J. Linguistics **61** (2025), 397–437.
doi:10.1017/S0022226723000282

L'é ciaro che se dise cusì. On Change in the System of Expletive Subject Clitics in Opitergino

CHIARA ZANINI
University of Zurich
GRETA BATTISTELLA
University of Zurich
FRANCESCO GARDANI
University of Zurich

(Received 20 April 2023; revised 10 September 2023; Accepted 22 September 2023)

Expletive subject clitics (ESCs) are pronominal elements that occur in impersonal contexts with which no individual reference is associated. Their presence strikingly distinguishes northern Italo-Romance varieties from standard Italian. We target this structural incongruence by studying the occurrence of ESCs in present-day Opitergino, a virtually unstudied Venetan variety. We explore the question of whether, in the wake of a profound transformation in the sociolinguistic environment that occurred between the first half of the 20th century and early 2020 years, the contact between Opitergino and now-dominant Italian has induced change in the Opitergino ESC system. To test whether change has occurred and to what extent, we compare the results of an extensive online survey we conducted in 2022 with the baseline rules we extracted from speakers born before 1942. We observe that while the system is overall stable, a thread of change is ongoing and manifests in (a) rule weakening in declaratives and (b) erosion of the obligatoriness of ESCs in interrogatives. We argue that this change is likely to be an effect of contact, resulting in structural convergence but not in loss, and affected the part of the ESC system that features more optionality, namely, the domain of declarative clauses.

KEYWORDS: clitics, dialect of Oderzo/Opitergino, expletive subject clitics, Italo-Romance, optionality

1. INTRODUCTION

Opitergino is a northern Italo-Romance (NIR) variety spoken in Oderzo (province of Treviso), a municipality of about 20,000 inhabitants located in the Veneto plain. One of the most salient features of NIR varieties in the Italo-Romance landscape is the existence of subject clitics, that is, unstressed pronominal particles that

obligatorily attach to the finite verb form and double the lexical subject whenever this is overtly expressed (Rohlfs 1968; Vanelli 1987; Brandi & Cordin 1989). NIR subject clitics, as found in Trevigiano (1a–b),[1] do not exist in standard Italian (cf. 1c) and in almost all central and southern Italo-Romance varieties.[2]

(1) (a) *el=parla italian* (Trevigiano)
3M.SG=speaks Italian
'He speaks Italian'.
(b) **Ø parla italian*
speaks Italian
'She/he speaks Italian.'
(c) *Ø parla italiano* (standard Italian)
speaks Italian
'She/he speaks Italian.'

Some NIR varieties also have expletive subject clitics (henceforth, ESCs), which are pronominal elements that occur in impersonal contexts with which no individual reference is associated (Manzini & Savoia 2005, I: 162), as in the Lombard variety of Monno (2a). Again, in these contexts, standard Italian rules out a phrasal subject (2b).

(2) (a) *el=plof* (variety of Monno)
ESC=rains
'It's raining.'
(b) *Ø piove* (standard Italian)
rains
'It's raining.'

The presence of ESCs is a striking structural difference between NIR varieties and standard Italian. In Italian, a subject cannot be expressed in impersonal constructions, such as (2b) (**LUI piove*) (it can only be realized as a strong pronoun under certain pragmatic conditions, such as under contrastive focalization, e.g. *LUI parla italiano* 'it's HIM that speaks Italian').

In this paper, we target this structural incongruence by studying the occurrence of ESCs in present-day Opitergino, a virtually unstudied Venetan (thus, NIR) variety. We explore the question of whether, in the wake of a profound transformation in the sociolinguistic environment that occurred between the first half of the 20th century and early 2020 years, the contact between Opitergino and now-dominant Italian has induced change in the Opitergino ESC system.

[1] If not otherwise indicated, the abbreviations in the glosses follow the Leipzig Glossing Rules (https://www.eva.mpg.de/lingua/pdf/Glossing-Rules.pdf).

[2] In NIR varieties, subject clitics are organized in paradigms that can be complete or display different degrees of defectiveness: When just one person is attested, this is usually the third or second person (cf. Manzini & Savoia 2005, I: 69–121; Renzi & Vanelli 1983).

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

The paper is organized as follows. Section 2 provides background information on ESCs in NIR varieties (Section 2.1) and the sociolinguistic change affecting the use of Italo-Romance varieties starting from the 1950s (Section 2.2). Section 3 presents the research questions and hypotheses (Section 3.1), the baseline ('conservative' Opitergino) (Section 3.2), the study's design (Section 3.3), and the methods used for data analysis (Section 3.4). Section 4 presents the results of the survey. Section 5 analyses the results and evaluates the hypotheses. Section 6 concludes the paper.

2. Background

The present study focuses on (a) ESCs because research on NIR has failed to produce fine-grained investigations into single varieties and to closely investigate the role of language contact in change (Section 2.1); (b) Opitergino because of the sociolinguistic peculiarities that set Venetan varieties apart from most Italo-Romance speaking areas in Italy (Section 2.2), making them a privileged object of investigation into microvariation – all the more because no linguistic research has been conducted on Opitergino so far.

2.1 *ESCs in NIR varieties*

The first work to systematically explore the occurrence of ESCs in NIR varieties is Renzi & Vanelli (1983). They considered three types of impersonal constructions: weather verbs (e.g. *it is raining*), existentials (e.g. *there is a girl*), and impersonal deontics (e.g. *it is necessary that*). Despite being based on a rather small sample of varieties (30) and respondents (one or two informants per variety), Renzi & Vanelli's survey captured some trends: the NIR varieties that had ESCs in all three impersonal contexts also had subject clitics with both preverbal and postverbal subjects and displayed full-fledged paradigms.[3] This is the case, for example, of the variety of Monno, as in (3) (data from Pescarini & Loporcaro 2022: 153–154): here, both preverbal (3a) and postverbal (Example 3b) subject is mandatorily doubled by the subject clitic and ESCs occur in impersonal constructions (3c).

[3] Exploring the relation between expletives and null subjects is beyond the scope of this article. Here, suffice to say, that the existence of subject clitics and expletive clitics is not a sufficient condition for a language to qualify as non-pro-drop. Most NIR varieties can have defective subject clitic paradigms and have expletives in some impersonal contexts but not in others. For these reasons, NIR varieties are considered pro-drop languages, other than languages such as French. In this regard, Loporcaro (2012: 176) distinguishes between 'subject clitics,' in the sense of nonargumental clitic markers that realize agreement features (as occurring in NIR varieties, e.g. Monno *el plof* 'it rains'), and 'clitic subjects,' in the sense of full-fledged subject pronouns that happen to be clitic and are not subject to the *that*-trace effect (Perlmutter 1968) (as is the case of non-pro-drop languages, e.g. Fr. *il pleut* 'it rains') (cf. Pescarini & Loporcaro 2022: 151–153).

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

(3) (a) *le matele le= lavarè zo i piacc* (variety of Monno)
the girls 3F.PL= wash.FUT.3 down the dishes
'The girls will wash the dishes.'
(b) *el= salta zo le foe*
3M.SG= drop.3 down the leaves
'Leaves are falling.'
(c) *el= plof*
ESC= rain.3
'It's raining.'

Other varieties had defective paradigms. This is the case of the variety of Trieste (Friuli) (4), where the subject proclitic *le* is optional with preverbal subjects (4a), but mostly excluded with postverbal subjects (4b), and there is no ESC (4c) (data from Pescarini & Loporcaro 2022: 153–154).

(4) (a) *le mule (le=) laverà i piati* (Triestino)
the girls 3F.PL= wash.FUT.3 the dishes
'The girls will wash the dishes.'
(b) *Ø casca le foie*
drop.3 the leaves
'Leaves are falling.'
(c) *Ø piovi*
rain.3
'It's raining.'

Renzi & Vanelli (1983) also noted that – overall – ESCs are mostly found with weather verbs, while they are scarcely attested with impersonal deontics. Based on their findings, they designed an implicational scale according to which having ESCs with weather verbs is a prerequisite for having ESCs in all other impersonal constructions (Renzi & Vanelli 1983: 135–136).

Capitalizing on Renzi & Vanelli's work, Pescarini (2012; 2014; 2016) explored the distribution of ESCs in a sample of 140 NIR varieties that included two additional impersonal contexts, namely, epistemic (e.g. *it seems that*) and impersonal *si* (e.g. *one says that*) constructions. The implicational scale proposed by Pescarini (2012; 2014; 2016) (cf. 5) resonates with Renzi & Vanelli's (1983): ESCs are more likely to occur with weather verbs. This possibility decreases as one moves to the right of the scale: ESCs 'are less common in other impersonal contexts like raising constructions, existential constructions, impersonal *si* constructions, and in combination with the impersonal modal of necessity' (Pescarini 2016: 747). This would imply a higher quantity of varieties having ESCs with weather verbs (68% of Pescarini's sample) than having ESCs with other constructions (existentials: 59%; epistemics: 54%; impersonal *si*: 32%; impersonal deontics: 22%); and the distribution of ESCs in each variety would follow an implicational scale. NIR varieties

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

that admit ESCs with one construction require that ESCs also occur with all other constructions to its left in the scale.

(5) weather verbs > existentials > epistemics > impersonal *si* > impersonal deontics

Another milestone in research on subject clitics is Manzini & Savoia (2005, I: Ch. 2). It included data from 187 Italo-Romance varieties (as counted by Pescarini 2016) and involved three types of impersonal context: weather verbs, constructions with a postverbal phrasal subject that does not agree with the verb (corresponding to, e.g. **here comes the boys*), and impersonal constructions that take a clause as a complement (e.g. *it is better that*). The authors merely observed that NIR varieties show different behaviors: some admit ESCs in all three contexts, others only in some (Manzini & Savoia 2005, I: 193–191). Some of the data in Manzini & Savoia seem not to comply with Pescarini's (2012; 2014; 2016) scale. For example, in some varieties, ESCs do not occur with weather verbs and yet they occur in other contexts. This is the case of the variety of Povegliano (Veneto), exemplified in (6) (data from Manzini & Savoia 2005, I: 193):

(6) (a) *l=e ˈmɛjo anˈdar ˈkaza* (variety of Povegliano)
ESC=is better go home
'It is better to go home.'
(b) *Ø a pjoˈvest*
has rained
'It rained.'

Recently, Pescarini (2022) statistically analyzed data from 350 data points. While this is noteworthy, the database is an aggregation of the data from Manzini & Savoia (2005) and the ASIt database.[4]

The works hitherto cited certainly give an account of the diatopic variation and complexity of the phenomenon. Data were collected using questionnaires (sometimes written questionnaires) from generally no more than one or two (rarely, up to five) informants per data point. Since the purpose was to delineate the systems of ESC in the NIR varieties under investigation, the selected informants plausibly represented their conservative incarnations. There is a downside to this approach, however: empirical evidence gathered this way does not allow us to (a) capture sociolinguistic variation and (b) understand whether the absence of ESCs in a given context of a given variety is systematic or due to idiolectal variation. As Ferrarotti (2019: 111) puts it, if an ESC 'does not appear in a given context, that is not necessarily to say that it is unacceptable, for it may simply have been omitted'.

To the best of our knowledge, the only investigation into the topic that involves a representative population sample is Casalicchio & Frasson's (2018) survey on the

[4] ASIt = *Atlante Sintattico d'Italia*, Università di Padova, Università di Venezia. URL http://asit.maldura.unipd.it/

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

four Venetan varieties of Paduan, Trevigiano, Venetian, and Vicentino. The authors reported that, while 91% of the respondents judged the absence of ESCs such as in Trevigiano *ø ze drio nevegar* 'it is snowing' as correct, a substantial portion of respondents (42.1%) did accept ESCs with this construction (cf. Trevigiano *el=ze drio piovar* 'it is raining'). The authors interpreted this result as an overextension of the syntactic rule on the use of subject clitics. However, Casalicchio & Frasson's (2018) study is problematic in two respects: first, it deals with just one type of impersonal construction, namely, weather verbs, leaving out at least five other relevant contexts that can license the occurrence of ESCs; second, it used four questionnaire variants, based on the varieties of the four towns of Padua, Treviso, Venice, and Vicenza; however, the questionnaires were administered to speakers living in all municipalities of these provinces. This procedure amounts to largely ignoring the extant diatopic variation.

2.2 *Change in the sociolinguistic environment*

For more than 70 years now, a heavy sociolinguistic change has invested the entire Italo-Romance landscape (in Italy). After the Second World War, several factors boosted the use of (mainly Tuscan-based) standard Italian across the population. One of these factors was the purchase of televisions on a large scale, made possible in the wake of the 'economic boom' enduring until the end of the 1960s, which sustained the population's exposure to the standard variety used on national television programs. As a result of this exponentially increased exposure, the speakers' competence in local varieties (generally referred to as 'dialects') has changed dramatically over the past 70 years. The contact between the dialects and the standard has had bidirectional effects: On the one hand, dialect-dominant speakers have imposed traits of their local Italo-Romance varieties onto Italian, leading to the emergence of REGIONAL ITALIAN varieties (for details, cf. Cerruti 2011); on the other, standard Italian has ousted the local Italo-Romance varieties, gradually replacing them, and establishing itself as the dominant language, and this came along with a general contraction of dialect speakers in the Italian peninsula. As statistics reported in De Mauro (2014: 113) show, the use of dialects as a sole means of communication decreased from 64.0% in 1955 to 5.4% in 2006. At the same time, the percentage of people using Italian as a sole means of communication increased from 18.0% in 1955 to 45.5% in 2006. The most recent data (ISTAT 2017) from 2015 indicate that in the domain 'in the family,' 45.9% of the population (from age 6 onwards) predominantly use Italian, while 14.0% predominantly use a dialect, and 32.2% use both Italian and a dialect. Similar dynamics are found in the domains 'with friends' and 'with strangers'.

The exclusive or predominant use of local varieties decreases in all age groups: as of 2015, 32.0% of speakers aged 75 or above state that they exclusively or predominantly speak dialect in the family, while in 2006, it was 37.1%. From a diastratic viewpoint, dialect use in the family and with friends is maintained to a

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

larger extent among speakers with a low educational qualification, irrespective of age: among speakers with 8 years of schooling, 24.8% exclusively use dialect in the family and 33.7% with friends, whereas among speakers with a bachelor or a higher degree, the figures are 3.1% in the family and 2.7% with friends.

However, this decrease has not affected all areas of the peninsula homogeneously. The ISTAT (2017) 2015 data show that Veneto, despite a progressive increase in the use of Italian, ranks among the regions in which the use of dialects is still relatively high: in the domain 'in the family,' 30.6% of the population (from age 6 onwards) predominantly use a dialect, while, for example, in neighboring Lombardy, it is 5.6%. If we lump together the declared sole or predominant use of dialect in the family with that of both Italian and dialect, we get 62.0% in Veneto versus 31.7% in Lombardy. Nowadays, Veneto is characterized by a situation of diglossia involving standard Italian and local varieties, which are vital and spoken on a daily basis: Older generations (aged 80 onwards) tend to have a dialect as L1 and the dominant language, both qualitatively (for example, in terms of proficiency) and quantitatively (in terms of social domains in which it is spoken), while in younger generations, standard Italian is rather the dominant language.

3. The study

3.1 *Research questions and hypotheses*

In this case study, we move from the following sociolinguistic and contact-theoretical considerations: (a) Starting from the 1950s, standard Italian has increasingly become the dominant language in the wake of the massive sociolinguistic change that has invested Italy, even in areas such as Veneto in which Italo-Romance varieties have been more resilient. (b) In language contact, the source of change is the (individually or socially) dominant language, which might encroach upon the structure of the recipient language in a systematic way. (c) However, contact does not necessarily lead to change. As a matter of fact, the very prerequisite for contact-induced change to happen is the existence of divergent structures (cf. Gardani 2022: 847–850). The presence of ESCs in Opitergino and their absence in its contact language, standard Italian, represents precisely such a case of structural incongruence, which can – though need not – lead to change.

Based on this, we pursue the following main research question (RQ1): Has the contact with standard Italian induced change in the ESC system of Opitergino? At this point, it should be stressed that 'deterministic and absolute predictions' (Thomason 2000: 175) on the linguistic outcome of language contact are impossible to make – at least, as concerns single contact settings (whereas making predictions over crosslinguistic samples is a viable option, as shown, e.g. by Gardani 2008; 2012). Instead, we can formulate hypotheses about what is reasonable to expect. In our specific case, the predominance of standard Italian over local varieties and its being the moderately dominant language in large parts of the population prompts us

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

to discard the null hypothesis and to deem it probable that (H1) SOME DEGREE OF CONTACT-INDUCED CHANGE HAS INDEED OCCURRED.

Secondly, if change has occurred, (RQ2) WHAT SORT OF CHANGE IS IT REASONABLE TO EXPECT? In that respect, we cannot sufficiently rely on results from previous investigations, as research on the behavior of clitics in language contact is generally meager.[5] A likely outcome could be a convergence towards the structure of the dominant language, standard Italian. However, the interplay of sociological and structural factors provides an extra layer of intricacy. The sociolinguistic structure of a speech community is known to play a decisive role in determining not just the directionality of change but also – crucially – the extent to which structural change can occur (cf. Weinreich's 1953: 112 'twin approach'). In areas in which Italian is clearly the dominant language (such as Lombardy), we could – in principle – expect a net convergence of Italo-Romance varieties towards Italian, resulting, say, in the loss of ESC.[6] But in Oderzo, the sociolinguistic dynamics are different; hence, such a clear-cut outcome is less probable: Speakers of Opitergino are proficient speakers,[7] and proficient bilinguals tend to be more preserving than nonproficient speakers. Based on these observations, we expect that (H2) CONTACT HAS PRODUCED SLIGHT STRUCTURAL CONVERGENCE TOWARDS STANDARD ITALIAN NOT RESULTING IN LOSS.

Thirdly, if convergence has occurred, where and how did it become manifest? We focus on syntactic change, specifically on the obligatoriness of rules, which we understand as context-sensitivity. We ask (RQ3): DOES OBLIGATORINESS (VS OPTIONALITY) PLAYS A ROLE IN CONTACT-INDUCED CHANGE? Note that this question has been largely disregarded by theoretical debates on language contact, a remarkable exception being Thomas (2012). The key element here is that rule optionality may generate a certain amount of uncertainty. Matter-of-factly, Thomas (2012: 208) 'found that like monolinguals, bilinguals tend to limit optional grammars, but do so in different ways according to the unique cognitive and experiential factors associated with their age of exposure to bilingualism'. A possible speakers' reaction to uncertainty is reducing it, which can be implemented via either obliteration of optionality (optional rules become obligatory) or rule generalization ('removing a condition from a rule,' according to Harris & Campbell 1995: 102). As we will see (Section 3.2), in Opitergino, ESCs are context-sensitive in declarative clauses, whereas they are context-insensitive in interrogative clauses. We therefore expect that (H3) CHANGE HAS AFFECTED ESCS IN THE DOMAIN OF DECLARATIVE CLAUSES TO A GREATER EXTENT THAN IN INTERROGATIVE CLAUSES.

[5] While some studies exist (e.g. Klee 1990; Souag 2017; Casalicchio & Frasson 2018), the most recent investigations have focused on heritage speakers (e.g. Montrul 2010; Frasson, D'Alessandro & van Osch 2021; Ivanova-Sullivan et al. 2022), who are clearly unbalanced bilinguals, other than Opitergino-Italian bilinguals.

[6] Structural convergence results from pattern borrowing (Matras & Sakel 2007; Sakel 2007; Gardani 2020). It can manifest as addition or loss (cf. Gardani 2022: 848) and affect several levels of the grammar of ESCs. For example, possible scenarios of change in morphology might encompass paradigmatic leveling, syncretism, and defectiveness (all these possibilities are strongly filtered by sociolinguistic factors).

[7] Note that we targeted speakers who self-declared as proficient in Opitergino.

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

Besides, we cannot rule out secondary 'internal rearrangements' (Heath 1984: 368; called 'later changes' in Thomason 2001: 62), but any hypothesis to this effect would be impressionistic.

To test the three hypotheses, in the remainder of the paper, we (a) provide a baseline of ESCs in conservative Opitergino (Section 3.2), (b) present the results of an online survey we conducted in 2022 with 100 informants (Section 4), and (c) compare them (Section 5).

3.2 *The baseline*

There are virtually no studies focusing on Opitergino. While some studies exist on other varieties belonging to Liventino, the dialect group to which Opitergino belongs (Zamboni 1974; 1979), and more generally on northern Venetan dialects (also including Trevigiano, Feltrino, and Bellunese), the documentation of the Liventino varieties is scarce. The oldest Opitergino text we found is a short text by Francesco Carlo Gasparinetti from the 19th century, recorded in Papanti (1875: 515); unfortunately, this text does not contain any impersonal construction, thus no ESCs. There are a few more texts in Opitergino, but besides their being poems, they do not match the chronological depth necessary to build the baseline (authors were born after the 1960s and, in some cases, left Veneto in early childhood, e.g. Franzin 2013). Therefore, we set to establish our baseline ourselves, aiming at the conservative Opitergino spoken before the widespread Opitergino/standard Italian bilingualism of recent decades. Relying on an apparent time approach (cf. Labov 1963), we extracted the baseline from the speech of elder people who speak Opitergino as an L1 (while Italian is their L2). To create the relevant baseline, we conducted interviews with four speakers (m = 2) born before 1942, as reported in Table 1.

First, we put together the paradigms of strong personal subject pronouns and subject clitics (both proclitics and enclitics). They are given in Table 2. We also determined that ESCs exist in Opitergino. They have the same forms of the III M.SG, namely, proclitic *el* and *l* (the latter occurring before copular *é*) in declarative sentences (henceforth, 'declaratives'), and enclitic *eo* and *o* in interrogative sentences (henceforth, 'interrogatives').

As Table 2 shows, the paradigm of subject proclitics (found in declarative clauses) is defective: first singular, first plural, and second plural persons are lacking. This is in

Informant	Year of birth	Education (number of years)	Speaking dialect
SB	1930	5	Daily
PGP	1934	13	Daily
OS	1937	5	Daily
AC	1941	5	Daily

Table 1
Demographic profile of the baseline informants (conservative Opitergino).

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

		Strong/stressed	Clitic	
			Proclitic	Enclitic
1SG	'I'	mí	—	(=eo/=jo)
2SG	'you'	tí	te=	=tu
3M.SG	'he'	lú	el=/l=	=eo/=o
3F.SG	'she'	éa	la=/a=	=ea
1PL	'we'	no(i)áltri	—	=eo
2PL	'you'	vo(i)áltri	—	=o
3M.PL	'they'	lóri	i=	=i/=ei
3F.PL	'they'	lóre	e=	=e/=ee

Table 2
The system of subject pronouns in Opitergino.

line with the characteristics of the Liventino group (Zamboni 1979). By contrast, the paradigm of subject enclitics (in interrogative clauses) is not defective. This asymmetry between the proclitic and enclitic paradigms is also a common feature of almost all Venetan varieties (Zamboni 1974: 50). According to some authors (Renzi & Vanelli 1983; Pescarini 2012; 2014; 2016; cf. *intra* Section 2), NIR varieties with defective proclitic paradigm are expected to have no obligatory subject doubling and no ESCs.[8] However, in Opitergino, subject clitics are mandatory with preverbal subjects, that is, doubling does occur (7a), while they are mostly excluded in the presence of postverbal subjects (cf. *e fòje* in 7b).

(7) (a) *e tose e= lavarà i piati*
the girls 3F.PL= wash.FUT.3 the dishes
'The girls will wash the dishes.'
(b) *Ø casca e fòje*
drop.3 the leaves
'Leaves are falling.'

In a second step, we investigated which impersonal contexts license the presence of ESCs in Opitergino. Capitalizing on the studies mentioned in Section 2, we considered the following six constructions:[9]

– Subjectless Predicative Copular (SPC) construction (e.g. Italian *è una bambina* 'it is a girl'), which also includes existential construction (e.g. Italian *c'è una bambina* 'there is a girl'). SPC constructions display the following properties: they have no phrasal subject; they have a copula; and they contain

[8] Benincà & Poletto (2004) argue that doubling of non-dislocated preverbal subjects is probably obligatory.
[9] All previous literature only deals with existential constructions. We introduced 'Subjectless Predicative Copular' as a new type of syntactic construction to also account for non-existential constructions: Opitergino does not formally distinguish between *l'é un tosatel* as existential 'there's a boy' vs. non-existential 'it's a boy.'

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

a nonverbal predicate (noun or adjective phrase) placed after the copula (Salvi 2001, II: 174–175; Cennamo 2016: 979; Bentley 2017).

- Extrapositions (e.g. Italian *è ovvio che* 'it is clear that') have no phrasal subject and take a sentence as a complement to which ESC is related. The finite verb is inflected in the III SG (Manzini & Savoia 2005, I: 182–183).
- Weather verbs (e.g. Italian *piove* 'it is raining') are zero-argument verbs and have no phrasal subject. A 'semi-argument' can be realized and placed after the verb when referring to a non-prototypical referent (e.g. Italian *piovono soldi* 'it is raining money,' lit. money rain) (Chomsky 1981; Manzini & Savoia 2005, I: 182–183; Pescarini 2014: 239; 2015: 70–72).
- Epistemics (e.g. Italian *sembra che i ragazzi stiano bene* 'it seems that the kids are doing well') have no phrasal subject and take a sentence as a complement, the finite verb is inflected in the III SG. ESC is bound to the subject of the embedded clause, as is evident from the corresponding raising structures (e.g. 'it seems that the kids are doing well' corresponds to 'the kids seem to be doing well') (see, among others, Postal 1974).
- Impersonal *si* constructions (e.g. Italian *si dice* 'one says') have no phrasal subject, and the finite verb is inflected in the III SG. It is used for events with undefined, human, logically implied participants, marked by the clitic *si* but not explicitly realized (see, among others, Cennamo 1993; D'Alessandro 2007; Parry 1994; Pescarini 2015).
- Impersonal deontics (e.g. Italian *bisogna* 'it is necessary that') have no phrasal subject, and the finite verb is inflected in the III SG. Impersonal deontics are zero-argument, modal verbs with a defective paradigm and take a sentence as a complement (Benincà & Poletto 1994; Berizzi 2012).[10]

As for declaratives, all baseline informants consistently reject ESCs with impersonal *si* and impersonal deontic constructions (8e–f). ESCs are ruled out for weather verbs (8c), as well. The baseline informants consistently use ESCs with SPC and extraposition constructions (8a–b). With epistemic constructions, ESCs are optional (8d), with the variation being both intra- and interindividual: in spontaneous production, each informant uses this construction both with and without ESCs, even within the same conversation; in acceptability judgments, all informants positively rate epistemics both with and without ESCs.

[10] Benincà & Poletto (1994; 1997) observe that, in Venetan dialects, *toccare* 'it is necessary that' can be used as a modal verb indicating a state of necessity. They identify two verb types: *toca1* has an argument expressed by a dative clitic; *toca2* has no argument, is purely deontic, and is always inflected in the third person singular. In spite of this difference, we tested both *toca1* and *toca2* because they partake in the same grammaticalization process. For an overview of *toccare*, cf. Berizzi (2012).

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

(8) Degree of obligatoriness of ESC in declaratives in the Opitergino baseline

Obligatory: (a) subjectless predicative copular (SPC)
l'=é un/'na/do tosatel/toséta/tosàti
ESC=is a/a/two young_boy/girl/boys
'There's a boy/girl. / There are two boys.'

(b) extraposition
l'=é ciaro che i vien co mi!
ESC=is clear that they come with me
'It's clear that they'll come with me.'

Optional: (c) epistemic
(el=)par che a vegne doman
(ESC=)seems that she comes tomorrow
'She seems to be coming tomorrow.'

Absent: (d) weather verbs
piove
rains
'It's raining.'

(e) impersonal *si*
se dise cusì
one says so
'This is how to say it.'

(f) impersonal deontic
bisogna ndar
needs go
'We need to go.'

As for weather verbs, ESCs never occur in spontaneous production; in grammaticality judgments, three out of four baseline informants rule them out categorically. The only informant who positively rated ESCs in weather verb constructions is PGP. PGP is the most educated informant (holding a high school degree), worked in public institutions, and often interfaced with speakers of other Venetan varieties. Their judgment might have been influenced by the ESC systems of other varieties.

Note that in both constructions with obligatory ESC, the subject clitic attaches to copula *é* 'is'. Discussing other Venetan varieties, some scholars consider this case not a real ESC but an auxiliary clitic, that is, 'a "dummy clitic" for a position which needs to be realized once some phonological conditions (which differ across dialects) are met' (Poletto & Tortora 2016: 785; see also, Benincà 2007). This issue is controversial, though (cf. Poletto & Tortora 2016: 785 fn. 22). Auxiliary clitics are insensitive to agreement features, which would imply that only one clitic form exists regardless of the grammatical properties of the subject. This is not the case in Opitergino, where the form of ESC is always masculine singular, as expected (8a), whereas the forms of referential subject clitics differ

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

based on gender, both in compound forms (i.e. with an auxiliary) (9a, a') and non-compound forms (9b, b').[11]

(9) (a) *l'=é* *rivà* *a* *miodì* (b) *el=va* *doman*
3M.SG=be.3 arrived.M.SG at noon 3M.SG=go.3 tomorrow
'he arrived at noon' 'he'll go (there) tomorrow'
(a') *la=é* *rivada* *a* *miodì* (b') *a=va* *doman*
3F.SG=be.3 arrived.F.SG at noon 3F.SG=go.3 tomorrow
'she arrived at noon' 'she'll go (there) tomorrow'

Despite the lack of 19th-century Opitergino written texts, we were able to find chronologically comparable documentary evidence from adjoining Liventino varieties (which, as we said, belong to the same dialect group). The picture they provide is consistent with our baseline. In two texts from the nearby variety of Meduna di Livenza, ESC occurs twice in the two attested SPC constructions (2/2), once with one attested extraposition (1/1), while it does not occur with the one attested impersonal deontic [biˈzɔɲa] 'it is necessary that' (0/1) (Zamboni 1974: 88–89).[12] Also, data gathered between 1919 and 1928 (Jaberg & Jud 1928: 14–16) in San Stino di Livenza (ASIt; maps 366, 367; data point 356) show that ESC is absent with weather verbs, both in declaratives and interrogatives.[13]

The data in (8) occur in declaratives, in which ESCs appear to be context-sensitive. Context-sensitivity, that is, different degrees of obligatoriness of rule depending on the syntactic context, does not apply, however, to the use of ESC in interrogatives, which in conservative Opitergino is mandatory across impersonal contexts, cf. (10a–f).

(10) Use of ESC in interrogatives in the Opitergino baseline
(a) subjectless predicative copular (SPC)
é=o *un* *tosatel?*
is=ESC a young_boy
'Is there a boy?'

[11] As reported in Zamboni (1974: 59), Trevigiano has the auxiliary form *zé* (*al= z=é* 'he is'), while conservative varieties of northern Veneto lack it (e.g. Feltrino-Bellunese (*e*) *l=é* 'he is'). Opitergino behaves like these conservative varieties and does not display the auxiliary form *zé*. Notably, the auxiliary clitic *z=* is insensitive to agreement features even with referential subjects (cf. Trevigiano (source: ASIt) *El direttor z=è rivà* 'the.M director(M) arrived'; *A barca z=è 'ndada a fondo* 'the.F boat(F) sank'; *I pressi z=è andai su* 'prices have risen').

[12] Zamboni (1974: 88–89) does not provide any information about the age of the informants. However, judging (a) from the date of publication of the data and their previous elicitation and (b) from the sociocultural details of the autobiographical story told by the informant in the first text, we estimate that the two speakers were born no later than 1900.

[13] A similar pattern has been found for Cilense, the variety of Ceggia (Davanzo 2016: 110–112). However, the data collected by Davanzo are recent, thus not corresponding chronologically.

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

(b) extraposition
é=o ciaro che i vien co mi?
ESC=is clear that they come with me
'Is it clear that they'll come with me?'

(c) weather verbs
pióve=o?
rains=ESC
'Is it raining?'

(d) epistemic
pàr=eo che a vegne doman?
seems=ESC that she comes tomorrow
'Does she seem to be coming tomorrow?'

(e) impersonal *si*
se dìse=o cusì?
one says=ESC so
'Is this how to say it?'

(f) impersonal deontic
bisògn=eo 'ndar?
needs=ESC go
'Do we need to go?'

3.3 *Design*

3.3.1 *The questionnaire*

We designed a questionnaire to evaluate the acceptability of the occurrence of ESCs in the impersonal constructions illustrated in Section 3.2.[14] Availing ourselves of coauthor GB's native proficiency in Opitergino, we created ten sentences for each type of construction, and each of these sentences appeared once with an ESC and once without it. Half of the sentences were declaratives, half were interrogatives, amounting to 120 stimuli. The experimental conditions are summarized in Table 3. Given such a large number of stimuli, we chose not to add fillers. This aimed at both keeping the questionnaire fill-out time to a minimum and at avoiding loss of attention in informants.[15]

Because of the absence of corpora for Opitergino, the frequency of the stimuli could not be measured. To cope with this situation, another native speaker was asked to validate all sentences, which were rather short, composed of very familiar

[14] Our survey is centered on acceptability because we focus on syntactic change (cf. Section 3.1); for morphological change, production tasks would be more suitable.

[15] From a methodological point of view, the absence of fillers does not constitute any sort of problem because the stimuli were very different from each other, which excluded any automaticity issue in performing the task. Crucially, our study is not a psycholinguistic experiment but a survey drawing on semi-experimental methods (e.g. Meakins, Green & Turpin 2018: 254).

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

		Number of sentences per condition			
		Declaratives		Interrogatives	
Types of construction	Examples	ESC: Y	ESC: N	ESC: Y	ESC: N
Weather verbs	'it's raining'	5	5	5	5
SPC	'there is (a child)'	5	5	5	5
Epistemic	'it seems that…'	5	5	5	5
Extraposition	'it is clear that…'	5	5	5	5
Impersonal *si*	'one says…'	5	5	5	5
Impersonal deontic	'it is necessary that…'	5	5	5	5

Table 3
Experimental conditions (Y = with ESC; N = without ESC).

words, and typically used in conversational contexts. Whenever present, the ESCs never surfaced in either the initial or final sentence position. All experimental stimuli were audio-recorded by a native speaker, and an audio file was created for each sentence. All stimuli are listed in Appendix 1.

The questionnaire was administered online via the QuestionPro platform (https://www.questionpro.com/). It could be accessed through a link and filled out using PCs, smartphones, and tablets. Participants were invited to click on the play symbol, listen to each audio file, one at a time, and rate the sentence they had just heard on a 5-point Likert scale. The range of possible answers was: 1 = completely unacceptable; 2 = unacceptable; 3 = neither unacceptable nor acceptable; 4 = acceptable; 5 = perfectly acceptable. Participants were unaware of the precise objective of the study to avoid bias and automaticity in responses (they were only aware of the general objective of the study, which was to detect change in Opitergino). They were instructed to not express normative judgments but to evaluate sentences based on their perceptions, asking themselves, for instance, whether they would also use those sentences. If they would never and on no occasion use the sentence they heard because they judged it to be very bad, they were asked to assign the lowest score. If the sentence sounded very good, they should assign the highest score. If the sentence did not sound good, but not bad either, or it could have been said on some occasions, an intermediate score could be given. Once the score for a sentence was given, participants could listen to the next audio file. The fact that all stimuli were audio recordings ensured that all participants received the same stimuli under the same conditions. For the case of participants struggling with filling out the questionnaire on a digital device, help from younger persons was encouraged but limited to pressing the buttons in place of the participant. The audio files in the questionnaire were presented to each participant in a different random order. The software automatically generated a different order each time the questionnaire was accessed through the link.

Before listening to the audio files, participants were asked to anonymously answer some questions targeting their sociolinguistic profile: year of birth; gender; educational qualification; place of living; parents' provenance; proficiency in

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

standard Italian and in Opitergino; frequency of dialect use; and people with whom and places where the dialect is used. Relying on the Language Experience and Proficiency Questionnaire (LEAP-Q) (Marian, Blumenfeld & Kaushanskaya 2007), we considered these sociolinguistic measures a proxy for proficiency. In language contact situations, direct assessment of competence through language tests can be biased by many factors, while self-reported levels of competence, such as those adopted by Adamou & Shen (2019), prove to be a more reliable and less biased estimate (e.g. Meakins, Green & Turpin 2018: 257–258).

On average, completing the questionnaire took 17 minutes.

3.3.2 *Participants*

One hundred and seventy-six informants were recruited through social networks, the local press, the parish, and the municipality website; they consented to participate in the survey on a voluntary basis. Informed consent was sought from each participant before taking part in the study. One hundred informants (female, f = 60; male, m = 39; other, o = 1) completed the questionnaire. As shown in Figure 1a, they were fairly well distributed by age group: younger (18–40) = 29; middle-aged (41–60) = 33; and elderly (61–92) = 38. Sixty-eight informants declared to live in the municipality of Oderzo, while 11 resided outside the province of Treviso (Figure 1b), and more than half (n = 57) had at least one parent from Oderzo (Figure 1c). Most of them reported having 13 to 16–18 years of education, which roughly corresponds to a high school diploma (n = 45) and a bachelor's/master's degree (n = 39), respectively (Figure 1d).

Regarding the language profile, 69 out of 100 informants reported speaking standard Italian since before their schooling, although the proportion decreases as age increases (younger = 0.93; middle-aged = 0.64; elderly = 0.55; Figure 2a). The reverse trend is observable in the use of dialect: 61 informants reported speaking Opitergino since before their schooling (Figure 3a); however, in this case, the proportion increases as age increases (younger = 0.5; middle-aged = 0.64; elderly = 0.72; Figure 2b). While these data show the relative prevalence of dialect use across all age groups, they also confirm that other than older people, young people are more likely to have standard Italian as their L1.

Thirteen participants in the questionnaire declared passive competence of Opitergino (younger = 0.21; middle-aged = 0.06; elderly = 0.14; Figures 2b and 3a). These participants were excluded from the final analysis, which was therefore carried out on a total of 87 informants. The vast majority of informants speak a dialect (i) on a daily basis (77 out of 87; Figure 3b); (ii) mostly with family members and acquaintances (45 out of 87) but also with strangers (i.e. with everyone = 23 out of 87; Figure 3c); and (iii) in both private (i.e. home) and public (i.e. pubs, restaurants, shops, etc.) settings (32 out of 87) but also at work (i.e. everywhere = 29 out of 87; Figure 3d).

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

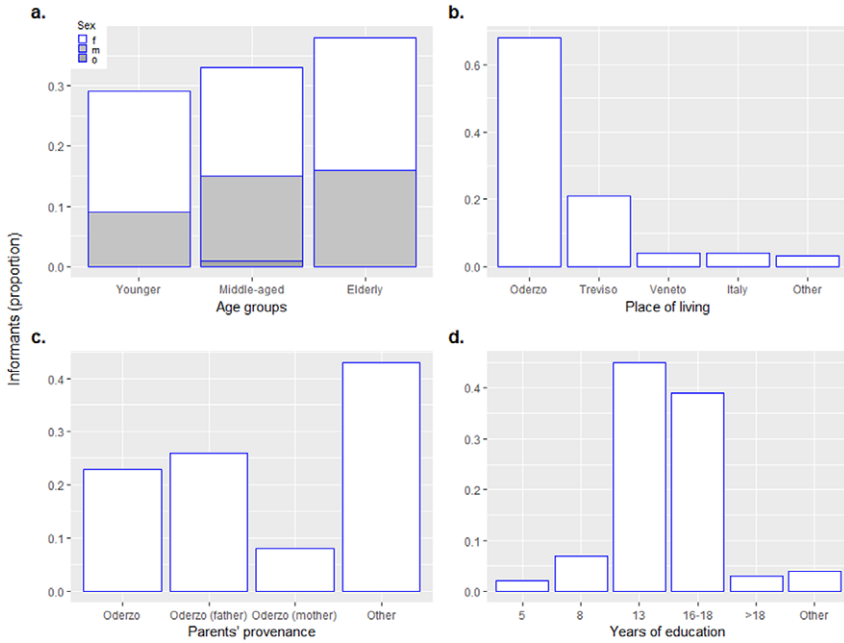

Figure 1
Demographic profile of the informants (n = 100) who completed the questionnaire.

3.4 *Methods for data analysis*

We analyzed the data collected from the questionnaire by means of the R software (R Core Team 2020). First, we calculated the proportion of answers per point of the Likert scale in relation to the presence or absence of ESCs and the type of impersonal construction in both declaratives and interrogatives.[16] Second, following Tagliamonte & Baayen (2012), we analyzed the differences in the rating of the ESCs across impersonal constructions by means of conditional inference trees and random forest, making use of the party package in R (Hothorn et al. 2006; Strobl et al. 2007; 2008). These non-parametric models are particularly suitable when the sample size is small, while the number of predictors is high and are robust in case of outliers. Conditional inference trees are recursive partitioning methods. The model selects 'the predictor which helps best to distinguish between different values of the response variable' and makes 'a split in this variable, splitting the data in several data sets' (Levshina 2021: 612). These two steps are repeated until there are no

[16] Note that answers in questionnaires on a Likert scale are ordinal (and thus categorical) variables (Agresti 2019; Veríssimo 2021). For this reason, we calculated proportions, but we did not provide measures, such as arithmetic mean or standard deviation. More generally, all statistical methods chosen in this study are suitable for categorical variables (for a comparison of different methods, see Janda & Endresen 2017).

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

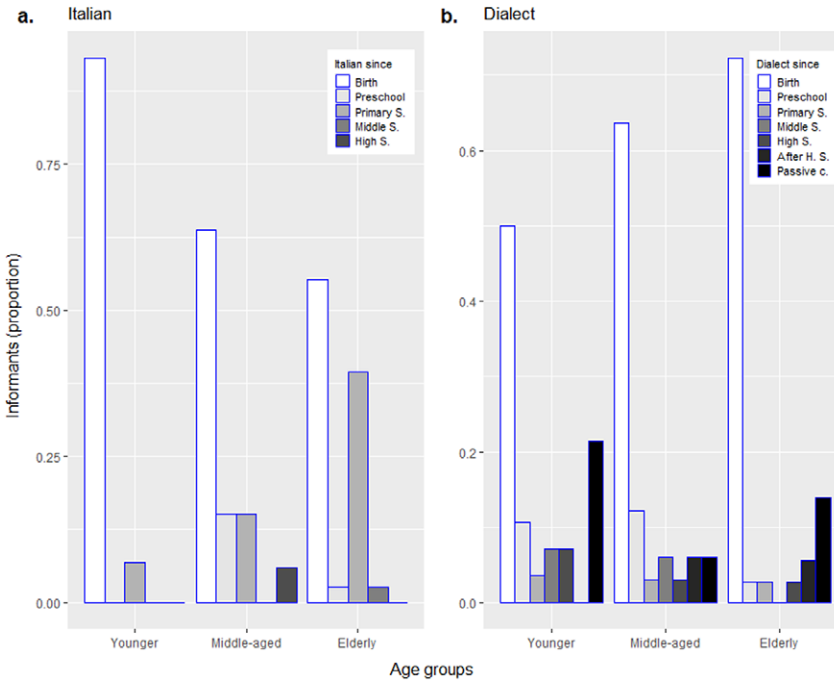

Figure 2
Language use by age group (n = 100): standard Italian (2a) vs. Opitergino (2b). Expanded legend: birth; preschool, primary school, middle school, high school, after high school, passive competence.

variables that are significantly associated with the outcome. The splits are visualized as branches in a tree structure. As for a random forest, it 'represents an ensemble method, by which many individual trees are "grown", and their predictions are averaged' (Levshina 2021: 612). Random forests give the impact of each predictor relative to all other predictors.

We fitted two random forest models, one for declaratives and one for interrogatives, to inspect the importance of linguistic and sociolinguistic predictors in terms of their ranking. We also fitted two conditional inference tree models, again, one for declarative and one for interrogative sentences, to check whether and how the predictors interact with each other. For all models, we used the rating scores as the response variable and the following factors as predictors (independent variables):

(α) ESC (presence of the expletive subject clitic: Y = yes, N = no);
(β) Construction (impersonal construction type: Deo = impersonal deontic; Epi = epistemic; Extra = extraposition; Imp = impersonal *si*; SPC = subjectless predicative copular; Wea = weather verb);
(γ) Age;
(δ) Sex (f = female, m = male, o = other/prefer not to disclose);

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

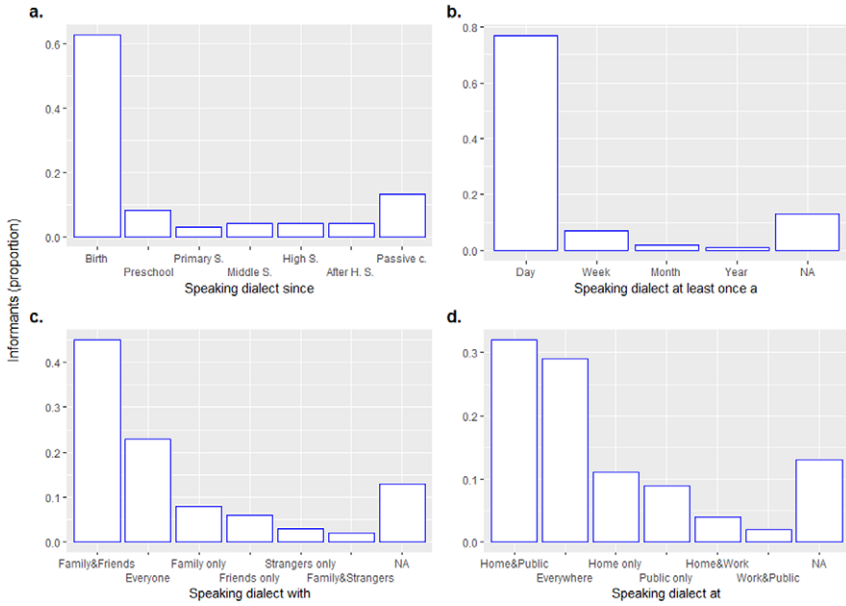

Figure 3
Use of dialect (n = 100). In Figure 3a, the considered levels on the x-axis are: birth, preschool, primary school, middle school, high school, after high school, passive competence. NA in Figures 3b–d refers to the proportion of informants reporting passive competence in Opitergino.

(ε) Education (years of education: 5, 8, 13, 16–18, >18; other);
(ζ) P_o_L (place of living: Oderzo = municipality of Oderzo, Treviso = other municipalities in the province of Treviso, Veneto = other provinces in the region of Veneto, Italy = other regions of Italy, Other = other countries);
(η) Parents_Origin (parents' provenance: Oderzo = both parents from Oderzo; Oderzo(father) = only the father from Oderzo; Oderzo(mother) = only the mother from Oderzo; Other = no parents from Oderzo);
(θ) Italian_Since (speaking Italian since: birth, preschool, primary school, middle school, high school, after high school);
(ι) Dialect_Since (speaking a dialect since: birth, preschool, primary school, middle school, high school, after high school);
(κ) Dialect_Freq (speaking a dialect at least once a day, week, month, year);
(λ) Dialect_With (speaking a dialect with family/friends/strangers);
(μ) Dialect_Where (speaking a dialect at home/at work or school/in public places).

Predictor (α) was accounted for by providing the informants with 120 sentences, of which 60 contained ESCs and 60 did not. Predictor (β) was captured by providing sentences with all types of impersonal constructions mentioned therein. Predictors

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

(γ)–(μ) were accounted for by the sociolinguistic questions of the questionnaire preceding the audio files (cf. Section 3.3.1).

4. Results

We collected ratings for 120 stimuli, consisting of 60 declaratives and 60 interrogatives. Each group contained 30 stimuli with ESCs and 30 stimuli without ESCs. The results of the analysis are presented in the following subsections. In each subsection, we first summarize the results of the rating by providing the proportions of responses per each point of the Likert scale per each condition. Then, we focus on the significance of differences between conditions considering linguistic and sociolinguistic predictors and describe the results of conditional inference trees and random forest models. At the end of each subsection, we summarize the results of the survey, by sentence type, namely, declaratives (Section 4.1) and interrogatives (Section 4.2).

4.1 *ESCs in declaratives*

We collected ratings of 60 declaratives, 30 with ESCs and 30 without, from 87 respondents (cf. Section 3.3.2), for a total of 5,220 answers. Overall, the acceptability ratios for declaratives without ESCs, of the type exemplified by (11a), were higher than those with ESCs, of the type exemplified by (11b). This is shown in Table 4.

(11) Context: declarative, SPC, −ESC (a) vs. +ESC (b)
(a) *é* *tardi*
is late
(b) *l'=é* *tardi*
ESC=is late
'It's late.'

Among the impersonal construction types, the one with the highest score in stimuli with ESCs is the SPC construction ('5–perfectly acceptable' = 0.7) as in (11b), followed by extraposition ('5–perfectly acceptable' = 0.33; Figure 4a) as in (12).

(12) Context: declarative, extraposition, +ESC
l'=é *ciaro* *che* *se* *fa* *cusì*
ESC=is clear that one does so
'It's clear that it has to be done like this.'

The other constructions (weather verb, epistemic, impersonal *si*, impersonal deontic) were rated poorly in sentences with ESCs: the proportion of scores equal to '5–perfectly acceptable' does not exceed 0.3 (epistemic: 0.24; impersonal *si*: 0.20; impersonal deontic: 0.18; weather verb: 0.18). See an example with an epistemic context in (13).

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

Context	Likert point	Proportion
Declarative with ESCs	5–perfectly acceptable	0.31
	4–acceptable	0.17
Declarative without ESCs	5–perfectly acceptable	0.48
	4–acceptable	0.23

Table 4
ESC acceptability ratios in declaratives.

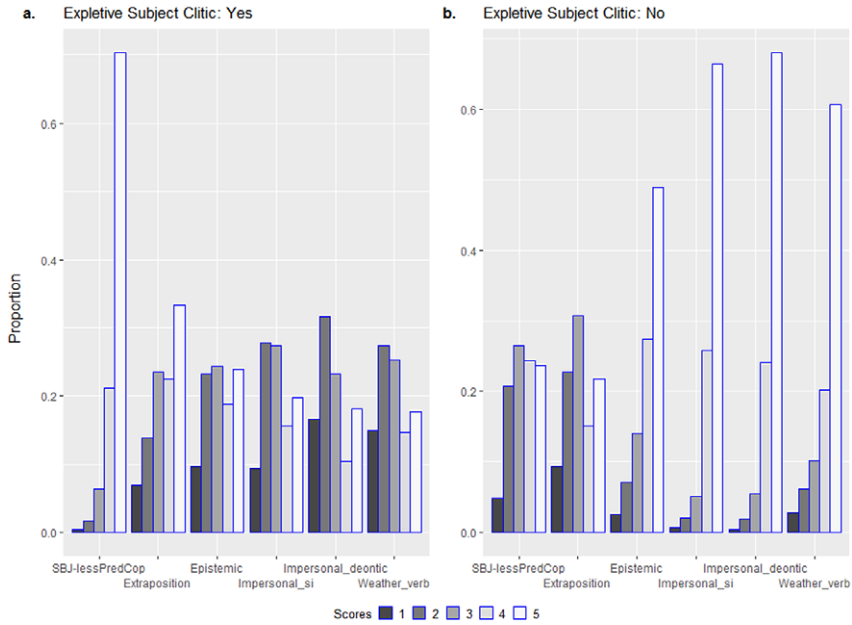

Figure 4
Rating scores across impersonal constructions in declaratives (n = 5,220). SBJ-lessPredCop stands for subjectless predicative copular (SPC).

(13) Context: declarative, epistemic, +ESC
el=par che se magne ben
ESC=seems that one eats well
'It seems like the food's good.'

In this regard, however, it is worth noting that the proportions of scores equal to '1–completely unacceptable' are similarly low (epistemic: 0.1; impersonal *si*: 0.09; impersonal deontic: 0.16; weather verb: 0.15) and more generally, ratings are distributed among the intermediate scores, peaking at scores '2–unacceptable' (epistemic: 0.23; impersonal *si*: 0.28; impersonal deontic: 0.31; weather verb:

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

0.27) and '3 – neither unacceptable nor acceptable' (epistemic: 0.24; impersonal *si*: 0.27; impersonal deontic: 0.23; weather verb: 0.25).

As concerns declaratives without ESCs (Figure 4b), we observe a reverse pattern. Here, weather verb, epistemic, impersonal *si*, and impersonal deontic constructions, e.g. as in (14), were rated higher ('5–perfectly acceptable' = 0.61, 0.49, 0.66, and 0.60, respectively) than SPC and extraposition constructions, e.g. as in (15) ('5–perfectly acceptable' = 0.24 and 0.22, respectively).

(14) Context: declarative, weather verb, −ESC
piove!
rains
'It's raining!'

(15) Context: declarative, extraposition, −ESC

é	*ciaro*	*che*	*se*	*fa*	*cusì*
is	clear	that	one	does	so

'It's clear that it has to be done like this.'

Even in this case, however, the proportion of scores equal to '1–completely unacceptable', relating to the lowest rated constructions, is low (SPC: 0.05; extraposition: 0.09), and ratings are rather distributed among the intermediate scores, peaking at score '3–neither unacceptable nor acceptable' (SPC: 0.26; extraposition: 0.31).

Moving to data analysis, in a first step, we applied the random forest model. The predictive power of the model is satisfactory: its out-of-bag (OOB) classification accuracy is 0.62 (with the baseline value being 0.2).[17] The impact of each predictor in the forest model is plotted in Figure 5. In the chart, the horizontal axis displays the conditional variable importance for each predictor considering all other predictors and their interactions. The score associated to a predictor is the average decrease in the prediction accuracy of the model when that predictor is permuted. In short, the more strongly a predictor is associated with the response variable, the greater the decrease in the prediction accuracy of the model (Levshina 2021: 617). Predictors are listed on the vertical axis according to their importance: a predictor at the top has a greater impact than a predictor further down. The variable importance scores 'should only be interpreted with regard to their ranking, and not as absolute values' (Levshina 2021: 636). The blue dashed line divides important scores from unimportant ones. Values of irrelevant predictors vary around zero. The cut-off value is that of the predictor associated with the lowest score.

In our model, the variable importance scores show that the independent variables (α) 'ESC' (0.117), that is, the presence or absence of an ESC, and (β) 'Construction'

[17] Certain data points are not used by the algorithm to build the model. They are not the predicted values for the training samples but those for the out-of-bag (OOB) samples. It is advisable to use OOB samples to evaluate the accuracy of the model because measures based on the training samples may be unreliable estimates (Levshina 2021; Strobl et al. 2009).

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

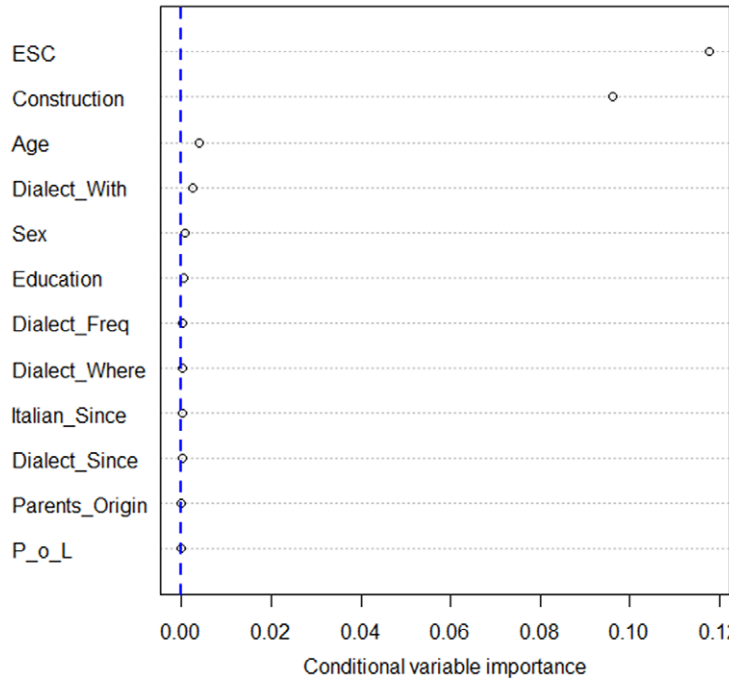

Figure 5
Conditional permutation importance of variables in the rating of declaratives (n = 5,220). On the vertical axis, Dialect_Freq refers to the frequency with which respondents speak Opitergino. P_o_L refers to the place where respondents live.

(0.096), that is, the type of construction in which an ESC is used or not used, are – by far – the most important predictors. Some degree of predictivity can be detected for the predictors (γ) 'Age' (0.003) and (λ) 'Dialect_With' (0.002), whereas the other sociolinguistic predictors (δ, ε, ζ, η, θ, ι, κ, μ) do not seem to contribute statistically significant effects.

Based on the random forest results, in a second step, we used the predictors that have discriminatory power, namely, (α) ESC, (β) Construction, (γ) Age, and (λ) Dialect_With, and grew a conditional inference tree to check whether and how these predictors interact with each other. The classification accuracy of this model is 0.43 (with the baseline value being 0.2). The tree and its splits are plotted in Figure 6.

The variables selected for the best split and the corresponding p-values are circled, while the branches specify the levels of the variables. The first split we observe separates declaratives with ESCs (right-hand branch) and without ESCs (left-hand branch). The next split (Node 2) is located in the left-hand branch and divides impersonal *si*, impersonal deontic, weather verb, and epistemic constructions from SPC and extraposition constructions. Node 3 further separates impersonal *si* and impersonal deontic constructions on the one side and weather

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

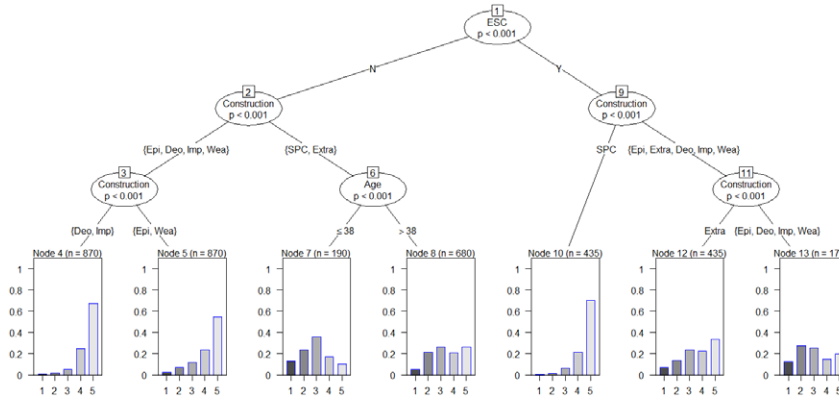

Figure 6
Conditional inference tree of the rating of declaratives (n = 5,220). Impersonal construction type: Deo = impersonal deontic, Epi = epistemic, Extra = extraposition, Imp = impersonal *si*, SPC = subjectless predicative copular, Wea = weather verb.

verb and epistemic constructions on the other. The bar plots at the bottom illustrate the proportion of scores for each point of the Likert scale in each end node that contains all observations for that combination of features. The plots at the bottom show that impersonal *si* and impersonal deontic constructions without ESCs have a high chance of being rated '5 – perfectly acceptable' (Node 4). This chance decreases at Node 5 (weather verb and epistemic constructions) and especially at Node 6 (SPC and extraposition constructions). Node 6 further parts the data based on age: Respondents aged over 38 years are more likely to accept SPC and extraposition constructions without ESCs (Node 8) than younger respondents (Node 7). Moving rightwards, Node 10 shows that SPC constructions are likely to be fully accepted when occurring with ESCs. By contrast, extraposition constructions (Node 12) and particularly the other types of construction (Node 13) have a lower chance of being rated '5 – perfectly acceptable'.

As concerns declaratives, the results of the data obtained through our 2022 survey can be summarized in the following way: (i) SPC and extraposition constructions are more likely to be accepted with ESCs. (ii) Impersonal *si* and impersonal deontic constructions are more likely to be accepted without ESCs. (iii) At the same time, however, SPC and extraposition constructions without ESCs, as well as impersonal *si* and impersonal deontic constructions with ESCs, were not rated as impossible (as a matter of fact, to these types of constructions informants, assigned intermediate rates). (iv) Epistemic and weather verb constructions are more likely to be accepted in the absence of ESCs, that is, they pattern with impersonal *si* and impersonal deontic constructions, although impersonal *si* and impersonal deontic constructions are slightly more likely to be rated as fully acceptable without ESCs than epistemic and weather verb constructions.

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

4.2 *ESCs in interrogatives*

Overall, the acceptability ratios for interrogatives with ESCs, as exemplified in (16a), were higher than for interrogatives without ESCs, as exemplified in (16b). This is shown in Table 5.

(16) Context: interrogative, impersonal *si*, +ESC (a) vs. −ESC (b)
(a) *se* *màgne=o* *anca* *a* *scorza?*
one eats=ESC also the peel
(b) *se* *màgna* *anca* *a* *scorza?*
one eats also the peel
'Do you eat the peel, as well?'

Weather verb and impersonal deontic constructions were accepted both with ESCs ('5–perfectly acceptable' = 0.65 and 0.53, respectively; Figure 7a) and without ESCs ('5–perfectly acceptable' = 0.54 and 0.63, respectively; Figure 7b). Impersonal *si* constructions were preferred when occurring with ESCs ('5–perfectly acceptable' = 0.58; Figure 7a), but they were not rated poorly in the other case, at least compared to the other constructions ('5–perfectly acceptable' = 0.37; Figure 7b).

The lowest scoring construction in stimuli with ESCs is the epistemic one ('5–perfectly acceptable' = 0.27; Figure 7a), as in (17), while the lowest scoring constructions in stimuli without ESCs are the SPC and extraposition constructions ('5–perfectly acceptable' = 0.26 and 0.20, respectively; Figure 7b), exemplified in (18), whose proportion of '5–perfectly acceptable' is higher in the contexts with ESC (SPC: 0.56; extraposition: 0.32; Figure 7a).

(17) Context: interrogative, epistemic, +ESC
pàr=eo *che* *se* *pose?*
seems=ESC that one can
'Does it seem possible?'

(18) Context: interrogative, extraposition, −ESC
é *ciaro* *che* *no* *se* *pol?*
is clear that NEG one can
'Is it clear that it's not possible?'

Context	Likert point	Proportion
Interrogative with ESCs	5–perfectly acceptable	0.49
	4–acceptable	0.24
Interrogative without ESCs	5–perfectly acceptable	0.39
	4–acceptable	0.23

Table 5
ESC acceptability ratios in interrogatives.

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

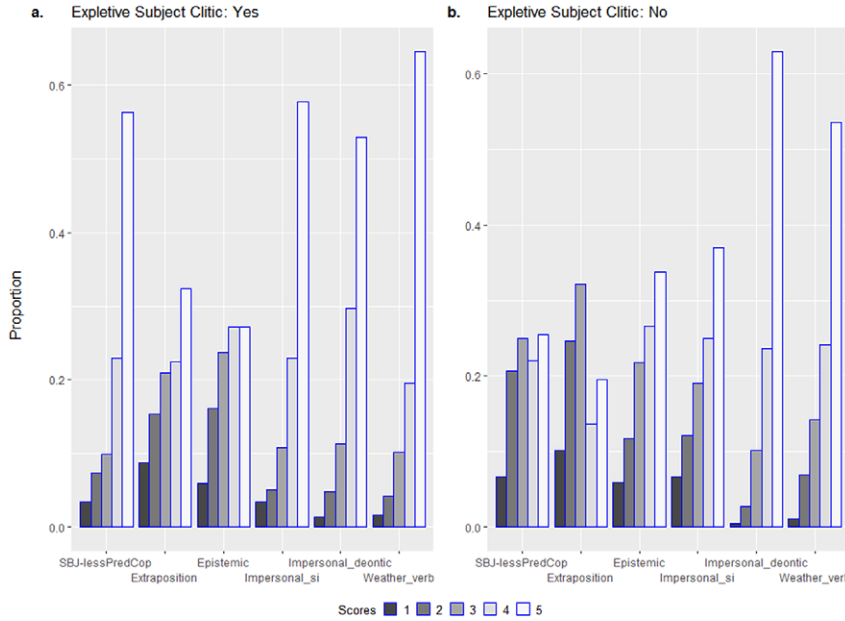

Figure 7
Rating scores across impersonal constructions in interrogatives (n = 5,220). SBJ-lessPredCop stands for subjectless predicative copular (SPC).

Applying the random forest model, the variable importance scores show that (β) 'Construction' (0.064) is the most important predictor, followed by (α) 'ESC' (0.02). Some predictivity can be spotted for (λ) 'Dialect_With' (0.002) and (γ) 'Age' (0.001). The other sociolinguistic predictors (δ, ε, ζ, η, θ, ι, κ, μ) do not seem to contribute statistically significant effects. As for the predictive power of the model, its OOB classification accuracy is 0.58 (with the baseline value being 0.2). The impact of variables is plotted in Figure 8.

Based on the random forest model, we grew a conditional inference tree using the predictors that have discriminatory power (Construction, ESC, Dialect_With, and Age). The tree, whose classification accuracy is 0.45 (with a baseline value of 0.2), is plotted in Figure 9. The first node parts the data set based on the type of construction. On the left-hand branch of the tree, impersonal deontic, weather verb, and impersonal *si* constructions split at Node 2, and the presence or absence of ESCs (Node 6) is likely to impact only the rating of impersonal *si* constructions. Node 3 separates 18-year-old respondent(s) from older respondents. No further split is observed.

Moving rightwards, Node 9 divides epistemic and SPC constructions from extraposition constructions. Eventually, the rating of epistemic, SPC, and extraposition constructions is likely to be modulated by the presence or absence of ESCs (Nodes 10 and 13). The plots at the bottom show that weather verb and impersonal deontic constructions have a high chance of being rated '5–perfectly acceptable',

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

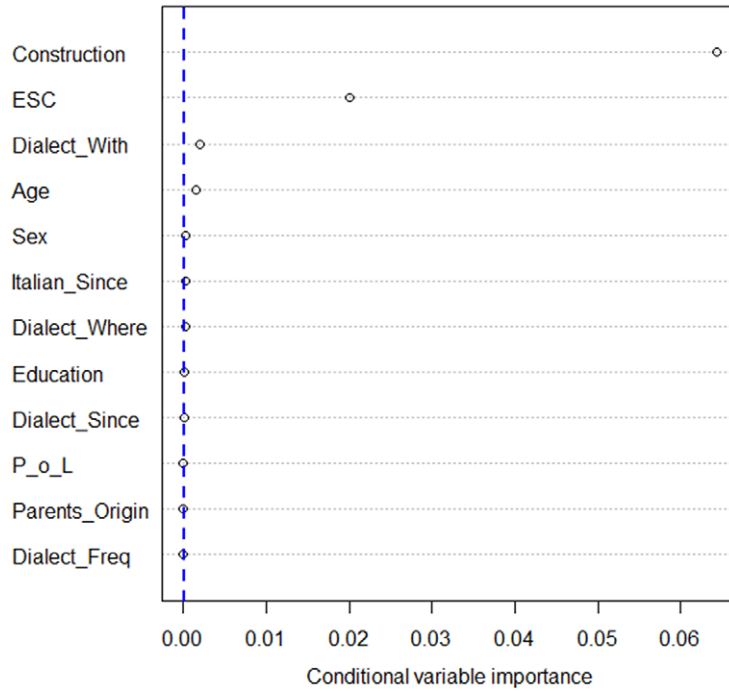

Figure 8
Conditional permutation importance of variables in the rating of interrogatives (n = 5,220). On the vertical axis, Dialect_Freq refers to the frequency with which respondents speak Opitergino. P_o_L refers to the place where respondents live.

regardless of the occurrence of ESCs. This possibility decreases as one moves to the right of the tree since the other constructions are more likely to be evaluated positively when occurring with ESCs.

As concerns interrogatives, the picture we get from the data obtained through our 2022 survey is rather nuanced and can be summarized along the following lines: (i) The presence of ESCs does not impact the acceptability ratings of weather verb and impersonal deontic constructions, which scored the highest. (ii) The presence of ESCs is more likely to positively impact all other constructions, namely, impersonal *si*, epistemic, SPC, and extraposition constructions. (iii) Among the latter, the presence of ESCs seems to impact the ratings of epistemic, SPC, and extraposition constructions more strongly than those of impersonal *si* construction.

5. Analysis of the results

In this section, we compare the results of our survey with the baseline and provide an analysis in terms of maintenance versus change in the Opitergino ESC rule system.

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

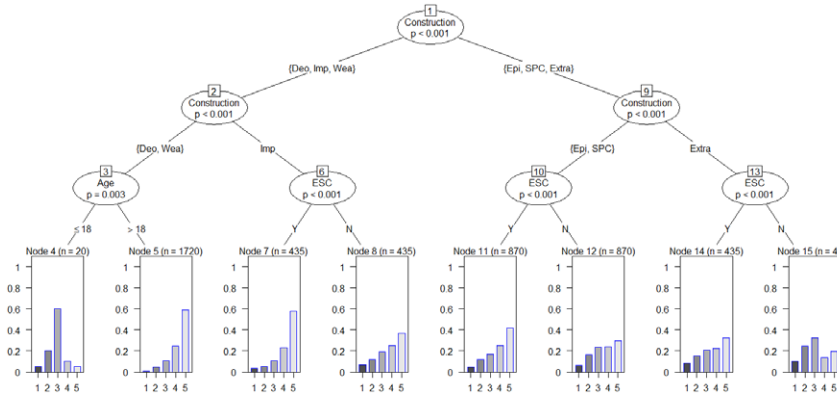

Figure 9
Conditional inference tree of the rating of interrogatives (n = 5,220). Impersonal construction type: Deo = impersonal deontic, Epi = epistemic, Extra = extraposition, Imp = impersonal *si*, SPC = subjectless predicative copular, Wea = weather verb.

We discuss declaratives in Section 5.1, interrogatives in Section 5.2, and the results of our analysis relating to the hypotheses formulated in Section 3.1 in Section 5.3.

5.1 *Change in declaratives*

As concerns declaratives, by comparing the results of our survey with the contexts licensing the occurrence of ESCs in the baseline (cf. Section 3.2), we observe the following trends:

I. In the baseline, ESCs occur mandatorily in SPC and extraposition constructions, and in the 2022 survey, SPC and extraposition constructions with ESCs are strongly preferred. Accordingly, no substantial change has occurred in these contexts. This would point to (a) an OVERALL STABILITY. At the same time, however, stimuli without ESCs are not rated as impossible in the 2022 survey, as one would expect, based on the baseline. This amounts to an increase in the contexts in which the absence of ESCs is accepted (rule generalization) and might point to (b) an ongoing CHANGE. Also, the fact that informants assign intermediate rates to these types of constructions highlights a certain degree of uncertainty on the applicability of the rule.
II. In the baseline, ECSs are optionally used with epistemics. In the 2022 survey, epistemics with ESCs are rather dispreferred. This possibly points to the obliteration of optionality and amounts to (a) an ongoing CHANGE. Besides, stimuli without ESCs are rated favorably, which (b) highlights a certain degree of uncertainty on the applicability of the rule.

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

III. In the baseline, ESCs do not occur in weather verbs, impersonal *si*, and impersonal deontic constructions, and in the 2022 survey, weather verbs, impersonal *si*, and impersonal deontic constructions are preferred without ESCs than with ESCs. Therefore, no substantial change has occurred in these contexts. This would point to (a) an OVERALL STABILITY. However, stimuli with ESCs are not rated as impossible. This amounts to an increase in the contexts in which the presence of ESCs is accepted (rule generalization) and might point to (b) an ongoing CHANGE.

In sum, the analysis reveals a partial match of the 2022 survey with the baseline. First, no substantial rule erosion (loss) has occurred. Rather, the participants in our study did not always make clear-cut judgments about the acceptability of ESCs. For example, weather verbs, impersonal *si*, and impersonal deontics have not been clearly rejected when occurring with ESCs. Second, this uncertainty manifested in terms of both rule generalization (in that an ESC was accepted when occurring in contexts in which it is agrammatical in the baseline) and its omission (in that an ESC was rejected when occurring in contexts in which it is mandatory in the baseline). This change can be understood as rule weakening in terms of decreasing context-sensitivity. Third, the increasing acceptance of the absence of ESCs is likely to be an effect of contact with Italian, as a result of pattern borrowing.

These findings are in line with what Casalicchio & Frasson (2018) report on the use of subject clitics (including ESCs) with weather verbs in Paduan, Trevigiano, Venetian, and Vicentino. The authors observed an overextended use of subject clitics with weather verbs (cf. Section 2.1), but they also noted 'a great deal of uncertainty and instability in speakers who make judgments contrary to the target: in several cases, their use of clitics seems to be optional, as they accept both contexts with clitics and without clitics' (Casalicchio & Frasson 2018: 127; our translation).

Other than Casalicchio & Frasson (2018), however, we analyzed not one but six types of impersonal constructions. While this was not sufficient to yield unambiguous results concerning the change, it still has allowed us to identify properties that might account for a higher degree of acceptability of ESCs. One such property is the fact that both constructions that scored the highest with ESCs (namely, SPC and extraposition) involve copula *é* 'is' of the auxiliary ESSERE 'BE'. The speakers' ratings seem to imply that the presence of a monosyllabic verbal form consisting of one vowel favors the occurrence of ESCs. At the same time, however, the ratings of the two copular constructions diverge significantly, with SPCs scoring much higher than extrapositions. This suggests that an explanation in phonological terms alone probably does not suffice.

The second property concerns processing. What really distinguishes SPC from all other construction types is the easiness with which speakers can identify a deictic, specific, and non-arbitrary subject argument, even though this is unexpressed. As a matter of fact, it has been argued that copular constructions Italian *è bello* 'it is nice' (which we subsume under SPC) have an implicit deictic subject: It could be paraphrased as *qui è bello* (literally 'here is nice') (Salvi 2001, 2: 174–

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

175). This is clearly not the case for the impersonal *si* construction and impersonal deontics, for which a subject argument denoting a specific, non-arbitrary referent can hardly be retrieved (cf. Benincà & Poletto 1994; 1997). Now, in Opitergino, ESCs are rated higher in those impersonal contexts where a deictic subject denoting a non-arbitrary referent can be recovered. From a syntactic viewpoint, this is consistent with the status of Opitergino as a pro-drop language. This would be unexpected in a non-pro-drop language, where an expletive subject pronoun (whether clitic or stressed) is mandatorily expressed and where, therefore, the possibility of tracing a semantic reference is irrelevant.

With respect to the ability to identify a deictic, specific, and non-arbitrary subject, constructions with weather verbs and with epistemics have a status that can be described as intermediate or controversial: Weather verbs are generally referred to in the literature as zero-argument verbs,[18] while in epistemics, an ESC is linked to the argument of the embedded clause.[19] At the moment, however, we are not able to provide any conclusive evaluation of the role that, more general, processing-related factors such as ease of processing might have played in the observed change.

5.2 *Change in interrogatives*

Let's recall that in the Opitergino baseline (cf. Section 3.2), the presence of ESCs in interrogatives is mandatory in constructions of all sorts. The 2022 survey matches this picture as regards the ratings of stimuli with ESCs showing that no substantial rule erosion (loss) has occurred. This points to a STABILITY of the ESC rule system in interrogatives. This stability was further confirmed by the evaluation of those stimuli lacking ESCs in our survey, for which informants provided rather unfavorable ratings. These stimuli included SPCs, extrapositions, and epistemics. However, when asked to judge stimuli lacking ESCs that included weather verbs, impersonal deontics, and (to a lesser extent) impersonal *si*, informants considered them fitting. Overall, this amounts to just one main difference vis-à-vis the baseline: namely, the fact that the omission of ESCs has become tolerable to some extent. These are signs of an ONGOING CHANGE to the effect that, in interrogatives, the absence of ESCs with weather verbs, impersonal *si*, and impersonal deontics is tolerated in some measure.

5.3 *Trends of change and hypotheses testing*

In this section, we provide an assessment of the development of the overall ESC system in Opitergino and test the scenario of contact-induced change.

[18] Admittedly, some scholars have claimed of the existence of a 'semi-argument' linked to the thematic role of theme (Pescarini 2014: 239; 2015: 70–72; Puglielli & Frascarelli 2008: 111–113).

[19] This is evident in the corresponding raising structures (e.g. 'it seems that the kids are doing well' corresponds to 'the kids seem to be doing well') (see, among others, Postal 1974).

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

The composite picture we get by combining the results for declaratives and interrogatives allows us to observe that: (1) the Opitergino ESC system is a relatively stable rule system; (2) at the same time, it displays a quite high degree of intersubjective variation and uncertainty in the informants' judgments (which is particularly evident in declaratives); and (3) in spite of this overall stability, a thread of change is recognizable.

The fact that a thread of change is recognizable leads us to test the hypotheses made in Section 3.1. First, we hypothesized (H1) an effect of contact with standard Italian. That change has been ongoing in Opitergino is a fact. Whether this change has been driven by contact cannot be determined with absolute certainty. However, the speakers' increasing acceptance of the absence of ESCs is a strong clue in support of the contact hypothesis. The effect of contact between Opitergino and standard Italian is most patent in interrogatives where ESC is baseline-obligatory across all construction types.

Second, we hypothesized (H2) that contact would produce slight structural convergence towards standard Italian but would not result in the loss of ESCs. This expectation is largely backed by the observation that no substantial rule erosion has occurred. However, the increasing tolerability of the omission of ESCs, while moderate, is a symptom of an ongoing erosion.

Third, we hypothesized (H3) that change would affect ESCs in the domain of declarative clauses to a greater extent than in interrogative clauses. As a matter of fact, rule weakening has affected especially declaratives: Other than in interrogatives, where an ESC is baseline-obligatory, in declaratives, the rules' context-sensitivity makes the situation more complex, and this has produced a change of different types (obliteration of optionality and rule generalization). This backs our hypothesis.

In addition, as regards ESCs in interrogatives, we noted that the omission of ESCs in interrogatives is rated favorably in precisely the same impersonal contexts in which the occurrence of ESCs in declaratives is rated as unfitting in our survey. These contexts are weather verbs, impersonal *si*, and impersonal deontic. The fact that informants gave higher scores to these constructions than to all other constructions implies that they generally accept them well, both with and without ESCs. This picture suggests that the ongoing erosion does not proceed randomly but mirrors the use of ESCs in declaratives. Given the fact that Opitergino and Italian maximally diverge relating to subject clitics and ESCs, the ways which change manifests in the recipient language (Opitergino) cannot be entirely explained by resorting to the system of the source language (Italian). Otherwise, we would only observe the loss of the ESC system, which we have not. Therefore, while contact most likely has led to change, the partial and ongoing alignment between ESCs in declaratives and ESCs in interrogatives can plausibly be ascribed to a secondary restructuring, which is internal to the Opitergino system and not due to contact.

In sum, we showed that syntactic change can manifest as decreasing context-sensitivity, leading to rule weakening. In Opitergino, this is observable both in the

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

domain of declaratives – where change is contact-induced – and in the domain of interrogatives – where change is internally driven.

6. Conclusion

This paper explored the occurrence of the expletive subject clitics in Opitergino, a Venetan variety spoken in Oderzo. By means of an extensive online survey, we studied the acceptability of ESCs in six types of impersonal constructions in declaratives and interrogatives. We observed a high degree of intersubjective variation across sociolinguistic predictors. We aimed to understand whether contact with standard Italian has had an impact on Opitergino and put forward three hypotheses. To test whether change has occurred and to what extent, we compared the results of our 2022 survey with the baseline rules we extracted from speakers born before 1942, whose L1 is Opitergino and L2 is Italian. We observed that while the system is overall stable, a thread of change is ongoing and manifest in (a) rule weakening in declaratives and (b) erosion of the obligatoriness of ESCs in interrogatives. We argued that this change is likely to be an effect of contact, resulting in structural convergence but not loss, and affected the part of the ESC system that features more optionality, namely, the domain of declarative clauses. These results supported our hypotheses.

Our study innovatively combines dialectological with language contact research. Theoretically, it contributes to research on syntactic change, addressing the role of context-sensitivity of rules; methodologically, it contributes to the analysis of microvariation by resorting to a pioneering application of data elicitation and statistical methods. One of the main findings of this study was to bring out the role of context-sensitivity in syntactic change. We showed that change can manifest as decreasing context sensitivity, leading to rule weakening. This study is suitable for replication, *mutatis mutandis*, in other areas and for subsequent comparison to control for the role of different sociolinguistic settings.

Acknowledgments

We are grateful to all participants in the survey for their valuable contribution to this study. We would also like to thank Ms. Adelina Bonacin for her help in reviewing the stimuli and Mayor Maria Scardellato, Monsignor Pierpaolo Bazzichetto, and Stefano Rui as well as the press (Maria Teresa Tolotto for *Il Dialogo* and Andrea Pizzinat for *L'Azione*) for helping us promote the survey. Finally, heartful thanks to GB's grandparents Albina, Odetta, Niccolò, Pier Giorgio, and Santo, who patiently answered hundreds of questions to establish the baseline – two of them recently passed away, but something of them still lives on.

All three authors contributed to the research design. CZ designed the questionnaire and was responsible for data collection and analysis. GB selected the stimuli, created the audio files, and promoted the investigation in Oderzo. FG initiated the research and supervised all phases of the work. FG and CZ wrote the paper jointly.

All authors provided feedback on the draft and approved the final version of the manuscript.

REFERENCES

Adamou, Evangelia & Xingjia Rachel Shen. 2019. There are no language switching costs when codeswitching is frequent. *International Journal of Bilingualism* 23.1, 53–70.

Agresti, Alan. 2019. *An introduction to categorical data analysis*. Hoboken, NJ: Wiley.

Benincà, Paola. 2007. Clitici e ausiliari: Gh ò, z é. *The Italianist* 27, 27–47.

Benincà, Paola & Cecilia Poletto. 1994. Bisogna and its companions: The verbs of necessity. In Guglielmo Cinque, Jan Koster, Jean-Yves Pollock, Luigi Rizzi & Raffaella Zanuttin (eds.), *Paths towards Universal Grammar. Studies in honor of Richard S. Kayne*, 35–58. Washington: Georgetown University Press.

Benincà, Paola & Cecilia Poletto. 1997. The diachronic development of a modal verb of necessity. In Ans van Kemenade & Nigel Vincent (eds.), *Parameters of morphosyntactic change*, 94–118. Cambridge: Cambridge University Press.

Benincà, Paola & Cecilia Poletto. 2004. Topic, focus and V2: Defining the CP sublayers. In Luigi Rizzi (ed.), *The structure of CP and IP*, 52–75. Oxford: Oxford University Press.

Bentley, Delia. 2017. Copular and existential constructions. In Andreas Dufter & Elisabeth Stark (eds.), *Manual of Romance morphosyntax and syntax*, 332–366. Berlin: De Gruyter.

Berizzi, Mariachiara. 2012. Toccare come verbo deontico nei dialetti Italiani. *Quaderni di Lavoro ASIt* 14, 191–208.

Brandi, Luciana & Patrizia Cordin. 1989. Two Italian dialects and the null subject parameter. In Osvaldo A. Jaeggli & Kenneth J. Safir (eds.), *The null subject parameter*, 111–142. Dordrecht: Kluwer.

Casalicchio, Jan & Alberto Frasson. 2018. Cambiamenti nell'uso dei soggetti clitici veneti: Il ruolo del contatto con l'Italiano. In Annamaria Chilà & Alessandro De Angelis (eds.), *Capitoli di morfosintassi delle varietà romanze d'Italia: Teoria e dati empirici*, 117–133. Palermo: Centro di Studi Filologici e Linguistici Siciliani.

Cennamo, Michela. 1993. *The reanalysis of reflexives: A diachronic perspective*. Naples: Liguori.

Cennamo, Michela. 2016. Voice. In Adam Ledgeway & Martin Maiden (eds.), *The Oxford guide to the Romance languages*, 967–980. Oxford: Oxford University Press.

Cerruti, Massimo. 2011. Regional varieties of Italian in the linguistic repertoire. *International Journal of the Sociology of Language* 210, 9–28.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

D'Alessandro, Roberta. 2007. *Impersonal* si *constructions: Agreement and interpretation*. Berlin: Mouton de Gruyter.

Davanzo, Giacomo. 2016. *Il dialetto di Ceggia. Analisi fono-morfologica e sintattica.* Ph.D. dissertation, Ca' Foscari University of Venice.

De Mauro, Tullio. 2014. *Storia linguistica dell'Italia repubblicana: Dal 1946 ai nostri giorni*. Rome: Laterza.

Ferrarotti, Lorenzo. 2019. Selection and morphology of expletive subject clitics in northern Italian dialects. In Silvio Cruschina, Adam Ledgeway & Eva-Maria Remberger (eds.), *Italian dialectology at the interfaces*, 101–112. Amsterdam: John Benjamins.

Franzin, Fabio. 2013. *Bestie e stranbi: Nel dialetto veneto-trevigiano dell'Opitergino-Mottense*. Martinsicuro: Di Felice.

Frasson, Alberto, Roberta D'Alessandro & Brechje van Osch. 2021. Subject clitics in microcontact: A case study from heritage Friulian in Argentina and Brazil. *Heritage Language Journal* 18.1, 1–36.

Gardani, Francesco. 2008. *Borrowing of inflectional morphemes in language contact*. Frankfurt: Peter Lang.

Gardani, Francesco. 2012. Plural across inflection and derivation, fusion and agglutination. In Lars Johanson & Martine Robbeets (eds.), *Copies versus cognates in bound morphology*, 71–97. Leiden & Boston: Brill.

Gardani, Francesco. 2020. Borrowing matter and pattern in morphology. An overview. *Morphology* 30.4, 263–282.

Gardani, Francesco. 2022. Contact and borrowing. In Adam Ledgeway & Martin Maiden (eds.), *The Cambridge handbook of Romance linguistics*, 845–869. Cambridge: Cambridge University Press.

Harris, Alice C. & Lyle Campbell. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.
Heath, Jeffrey. 1984. Language contact and language change. *Annual Review of Anthropology* 13, 367–384.
Hothorn, Torsten, Peter Buehlmann, Sandrine Dudoit, Annette Molinaro and Mark Van Der Laan. 2006. Survival ensembles. *Biostatistics* 7, 355–373.
ISTAT. 2017. *Anno 2015. L'uso della lingua Italiana, dei dialetti e delle lingue straniere*. Rome: Istituto Nazionale di Statistica.
Ivanova-Sullivan, Tanya, Irina A. Sekerina, Davood Tofighi & Maria Polinsky. 2022. Language-internal reanalysis of clitic placement in heritage grammars reduces the cost of computation: Evidence from Bulgarian. *Languages* 7.1, 24.
Jaberg, Karl & Jakob Jud. 1928. *Der Sprachatlas als Forschungsinstrument. Kritische Grundlegung und Einführung in den Sprach- und Sachatlas Italiens und der Südschweiz*. Halle: Niemeyer.
Janda, Laura A. & Anna Endresen. 2017. Five statistical models for Likert-type experimental data on acceptability judgments. *Journal of Research Design and Statistics in Linguistics and Communication Science* 3.2, 217–250.
Klee, Carol A. 1990. Spanish-Quechua language contact: The clitic pronoun system in Andean Spanish. *Word* 41.1, 35–46.
Labov, William. 1963. The social motivation of a sound change. *Word* 19, 273–309.
Levshina, Natalia. 2021. Conditional inference trees and random forests. In Magali Paquot & Stefan Th. Grice (eds.), *A practical handbook of corpus linguistics*, 611–643. Cham: Springer.
Loporcaro, Michele. 2012. A new strategy for progressive marking and its implications for grammaticalization theory: The subject clitic construction of Pantiscu. *Studies in Language* 36.4, 747–784.
Manzini, Maria Rita & Leonardo Maria Savoia. 2005. *I dialetti Italiani e Romanci. Morfosintassi generativa*. Alessandria: Edizioni dell'Orso.
Marian, Viorica, Henrike K. Blumenfeld & 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, 940–967.
Matras, Yaron & Jeanette Sakel. 2007. Investigating the mechanisms of pattern replication in language convergence. *Studies in Language* 31.4, 829–865.
Meakins, Felicity, Jennifer Green & Myfany Turpin. 2018. *Understanding linguistic fieldwork*. London: Routledge.
Montrul, Silvina. 2010. How similar are adult second language learners and Spanish heritage speakers? Spanish clitics and word order. *Applied Psycholinguistics* 31.1, 167–207.
Papanti, Giovanni. 1875. *I parlari Italiani in certaldo alla festa del V centenario di messer Giovanni Boccacci: Omaggio*. Livorno: Francesco Vigo.
Parry, Mair M. 1994. Piedmontese subject clitics: A diachronic perspective. *Vox Romanica* 52, 96–116.
Perlmutter, David M. 1968. *Deep and surface structure constraints in syntax*. Ph.D. dissertation, Massachusetts Institute of Technology.
Pescarini, Diego. 2012. Clitici espletivi nei dialetti settentrionali: I dati dell'Atlante Sintattico d'Italia (ASIt). *Atti del Sodalizio Glottologico Milanese* 7, 42–56.
Pescarini, Diego. 2014. La distribuzione dei clitici soggetto espletivi: Tipologia e prospettive parametriche. *L'Italia Dialettale* 75, 229–246.
Pescarini, Diego. 2015. *Le costruzioni con «si». Italiano, dialetti e lingue Romanze*. Rome: Carocci.
Pescarini, Diego. 2016. Clitic pronominal systems: Morphophonology. In Adam Ledgeway & Martin Maiden (eds.), *The Cambridge handbook of Romance linguistics*, 742–757. Cambridge: Cambridge University Press.
Pescarini, Diego. 2022. Expletive subject clitics in northern Italo-Romance. *Languages* 7, 265.
Pescarini, Diego & Michele Loporcaro. 2022. Variation in Romance. In Adam Ledgeway & Martin Maiden (eds.), *The Cambridge handbook of Romance linguistics*, 150–180. Cambridge: Cambridge University Press.
Poletto, Cecilia & Christina Tortora. 2016. Subject clitics: Syntax. In Adam Ledgeway & Martin Maiden (eds.), *The Oxford guide to the Romance languages*, 772–785. Oxford: Oxford University Press.
Postal, Paul M. 1974. *On raising*. Cambridge, MA: MIT Press.
Puglielli, Annarita & Mara Frascarelli. 2008. *L'analisi linguistica. Dai dati alla teoria*. Cesena: Caissa.
R Core Team. 2020. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. URL https://www.R-project.org/.

Renzi, Lorenzo & Laura Vanelli. 1983. I pronomi soggetto in alcune varietà Romanze. In Paola Benincà, Manlio Cortelazzo, Aldo Luigi Prosdocimi, Laura Vanelli & Alberto Zamboni (eds.), *Scritti linguistici in onore di Giovan Battista Pellegrini*, vol. 1, 121–145. Pisa: Pacini.
Rohlfs, Gerhard. 1968. *Grammatica storica della lingua Italiana e dei suoi dialetti, II*: *Morfologia*. Turin: Einaudi.
Sakel, Jeanette. 2007. Types of loan: Matter and pattern. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 15–29. Berlin: Mouton de Gruyter.
Salvi, Giampaolo. 2001. Le frasi copulative. In Lorenzo Renzi, Giampaolo Salvi & Anna Cardinaletti (eds.), *Grande grammatica Italiana di consultazione*, vol. 2, 163–190. Bologna: il Mulino.
Souag, Lameen. 2017. Clitic doubling and language contact in Arabic. *Zeitschrift für arabische Linguistik* 66, 45–70.
Strobl, Carolin, Anne-Laure Boulesteix, Achim Zeileis & Torsten Hothorn. 2007. Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics* 8, 25.
Strobl, Carolin, Anne-Laure Boulesteix, Thomas Kneib, Thomas Augustin & Achim Zeileis. 2008. Conditional variable importance for random forests. *BMC Bioinformatics* 9, 307.
Strobl, Carolin, James Malley & Gerhard Tutz. 2009. An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods* 14.4, 323–348.
Tagliamonte, Sali A. & R. Harald Baayen. 2012. Models, forests, and trees of York English: *Was*/*were* variation as a case study for statistical practice. *Language Variation and Change* 24.2, 135–178.
Thomas, Danielle L. T. 2012. *Grammatical optionality and variability in bilingualism: How Spanish-English bilinguals limit clitic-climbing*. Ph.D. dissertation, University of Toronto.
Thomason, Sarah G. 2000. On the unpredictability of contact effects. *Sociolinguistic Studies* 1.1, 173–182.
Thomason, Sarah G. 2001. *Language contact: An introduction*. Edinburgh: Edinburgh University Press.
Vanelli, Laura. 1987. I pronomi soggetto nei dialetti Italiani settentrionali dal Medio Evo a oggi. *Medioevo Romanzo* 12, 173–211.
Veríssimo, João. 2021. Analysis of rating scales: A pervasive problem in bilingualism research and a solution with Bayesian ordinal models. *Bilingualism: Language and Cognition* 24.5, 842–848.
Weinreich, Uriel. 1953. *Languages in contact, findings and problems*. New York: Linguistic Circle of New York.
Zamboni, Alberto. 1974. *Veneto*. Pisa: Pacini.
Zamboni, Alberto. 1979. Le caratteristiche essenziali dei dialetti Veneti. In Manlio Cortelazzo (ed.), *Guida ai dialetti Veneti*, vol. 1, 9–43. Padua: Cleup.

Appendix

The Appendix lists all the stimuli used. Stimuli are grouped by type of construction. The order in which the stimuli appear in the Appendix was not randomized; as such, it does not reflect the order in which the stimuli were administered to the informants.

An expletive subject clitic (ESC) is marked by '='. The lack of ESC is marked by '*Ø*'.

Stimuli: Declaratives

Stimuli		Type of construction	ESC
1.	*Sèra e fenèstre! Ø Piove!* 'Close the windows! It's raining!'	Weather verbs	N
2.	*Varda fora: Ø névega!* 'Look outside: it's snowing!'	Weather verbs	N
3.	*No serve bagnar: Ø tonedéa.* 'No need to water [the garden]: it's thundering.'	Weather verbs	N
4.	*No sta 'ndar: Ø névega!* 'Don't go: it's snowing!'	Weather verbs	N
5.	*Mi stae casa, Ø piove.* 'I'll stay at home, it's raining.'	Weather verbs	N
6.	*In te sto posto, Ø par che se magne ben.* 'In this place, it seems like the food's good.'	Epistemic	N
7.	*Scolta… Ø Par che a vegne.* 'Listen… it seems like she's coming.'	Epistemic	N
8.	*A Lina, Ø par che a stae ben…* 'Lina seems to be fine.'	Epistemic	N
9.	*No ò fat: Ø par che no se pose.* 'I haven't done it: it seems that there's no way to do it.'	Epistemic	N
10.	*No savarìe dirte, Ø par de sì.* 'I can't tell, it seems like this.'	Epistemic	N
11.	*Sèra e fenèstre! El=piove!* 'Close the windows! It's raining!'	Weather verbs	Y
12.	*Varda fora: el=névega!* 'Look outside: it's snowing!'	Weather verbs	Y
13.	*No serve bagnar: el=tonedéa.* 'No need to water [the garden]: it's thundering.'	Weather verbs	Y
14.	*No sta 'ndar: el=névega!* 'Don't go: it's snowing!'	Weather verbs	Y
15.	*Mi stae casa, el=piove.* 'I'll stay at home, it's raining.'	Weather verbs	Y
16.	*In te sto posto, el=par che se magne ben.* 'In this place, it seems like the food's good.'	Epistemic	Y
17.	*Scolta… El=par che a vegne.* 'Listen… it seems like she's coming.'	Epistemic	Y
18.	*A Lina, el=par che a stae ben…* 'Lina seems to be fine.'	Epistemic	Y
19.	*No ò fat: el=par che no se pose.* 'I haven't done it: it seems that there's no way to do it.'	Epistemic	Y

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

Continued

Stimuli		Type of construction	ESC
20.	*No savarìe dirte, el=par de sì.* 'I can't tell, it seems like this.'	Epistemic	Y
21.	*Magnén qua: Ø é un bel posto!* 'Let's eat here: it's a nice place!'	SPC	N
22.	*No se pol. Ø É scuro!* 'It isn't possible. It's (already) darkened!'	SPC	N
23.	*Ades basta. Ø É tardi.* 'Enough now. It's late.'	SPC	N
24.	*Varda: Ø é 'na toséta!* 'Look: there's a little girl!'	SPC	N
25.	*Scolta: Ø é zà do boti.* 'Listen: it's already two o'clock.'	SPC	N
26.	*A Maria, Ø é ciaro che a vien.* 'It's clear that Maria is not coming.'	Extraposition	N
27.	*Scólteme! Ø É ciaro che se fa cusì.* 'Listen! It's clear that it has to be done like this.'	Extraposition	N
28.	*A nona, poréta, Ø é ciaro che a sta mal.* 'It's clear that grandmother is ill—poor thing.'	Extraposition	N
29.	*Tasi! Ø É ciaro che i vol cusì!* 'Shut up! It's clear that they want it this way!'	Extraposition	N
30.	*Brute robe. Ø é ciaro che no se e fa.* 'Bad things. It's clear that one must not do them.'	Extraposition	N
31.	*Magnén qua: l'=é un bel posto!* 'Let's eat here: it's a nice place!'	SPC	Y
32.	*No se pol. L'=é scuro!* 'It isn't possible. It's (already) darkened!'	SPC	Y
33.	*Ades basta. L'=é tardi.* 'Enough now. It's late.'	SPC	Y
34.	*Varda: l'=é 'na toséta!* 'Look: there's a little girl!'	SPC	Y
35.	*Scolta: l'=é zà do boti.* 'Listen: it's already two o'clock.'	SPC	Y
36.	*A Maria, l'=é ciaro che a vien.* 'It's clear that Maria is not coming.'	Extraposition	Y
37.	*Scólteme! L'=é ciaro che se fa cusì.* 'Listen! It's clear that it has to be done like this.'	Extraposition	Y
38.	*A nona, poréta, l'=é ciaro che a sta mal.* 'It's clear that grandmother is ill—poor thing.'	Extraposition	Y
39.	*Tasi! L'=é ciaro che i vol cusì!* 'Shut up! It's clear that they want it this way!'	Extraposition	Y
40.	*Brute robe. L'=é ciaro che no se e fa.* 'Bad things. It's clear that one must not do them.'	Extraposition	Y
41.	*Son secùra. Ø Se dise cusì.* 'I'm sure. One says like this.'	Impersonal *si*	N
42.	*Ho capìo. Ø Se fa tut come prima.* 'I understand. We do it as we've always done.'	Impersonal *si*	N

(*Continued*)

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

Continued

Stimuli		Type of construction	ESC
43.	*Va ben. Ø Se fa cusì da tant.* 'Okay. We've done it like this for long.'	Impersonal *si*	N
44.	*Sa vutu… Ø se dise tante robe…* 'You know … one says many things…'	Impersonal *si*	N
45.	*Basta! Ø Se fa cusì!* 'Enough! That's how we got to do it!	Impersonal *si*	N
46.	*Tosàti, Ø toca 'ndar.* 'Guys, we got to go.'	Impersonal deontic	N
47.	*No so. Ø Bisogna domandarghe.* 'I don't know. We need to ask him/her/them.'	Impersonal deontic	N
48.	*Dai mama, Ø bisogna ciamàrli.* 'Come on, mom, we need to call them.'	Impersonal deontic	N
49.	*Poréta! Ø Toca ciamàr el dotor.* 'She, poor thing! We need to call the doctor.'	Impersonal deontic	N
50.	*Spèta. Ø Toca a mì 'des.* 'Wait. It's my turn now.'	Impersonal deontic	N
51.	*Son secùra. El= se dise cusì.* 'I'm sure. It one says like this.'	Impersonal *si*	Y
52.	*Ho capìo. El= se fa tut come prima.* 'I understand. We do it as we've always done.'	Impersonal *si*	Y
53.	*Va ben. El= se fa cusì da tant.* 'Okay. We've done it like this for long.'	Impersonal *si*	Y
54.	*Sa vutu… el= se dise tante robe…* 'You know … one says many things…'	Impersonal *si*	Y
55.	*Basta! El= se fa cusì!* 'Enough! That's how we got to do it!'	Impersonal *si*	Y
56.	*Tosàti, el= toca 'ndar.* 'Guys, we need to go.'	Impersonal deontic	Y
57.	*No so. El= bisogna domandarghe.* 'I don't know. We need to ask him/her/them.'	Impersonal deontic	Y
58.	*Dai mama, el= bisogna ciamàrli.* 'Come on, mom, we need to call them.'	Impersonal deontic	Y
59.	*Poréta! El= toca ciamàr el dotor.* 'She, poor thing! We need to call the doctor.'	Impersonal deontic	Y
60.	*Spèta. El= toca a mì 'des.* 'Wait. It's my turn now.'	Impersonal deontic	Y

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

STIMULI: INTERROGATIVES

Stimuli		Type of construction	ESC
1.	*PioveØ fora?* 'Is it raining outside?'	Weather verbs	N
2.	*NévegaØ anca doman?* 'Is it snowing tomorrow, too?'	Weather verbs	N
3.	*Che scuro! TonedéaØ anca?* 'So dark! Is it thundering, too?'	Weather verbs	N
4.	*Parché atu serà? PioveØ?* 'Why did you close [the door/window]? Is it raining?'	Weather verbs	N
5.	*NévegaØ o pioveØ?* 'Is it snowing or is it raining?'	Weather verbs	N
6.	*Che bèa: éØ 'na toséta?* 'So beautiful: is it a girl?'	SPC	N
7.	*Ditu che? ÉØ masa tardi?* 'What do you mean? Is it too late?'	SPC	N
8.	*Dime, éØ anca a Maria?* 'Tell me, is Maria there, too?'	SPC	N
9.	*Bondì sióra, éØ Toni?* 'Good morning, ma'am, is Toni there?'	SPC	N
10.	*Spèta, éØ zà un bot?* 'Wait, is it already one o'clock?'	SPC	N
11.	*Ma parØ sol che a mì che a sìe cusì?* 'Am I the only one that thinks that it's like this?'	Epistemic	N
12.	*Ditu che? ParØ che a vegne?* 'What do you mean? Does it seem like she's coming?'	Epistemic	N
13.	*ParØ zà che i vae via?* 'Does it seem like they're already leaving?'	Epistemic	N
14.	*ParØ che se pose?* 'Does it seem possible?'	Epistemic	N
15.	*Scolta, parØ che a stàe qua?* 'Listen, does it seem like she lives here?'	Epistemic	N
16.	*Se faØ cusì?* 'Is it done this way?'	Impersonal *si*	N
17.	*Se diseØ come 'sta roba?* 'How do you say it?'	Impersonal *si*	N
18.	*Se magnaØ anca a scorza?* 'Do you eat the peel, as well?'	Impersonal *si*	N
19.	*Se diseØ sempre e stese robe?* 'Do they always say the same things? (*lit.* One says…)'	Impersonal *si*	N
20.	*Se faØ come?* 'How do you do it? (*lit.* One does how?)'	Impersonal *si*	N
21.	*BisognaØ far cusì?* 'Does it have to be done like this?'	Impersonal deontic	N
22.	*TocaØ a mì?* 'Is it my turn?'	Impersonal deontic	N

(*Continued*)

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

Continued

Stimuli		Type of construction	ESC
23.	*TocaØ propio 'ndar?* 'Do we really need to go?'	Impersonal deontic	N
24.	*TocaØ fàrghine 'ncora?* 'Do we need to do more?'	Impersonal deontic	N
25.	*BisognaØ cavàrla via?* 'Do we need to remove it?'	Impersonal deontic	N
26.	*Scolta, éØ ciaro che no se pol?* 'Listen, is it clear that it's not possible?'	Extraposition	N
27.	*E ora, éØ ciaro che se fa cusì?* 'So, is it clear that it has to be done this way?'	Extraposition	N
28.	*Varda qua, éØ ciaro che no a vol?* 'Look, is it clear she doesn't want?'	Extraposition	N
29.	*Spèta, éØ ciaro che i vien co tì?* 'Wait, is it clear that they'll come with you?'	Extraposition	N
30.	*Scólteme, éØ ciaro che sta roba no a va?* 'Listen, is it clear that this thing doesn't work?'	Extraposition	N
31.	*Pióve=o fòra?* 'Is it raining outside?'	Weather verbs	Y
32.	*Néveghe=o anca doman?* 'Is it snowing tomorrow, too?'	Weather verbs	Y
33.	*Che scuro! Tonedé=o anca?* 'So dark! Is it thundering, too?'	Weather verbs	Y
34.	*Parché atu serà? Pióve=o?* 'Why did you close [the door/window]? Is it raining?'	Weather verbs	Y
35.	*Néveghe=o o pióve=o?* 'Is it snowing or is it raining?'	Weather verbs	Y
36.	*Che bèa: é=a 'na toséta?* 'So beautiful: is it a girl?'	SPC	Y
37.	*Ditu che? É=o masa tardi?* 'What do you mean? Is it too late?'	SPC	Y
38.	*Dime, é=o anca a Maria?* 'Tell me, is Maria there, too?'	SPC	Y
39.	*Bondì sióra, é=o Toni?* 'Good morning, ma'am, is Toni there?'	SPC	Y
40.	*Spèta, é=o zà un bot?* 'Wait, is it already one o'clock?'	SPC	Y
41.	*Ma pàr=eo sol che a mì che a sìe cusì?* 'Am I the only one that thinks that it's like this?'	Epistemic	Y
42.	*Ditu che? Pàr=eo che a vegne?* 'What do you mean? Does it seem like she's coming?'	Epistemic	Y
43.	*Pàr=eo zà che i vae via?* 'Does it seem like they're already leaving?'	Epistemic	Y
44.	*Pàr=eo che se pose?* 'Does it seem possible?'	Epistemic	Y
45.	*Scolta, pàr=eo che a stàe qua?* 'Listen, does it seem like she lives here?'	Epistemic	Y

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press

Continued

Stimuli		Type of construction	ESC
46.	*Se fà=eo cusì?* 'Is it done this way?'	Impersonal *si*	Y
47.	*Se dìse=o come 'sta roba?* 'How do you say it?'	Impersonal *si*	Y
48.	*Se màgne=o anca a scorza?* 'Do you eat the peel, as well?'	Impersonal *si*	Y
49.	*Se dìse=o sempre e stese robe?* 'Do they always say the same things? (*lit.* One says…)'	Impersonal *si*	Y
50.	*Se fà=eo come?* 'How do you do it?'	Impersonal *si*	Y
51.	*Bisògne=o far cusì?* 'Does it have to be done like this?'	Impersonal deontic	Y
52.	*Tóche=o a mì?* 'Is it my turn?'	Impersonal deontic	Y
53.	*Tóche=o propio 'ndar?* 'Do we really need to go?'	Impersonal deontic	Y
54.	*Tóche=o fàrghine 'ncora?* 'Do we need to do more?'	Impersonal deontic	Y
55.	*Bisògne=o cavàrla via?* 'Do we need to remove it?'	Impersonal deontic	Y
56.	*Scolta, é=o ciaro che no se pol?* 'Listen, is it clear that it's not possible?'	Extraposition	Y
57.	*E ora, é=o ciaro che se fa cusì?* 'So, is it clear that this has to be done this way?'	Extraposition	Y
58.	*Varda qua, é=o ciaro che no a vol?* 'Look, is it clear she doesn't want?'	Extraposition	Y
59.	*Spèta, é=o ciaro che i vien co tì?* 'Wait, is it clear that they'll come with you?'	Extraposition	Y
60.	*Scólteme, é=o ciaro che sta roba no a va?* 'Listen, is it clear that this thing doesn't work?'	Extraposition	Y

Authors' addresses: *(Zanini)*
Institute of Romance Studies, University of Zurich, Switzerland

(Battistella)
Institute of Romance Studies, University of Zurich, Switzerland

(Gardani)
Institute of Romance Studies, University of Zurich, Switzerland

https://doi.org/10.1017/S0022226723000282 Published online by Cambridge University Press