/10.1515/ling-2020-0081>
- Cerruti, Massimo, & Riccardo Regis. 2020. *Italiano e dialetto*. Roma: Carocci.
- Coseriu, Eugenio. 1980. "Historische Sprache" und "Dialekt". In J. Göschel, P. Ivić, & K. Kehr (eds), *Dialekt und Dialektologie. Ergebnisse des Internationalen Symposiums "Zur Theorie des Dialekts" (Marburg/Lahn, 5.-10. September 1977)*, 106-122. Wiesbaden: Steiner.
- De Angelis, Alessandro. 2023. The Strange Case of the Gallo-Italic Dialects of Sicily: Preservation and Innovation in Contact-Induced Change. *Languages* 8(3): 163. <https://doi.org/10.3390/languages8030163>
- De Mauro, Tullio. 1972. *Storia linguistica dell'Italia unita*. Bari: Laterza.
- Di Caro, Vincenzo Nicolò. 2022. *Il bilinguismo italiano-dialetto in Sicilia. Profilo sociolinguistico, nuove realtà comunicative e prospettive didattiche*. Venezia: Edizioni Ca' Foscari. <http://doi.org/10.30687/978-88-6969-649-7>
- Di Caro, Vincenzo Nicolò. 2023. Strategies of Indefiniteness Marking in Central Sicilian. Evidence from the Dialect of Delia. *Languages* 8(4): 227. <https://doi.org/10.3390/languages8040227>
- Di Caro, Vincenzo Nicolò. In prep. *CorDel (Corpus del dialetto deliano)*.
- Ferguson, Charles A. 1959. Diglossia. *Word* 15: 325-340.
- Ferguson, Charles A. 1991. Diglossia Revisited. *Southwest Journal of Linguistics* 10: 214-234.

Flege, James E., Mackay, Ian R. A., & Thorsten Piske. 2002. Assessing bilingual dominance. *Applied Psycholinguistics* 23(4): 567-598. <https://doi.org/10.1017/S0142716402004046>

Gertken, Libby M., Amengual, Mark, & David Birdsong. 2014. Assessing Language Dominance with the Bilingual Language Profile. In P. Leclercq, A. Edmonds, & H. Hilton (eds), *Measuring L2 Proficiency: Perspectives from SLA*, 208-225. Bristol: Multilingual Matters. <https://doi.org/10.21832/9781783092291-014>

Giammarco, Ernesto. 1979. *Abruzzo*. Pisa: Pacini.

Giusti, Giuliana. 2002. The Functional Structure of Determiners. A Bare Structure Approach. In G. Cinque (ed.), *Functional Structure in DP and IP. The Cartography of Syntactic Structures*, 54-90. Oxford/New York: Oxford University Press.

Giusti, Giuliana. 2015. *Nominal Syntax at the Interfaces*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Giusti, Giuliana. 2021. A Protocol for Indefinite Determiners in Italian and Italo-Romance. In T. Ihsane (ed.), *Disentangling bare nouns and nominals introduced by a partitive article*, 262-300. Leiden: Brill.

Gollan, Tamar H., Weissberger, Gali H., Runnqvist, Elin, Montoya, Rosa I., & Cynthia M. Cera. 2012. Self-ratings of Spoken Language Dominance: A Multi-Lingual Naming Test (MINT) and Preliminary Norms for Young and Aging Spanish-English Bilinguals. *Bilingualism: Language and Cognition* 15(3): 594-615. <https://doi.org/10.1017/S1366728911000332>

Grosjean, François. 2001. The Bilingual's Language Modes. In J. L. Nicol (ed.), *Explaining linguistics. One mind, two languages: Bilingual language processing*, 1-22. Oxford: Blackwell Publishing.

Kupisch, Tanja, Arona, Sebastiano, Besler, Alexandra, Cruschina, Silvio, Ferin, Maria, Gyllstad, Henrik, & Ilaria Venagli (2023). LexSIC: A quick vocabulary test for Sicilian. *Isogloss. Open Journal of Romance Linguistics* 9(1): 1-24. <https://doi.org/10.5565/rev/isogloss.302>

Ledgeway, Adam. 2009. *Grammatica diacronica del napoletano*. Tübingen: Niemeyer.

Lebani, Gianluca E., & Giuliana Giusti. 2022. Indefinite determiners in two northern Italian dialects. A quantitative approach. In O. Matushansky, L. Roussarie, M. Russo, E. Soare, & S. Wauquier (eds), *RLLT17*. Special issue of *Isogloss. Open Journal of Romance Linguistics* 8(2)/18: 1-19.

Maiden, Martin, & Mair Parry. 1997. *The dialects of Italy*. London; New York: Routledge.

- 
- Marian, Viorica, Blumenfeld, Henrike K., & Margarita Kaushanskaya. 2007. The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing Language Profiles in Bilinguals and Multilinguals. *Journal of Speech, Language, and Hearing Research*, 50(4): 940-967. [https://doi.org/10.1044/1092-4388\(2007/067\)](https://doi.org/10.1044/1092-4388(2007/067))
- McNally, Louise. 2020. Strong and Weak Nominals. In D. Gutzmann, L. Matthewson, C. Meier, H. Rullman, & T. E. Zimmerman (eds), *The Wiley Blackwell Companion to Semantics*. Oxford: Wiley.
- Merlo, Paola, & Giuseppe Samo. 2022. Exploring T3 languages with Quantitative Computational Syntax. *Theoretical Linguistics* 48(1-2): 73-83.
- Merlo, Paola, & Giuseppe Samo. Forthcoming. Generative Computational Modelling. In E. Leivada, & K. K. Grohmann (eds), *The Cambridge Handbook of Minimalism and Its Applications*. Cambridge: Cambridge University Press.
- Milsark, Gary. 1977. Peculiarities of the Existential Construction in English. *Linguistic Analysis* 3: 1-29.
- Pescarini, Diego. 2022. A Quantitative Approach to Microvariation: Negative Marking in Central Romance. *Languages* 7(2): 87.
- Poarch, Gregory J., Vanhove, Jan, & Raphael Berthele. 2019. The effect of bidialectalism on executive function. *International Journal of Bilingualism* 23(2): 612-628. <https://doi.org/10.1177/1367006918763132>
- Procentese, Cristina, Di Caro, Vincenzo Nicolò, & Gianluca E. Lebani. 2022. Assessing language dominance in Italo-Romance bilectals with the Dilalic Language Profile: a preliminary pilot study. Talk delivered at the *International conference on bilingualism with local languages (LABiC) 2022: Language Attitudes and Bi(dia)lectal Competence* on 12/09/2022 at Ca' Foscari University of Venice.
- Procentese, Cristina, Di Caro, Vincenzo Nicolò, & Gianluca E. Lebani. In prep. *The Dilalic Language Profile*. Ms., Ca' Foscari University of Venice.
- Rohlf, Gerhard. 1968. *Grammatica storica della lingua italiana e dei suoi dialetti*, vol. 2 *Morfologia*. Torino: Einaudi.
- Rohlf, Gerhard. 1969. *Grammatica storica della lingua italiana e dei suoi dialetti*, vol. 3 *Sintassi e formazione delle parole*. Torino: Einaudi.
- Rowe, Charley, & Kleanthes K. Grohmann. 2013. Discrete Bilectalism: Towards Co-Overt Prestige and Diglossic Shift in Cyprus. *International Journal of the Sociology of Language* 224: 119-142. <https://doi.org/10.1515/ijsl-2013-0058>

---

Samo, Giuseppe, & Fuzhen Si. 2022. Optionality of 的De in Chinese Possessive Structures: A Quantitative Study. *Qulso* 8: 37-53. <http://dx.doi.org/10.13128/QULSO-2421-7220-13602>

Samo, Giuseppe, & Paola Merlo. 2019. Intervention effects in object relatives in English and Italian: a study in quantitative computational syntax. In X. Chen, & R. Ferrer-i-Cancho (eds), *Proceedings of the First Workshop on Quantitative Syntax (Quasy, SyntaxFest 2019)*, 46-56. Paris: Association for Computational Linguistics. <https://aclanthology.org/W19-7906>

Schwarz, Bernhard. 2011. Long distance indefinites and choice functions. *Language and Linguistics Compass* 5: 880-897.

Selkirk, Elisabeth. 1977. Some remarks on noun phrase structure. In P. Culicover, T. Wasow, & A. Akmajian (eds), *Formal syntax*, 285-325. New York: Academic Press.

Solís-Barroso, Cecilia, & Sara Stefanich. 2019. Measuring Language Dominance in Early Spanish/English Bilinguals. *Languages* 4(3): 62. <https://doi.org/10.3390/languages4030062>

Treffers-Daller, Jeanine. 2015. Language dominance: The construct, its measurement, and operationalization. In C. Silva-Corvalán, & J. Treffers-Daller (eds), *Language Dominance in Bilinguals: Issues of Measurement and Operationalization*, 235-265. Cambridge: Cambridge University Press.

Treffers-Daller, Jeanine. 2019. What Defines Language Dominance in Bilinguals? *Annual Review of Linguistics* 5(1): 375-393. <https://doi.org/10.1146/annurev-linguistics-011817-045554>

Van Craenenbroeck, Jeroen, & Marjo Van Koppen. 2023. Quantitative approaches to syntactic variation. Ms., KU Leuven/Meertens Institute.