

Clitic syntax in areal linguistics: Pinning down the role of language contact via crowdsourcing

Romano Madaro

University of Verona

 ORCID ID: 0000-0001-5299-7760

romano.madaro@univr.it

Michele Cosentino

University of Messina

 ORCID ID: 0009-0002-9676-8577

m.cosentino@unime.it

Ermenegildo Bidese

University of Trento

 ORCID ID: 0000-0002-9521-3089

ermenegildo.bidese@unitn.it

Simone Barco

University of Trento

simone.barco@unitn.it



How to cite: Madaro, Romano, Cosentino, Michele, Bidese, Ermenegildo & Simone Barco. 2025.

Critic syntax in areal linguistics: Pinning down the role of language contact via crowdsourcing In *Italian dialects and language contact: Theory, typology, data*, eds. Mario Wild, Stefano Negrinelli, Federica Breimaier, Stefano Cristelli, Michele Loporcaro & Tania Paciaroni. Special issue of *Isogloss. Open Journal of Romance Linguistics* 11(6)/6, 1-24.

DOI: <https://doi.org/10.5565/rev/isogloss.514>

Abstract

In this study, we present the results of a linguistic survey conducted using *AlpiLinK*, a corpus of crowdsourced data, to highlight the potential of such a tool when dealing with areal contact. The data concern the syntax of clitics and prodrop phenomena in the Italo-Romance varieties spoken in North-Eastern Italy. We analyze (a) the

dropping of the 2nd person singular clitic in Gardenese Ladin and (b) the distribution of the 3rd person singular clitic in constructions with weather verbs in the Italo-Romance varieties of Northeastern Italy, in particular, Venetan, Trentino and Friulian, using a dataset of over 250 localities. The results demonstrate the potential of *AlpiLinK* in both qualitative and quantitative research, allowing us to draw some conclusions about the dynamics of Germanic-Romance language contact in this area.

Keywords: areal linguistics, language contact, Italo-Romance, crowdsourcing, Gardenese Ladin.

1. Introduction: Investigating linguistic variation through crowdsourcing

AlpiLinK (*Alpine Languages in Contact*)¹ is a research project developed in collaboration between the Universities of Verona, Trento, Bolzano-Bozen, Turin, and Aosta Valley and funded by the Italian Research Ministry as ‘Research Projects of National Relevant Interest’ (PRIN) for the period 2022-2025. It aims to explore, describe, and promote the Germanic, Romance, and Slavic varieties spoken across the Italian Alpine Space. Like its forerunner *VinKo* (*Varieties in Contact*),² which was although restricted to the Northeast of Italy, and other recent projects,³ *AlpiLinK* is based on the crowdsourcing of linguistic data. This means that any speaker of the investigated regional and minority varieties (e.g. Trentino dialects, Ladin, Cimbrian or Mòcheno) can register, independently and at any time, on the platform and participate in the linguistic questionnaire.⁴

The potential of crowdsourcing is remarkable. It is easy to see that a virtually large number of informants offers linguists unprecedented advantages: not only the possibility of obtaining a great number of geographically fine-grained data in a short time but also the opportunity to actively involve the speech communities and to immediately return elicited data in form of recordings, maps, images, tables or similar. Therefore, if on the one hand, the linguist can document micro-variation and ongoing changes in less described language varieties, on the other hand, this ‘give and take’ approach, besides promoting the so-called ‘citizen science’, helps the speakers to understand the linguistic richness of their territory (see Kruijt, Cordin & Rabanus 2023: 16-17).

Nevertheless, crowdsourcing also has some limitations. To name but a few: researchers cannot check back with the informants, audio data may be noisy and must be processed, only a few old speakers are familiar with online questionnaires, and speakers have time to plan the sentences that in natural settings are realized quickly and spontaneously (see Bidese, Cordin & Cosentino 2023). In addition, the distribution of questionnaires collected through crowdsourcing can be unsatisfactory: cities and urban areas are typically overrepresented, while less populated places and localities

¹ Retrievable at <https://alpilink.it/en/>.

² See <https://AlpiLinK.it/en/vinko/>.

³ See, for example, *Atlas zur deutschen Alltagssprache (Ada)* <https://www.atlas-alltagssprache.de/>, *Gschmöis* for the Swiss Dialects <https://www.gschmois.uzh.ch/de.html>, and *Microcontact (Language variation and change from the Italian heritage perspective)* <https://microcontact.sites.uu.nl/>.

⁴ See the registration form at <https://AlpiLinK.it/en/partecipa/>.

are absent or underrepresented. These critical aspects can be overcome through a self-administered sociolinguistic questionnaire that allows data analysis by location, gender and age groups. In the specific case of *AlpiLinK* the following speaker-related data are collected: age, gender, spoken variety, municipality of linguistic variety, period of residence in the town where the variety is spoken, contexts and frequency of use of the chosen language variety. There is also a parallel project, *VinKiamo*,⁵ developed by the *VinKo* and the *AlpiLinK* teams to support the informants who may not be able to use digital tools properly. *VinKiamo* activities are carried out in collaboration with upper secondary schools based on special agreements. They require students to recruit informants, especially from less-represented areas, and act as intermediaries between the *AlpiLinK* platform and the speakers themselves.

Finally, it is important to underline that *AlpiLinK*, like most other surveys based on crowdsourcing, aims at offering an in-depth comparison between remotely collected data and data traditionally collected through fieldwork.

In the present study, we use *AlpiLinK* data to demonstrate the potential of this tool in addressing areal contact. Specifically, we present two case studies concerning the syntax of clitics in (a) a Rhaeto-Romance variety spoken in Gardena Valley and (b) the Italo-Romance varieties spoken in North-Eastern Italy. In particular, we will analyze the dropping of the clitic of the 2nd person singular (henceforth 2SG) in Gardenese Ladin (Section 3) and the distribution of 3rd person singular (henceforth 3SG) clitic in constructions with weather verbs in the Italo-Romance varieties of Veneto, Trentino, and Friuli (Section 4), using a dataset of over 250 localities. In the conclusions (Section 5), we emphasize the importance of crowdsourcing data in demonstrating the effect of Germanic-Romance areal contact. Our framework is outlined in the following section (Section 2).

2. Areal Contact in North-Eastern Italy

The field of language contact is currently one of the most vibrant research areas in linguistics. Nevertheless, contact linguistics still lacks a definition of its specific research object and a widely accepted methodology (cf. Bidese 2023: 1-5). One of the main challenges concerns, on the one hand, the highly heterogeneous settings of contact – such as those found in Italy’s variegated linguistic situation (cf. Barco 2024) – and, on the other hand, the different levels at which the effects of language contact can be observed, for instance at the level of bilingual speakers or within a speech community where more than one language is in use. Muysken (2010) argued that there should be distinguished various ‘scenarios’ of language contact, in dependence of four aggregation levels: (i) the bilingual person; (ii) the bilingual community; (iii) the geographical region in which convergence effects between several languages can be observed; (iv) and, finally, larger areas of the world. Contact phenomena can operate at each level in different concrete ‘scenarios’ and in different temporal depths. The following table, taken from Muysken (2010: 268) shows the four levels of spatial aggregation related to the time span, the sources of investigation and the disciplines involved:

⁵ See <https://AlpiLinK.it/en/vinkiamo/>.

Table 1. Aggregation levels in language contact studies

	Space	Time	Source	Disciplines	Scenarios
Person	Bilingual individual	0-50 years	Recordings, texts, experiments	Psycholinguistics	Brain connectivity
Micro	Bilingual Community	20-200 years	Recordings, fieldwork observations	Sociolinguistics	Specific contact scenarios
Meso	Geographical Region	Generally 200-1,000 years	Comparative data, historical sources	Historical linguistics	Global contact scenarios
Macro	Larger areas of the world	Deep time	Typological data	Areal typology	Vague or no contact scenarios

Source: Muysken (2010: 268)

As already said, in this paper we will present two case studies concerning the loss of the 2SG enclitic pronoun in Gardenese Ladin and the distribution of the 3SG clitic with weather verbs in Trentino, Veneto and Friuli. We consider these two situations to belong to the third level of aggregation identified by Muysken (2010), where the contact between Germanic and Romance varieties is only areal, i.e. in the same geographical region, but, crucially, not in the community. Actually, the situation of Gardenese Ladin could also be placed on the second level of Muysken's typology, that of the bilingual community, where the contact languages are present in the same community. In fact, according to the sociolinguistic survey CLaM 2021 (see Cordin 2023; also, Fiorentini 2020), about 30% of the population of the Gardena Valley declare German as their mother tongue.⁶ Nevertheless, the general knowledge of Gardenese Ladin is assessed as very high (more than 90%) in almost all the municipalities of the Gardena Valley, with the only exception of Suredghes/Oltretorrente, which is the Ladin-speaking part of the German-speaking municipality of Ciastel/Castelrotto.⁷ For this reason, although the presence of German is very strong in the communities of the Gardena valley, we also include the case of Gardenese Ladin in the third level of aggregation, i.e. the level of geographical contact.⁸ The following map shows the area surveyed.

⁶ See the results of the question Q0301L1 "Which language do you consider to be your mother tongue?" (<https://cimbro-ladino-mocheno-2021.lett.unitn.it/en/gardenese-and-badiotto-ladin-competence#Q0301L1>).

⁷ See the results of the question Q0803L1C "Can you speak the Ladin of your valley?", in <https://cimbro-ladino-mocheno-2021.lett.unitn.it/en/gardenese-and-badiotto-ladin-competence#Q0803L1C>.

⁸ It is noteworthy that South-Bavarian/Tyrolean contact dialects of Gardenese Ladin exhibit the same phenomenon of dropping of the second person singular clitic. See the examples in section 4.3.

Figure 1. The area of investigation

Source: <https://alpilink.it/en/>

The first scholar to notice that even languages that are not genetically or typologically related can share part of the lexicon – interestingly, not the autochthonous one – and have a convergent syntax, was Wilhelm von Humboldt. His intuition was later formalized by Nikolai Trubetzkoy, who introduced the notion of sprachbund, i.e. ‘language union’ (see Trubetzkoy 1923 and, more recently, Aikhenvald 2007, McMahon 1994: 218 – 220 and Friedman and Joseph 2020, among many others). Bidese & Tomaselli (2021) revisited Schmidt’s (1872) famous *Wellentheorie* as a conceptual model for *Sprachbund* effects in areal contact.⁹ In this model, syntactic convergence phenomena between two typologically divergent systems are conceived as a consequence of areal superficial overlapping (*Sprachbund* synchronization) through the preservation of structural differences.

In the next chapters, we want to check this concept against two case studies of clitic syntax, namely, the dropping of the enclitic subject pronoun of 2nd person singular in Gardeneise Ladin (cf. Section 3), a phenomenon which is typical of the South-Bavarian dialects of the region, and the realization of the clitic pronoun of 3. person singular in the Italo-Romance dialects of Northeastern Italy (cf. Section 4), which is typologically much more typical of a pro-drop-language.

⁹ The *Wellentheorie* (wave model) was proposed by the German linguist Johannes Schmidt (1872) to explain the spread of new linguistic features among the languages of the Indo-European family. Unlike the *Stammbaumtheorie* (family tree model) (see Schleicher 1863), which describes linguistic innovation as the emergence of new groups of languages that share the same innovative features on the base of their phylogenetic relationship and models it like new branches on a tree, the *Wellentheorie* suggests that linguistic innovations spread outward from a central point like waves in water. It emphasizes overlapping influences among neighboring languages or dialects as well as among genetically non-related languages and recognizes the importance of transitional zones.

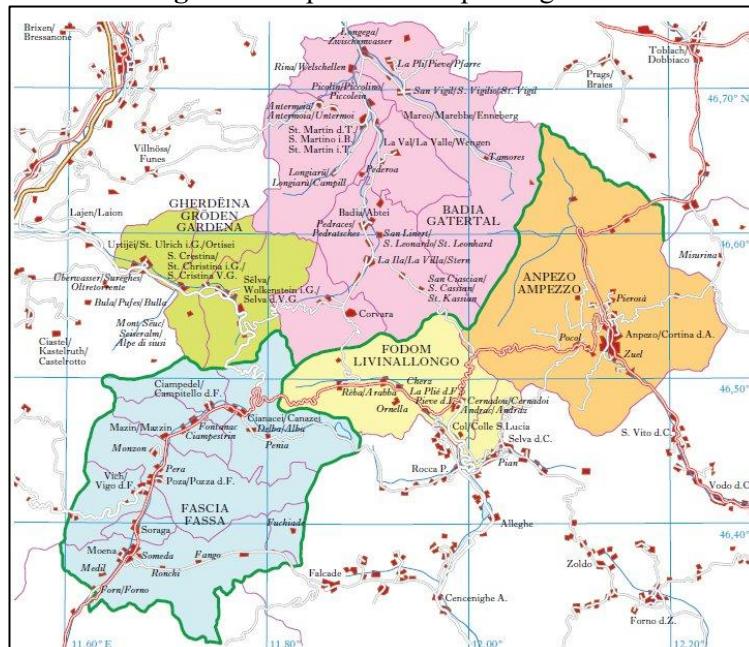
3. Clitic Syntax: The drop of the 2nd person singular clitic in Gardenese Ladin

3.1. The feature in diachrony

As many studies have confirmed (see Benincà 1994; Poletto 2002; Casalicchio & Cognola 2018; Casalicchio 2020), Gardenese and Badiotto Ladin share many syntactic features that typologically belong to verb-second [V2] languages. In addition to subject-verb inversion in main interrogatives – an aspect generally shared among northern Italo-romance varieties – Gardenese Ladin displays further restrictions on topicalized/left-dislocated elements in root declarative clauses, which also may trigger subject-inversion (see Casalicchio & Cognola 2018 for a comparison between Gardenese and Badiotto Ladin). This pattern aligns superficially with the Germanic ‘strict’-V2-varieties *modulo* the V-to-C-movement not being mandatory in all main clause contexts (hence defined as relaxed-V2).

Furthermore, if subject-verb inversion occurs, Gardenese differs from all other Ladin varieties by the complete absence of the 2SG enclitic subject pronoun (see Salvi 2020: 78). This feature has been traditionally attributed to the influence of Germanic varieties (German and Tyrolean dialects) [Salvi: *id.*], co-existing with Standard Italian in the linguistic repertoire of Ladin speech communities placed in South Tyrol, specifically in Gardena Valley and Badia Valley (see Belardi 1984; Berruto 2007; Fiorentini 2020).

Figure 2. Maps of Ladin-speaking areas



Source: Dell'Aquila 2010

It must be said, however, that an enclitic subject pronoun *-tɔ*, although in decline, was still present in the early 19th century as represented in example (1):

(1) Early 19th century Gardenese, Ploner, *L vedl mut*, from Casalicchio (2020: 174)

Sce	ne	ti	dès	böl	prëst	l-a	brëies,
if	not	her	give.2SG	well	soon	the-F	trousers

audirès=te tòst de biei svèies
hear.FUT.2SG=you soon some loud.PL screams
'If you don't give her the trousers immediately, you'll soon hear loud screams'

Casalicchio (2020) observed that in Ploner's texts, the drop of the enclitic is influenced by the presence of the interrogative particle *pa*:

(2) Early 19th century Gardenese, Ploner, *Kleine Erzählung*, from Casalicchio (2020: 174)
Co t'=es=pa poedù inamurè t' una
how REFL.2SG=have.2SG=Q could fall_in_love in a
tel persona?
such person
'How could you fall in love with such a person?'

In the second half of the century, the loss of *-tə* has by now become systematic (see Vian 1864).

Comparing the above presented historical data with those collected through *AlpiLinK* we will show in the next section how this feature has developed diachronically.

3.2 The feature in synchrony: the AlpiLinK data

Gardenese Ladin has the interrogative mark *pa*, which can also be found in many northern Italian dialects with different functions (see Munaro & Poletto 2002; Casalicchio & Cordin 2020). Derived from Latin POST ‘after’ (Gsell 1990), this particle underwent a process of grammaticalization, ultimately becoming obligatory in interrogative sentences in Gardenese Ladin (see Cordin & Dohi 2020). The sixteen Val Gardena speakers (aged between 13 and 87) who participated in *AlpiLink* confirm the mandatory use of this particle, as the next examples show:

(3) a. Present-day Gardenese (Ortisei, U0147)
 Foa *pa* Martina su-n plaza? L'=es=a
 was.3SG Q Martina on-the square her=have.2SG=Q
 ududa? *Čë* fajoa=la *pa*?
 seen what did.3SG=she Q
 'Was Martina in the square? Did you see her? What was she doing?'

As will be discussed in the next paragraph, syntactically, *pa* follows the verb, and phonetically, when it follows a sibilant, it undergoes an assimilation process that reduces it into the form '*a*' (Casalicchio 2020: 174), as illustrated in the following examples:

(3) b. Present-day Gardenese (Santa Cristina Valgardena, U0172)
Es=a udù coche Giuani à fa su si ciasa?
have.2SG=Q seen how John has made up his house
'Did you see how John built his house?'

c. Present-day Gardenese (Ortisei, U0142)

Es=a bel maià 1 pan o l=maies=a mo?
 have.2SG=Q already eaten the bread or it=eat.2SG=Q still
 'Have you already eaten the bread, or are you still going to eat it?'

Although there are currently few examples of this phenomenon, we can already report that, for instance, in translating the target sentence 'Did you see how John built his house?', as well as 'Martina was in the square, did you see her? What was she doing?', 100% of Gardenese speakers used *pa*.

Therefore, we must explain whether the absence of the enclitic in Gardenese is due to areal contact or represents an internal change (Section 3.3).

3.3. Discussion

Concerning the pro-drop phenomenon, Gardenese Ladin has been categorized as 'partial pro-drop' language (see Casalicchio 2021). The realization of the 3rd person subject is still mandatory (see Salvi 2016: 165; 2020: 93), whereas the absence of the 2nd person enclitic is assumed to be due to contact with the South-Bavarian variety spoken in the region for the 2nd singular (cf. footnote 8, above).¹⁰

As already stated in Casalicchio (2020), the interrogative particle *pa* correlates with the loss of the enclitic subject. This feature may be due to phonotactic reasons. If we consider that in Gardenese according to Casalicchio (2020: 173) "the atonic *e* gets weakened resulting in a change to *ə* or disappears" (It.: *la e atona subisce un indebolimento e passa a ə o cade*), a sequence made up of 2nd person singular verbal ending + 2nd person enclitic + *pa* could be rendered phonetically both as *-stəpa-* and as *-stpa-*. In the latter case, the alveolar consonant /t/ may drop producing a new sequence *-spa-* and this development can be easily found in the 19th-century texts (cf. 2). The present-day sequence *-sa-* arises from the fact that in Gardenese *pa* undergoes apheresis (> 'a') of the initial plosive after sibilants (cf. 4-7) (see Belardi 1984: 338; ALD II; Casalicchio 2020). The following schema reproduces the diachronic development of the string:

*-s-tə-pa > -s-t-pa > -s-pa > -s-a*¹¹

Let us now look at the specific properties of Gardenese *pa*. First, it is mandatory in all types of interrogative sentences (wh- and yes-no questions) as an 'interrogative marker' (cf. 4-7) (see Dohi 2017):

¹⁰ It should be said, however, that also the 2PL lacks a clitic pronominal form (Belardi 1984: 337), but in this case it is assumed to be an internal change that brings together Gardenese and Fassan Ladin (Salvi 2020: 78).

¹¹ However, it must be said that in 19th century Gardenese the 2nd person enclitic can drop even in the absence of the particle *pa*. The phonetic explanation, therefore, should be integrated with already existent approaches which take v-movement into account (see Benincà 1994).

(4) Ortisei, VinKo
 Ćë më des=a pon a mi?
 what me give.2SG=Q then to me
 ‘Thus what do you give me?’

(5) Gardenese Ladin, Bernardi (2002: 43)
 Jëis=a sëñ?
 go.2PL=Q now
 ‘Are you going now?’

(6) Gardenese Ladin, Belardi (1984: 338)
 Cie te dajon=s=a a ti?
 What you give=we=Q to you
 ‘What do we give you?’

(7) Gardenese Ladin, Personal communication
 Ćë më da=les=a a mi?
 what me give=they=Q to me?
 ‘What do they give me?’

Diachronically, *pa* in Gardenese grammaticalizes by changing from a modal particle with an illocutive / discourse functional value to a general question marker (see Hack 2014: 74). A similar process, although not generalized, occurs in Badiotto Ladin, where *pa* is obligatory only in wh-questions. In Fassan *pa* loses its modal value, but it is not reanalyzed as an obligatory interrogative marker in any type of root interrogatives (see Dohi 2017; 2017-18). Contrastively, *pa* does undergo phonetic reduction processes only in Gardenese:

(8) Badiotto, Poletto (2000: 58)
 Ula vas=t pa?
 where go=you Q
 ‘Where do you go?’

(9) Fassan, Dohi (2017-18: 128)
 Rùes=te pa doman?
 come=you Q tomorrow
 ‘Are you coming tomorrow?’

As we can see, in Badiotto and Fassan the 2nd person enclitic is lexically realized, even though *pa* is present. More precisely, the subject clitic is obligatorily lexicalized in Fassan, while it has become optional in Badiotto.

We should therefore conclude that, although in the history of Gardenese the particle *pa* is not the only key factor for the clitic drop, its degree of grammaticalization as an interrogative marker can determine whether prodrop occurs or not. It is not a coincidence that in Badiotto and Fassan, where the process of grammaticalization is not yet completed, the 2nd person singular subject clitic can still be found.

3.4. How areal contact works

The discussed case allows us to explain how areal contact works. In fact, the dynamics of contact with Germanic varieties might have acted as a simple reinforcing mechanism of a change already in progress, e.g. via a superficial overlap of morphosyntactic patterns. It is well-known, that Bavarian shares the same superficial string, namely ‘2SG=[*pro*]=Q’ (cf. 10 and, in particular, 11 and 12):¹²

(10) Bavarian, Bayer (1984: 211)

Kummst [pro_{ref}] noch Minga, dann muaßt
 come.2SG [pro_{ref.2SG}] to Munich then must.2SG
 [pro_{ref}] me b’suacha.
 [pro_{ref.2SG}] me visit
 ‘If you come to Munich, you must visit me.’

(11) Bavarian, Bayer (2012)

Wo wohnst=[pro_{ref}]=n (du)?
 where live.2SG=[pro_{ref.2SG}]=Q you
 ‘Where do you live?’

(12) South Bavarian (Merano/Meran)¹³

Wen hosch=[pro_{ref}]=n gsegn?
 who.ACC have.2SG=[pro_{ref.2SG}]=Q seen
 ‘Who did you see?’

However, the reasons for the emergence of the string ‘2SG=[*pro*]=Q’ in Bavarian are completely different. In fact, *pro* is the consequence of the reanalysis of the coalescence between the old inherited ending *-s* and the 2nd person singular pronoun *t(hu)* ‘you’, resulting in the new enlarged ending *-st* (see Braune 2004) whose phi-features are rich enough to assign the nominative case without overt realization of

¹² As well-known the correspondent sentences in Standard German do not admit the null subject pronoun:

(1) Kommst du nach München, dann musst du
 come.2SG you to Munich then must.2SG you
 mich besuchen.
 me visit
 ‘If you come to Munich, you must visit me.’

(2) Wo wohnst du denn?
 where live.2SG you PTC
 ‘Where do you live?’

(3) Wen hast du denn gesehen?
 who.ACC have.2SG you PTC seen
 ‘Who did you see?’

¹³ We are particularly grateful to our colleague Tamara Bassighini for the primary data on Meranese.

the pronoun (see Fuß 2004). The prototypical coalescence contexts between the ending and the pronoun were those in which the finite verb rises to C. From the verb the new ending also extended to the complementizers in C (cf. 13):¹⁴

(13) Bavarian, Bayer (2013: 30)

..., ob-st (du) des ned spuin kon-st
if-2SG you this not play can-2SG
'..., if you cannot play this.'

The very same structural development can be observed with the 2nd person plural: the old inherited ending *-t* and the 2nd person plural pronoun (*e*)s 'you' gave rise to the new ending *-ts* which exhibits the same syntactic pattern (cf. 14):

(14) Bavarian, Bayer (2013: 30)

..., ob-ts (es) des ned spuin ken-ts
if-2PL you this not play can-2PL
'..., if you cannot play this.'

Furthermore, the interrogative clitic *-n* is the result of the evolution of the modal particle *denn* into a question marker (see Bayer 2012). In many Bavarian varieties, it is almost mandatory in wh-questions (cf. 11 and 12 above); unlike the modal particle, the interrogative clitic *-n* has no connection to a semantic-pragmatic meaning (common ground reference) anymore.

To sum up: although the drop of the 2SG subject clitic in Gardenese Ladin looks like the same phenomenon in the regional variety of German, the triggers which led to the same superficial string seem to be reconducted to different (structural/internal) reasons, which are consistent with the divergent systems of the languages in contact.

¹⁴ As many researchers have pointed out (see among others Weiß 1998; Fuß 2004), this specific pattern is due to the so called Complementizer-Agreement (CA). Briefly, the [+pronominal] nature of the complementizer leads it to enter into an agreement relation with the Vfin, thus causing the inflection not only on the verb, but on the complementizer itself. This allows licensing an empty pronominal subject in specific syntactic contexts (not only in the 2SG/PL, but in some varieties also in the 1PL). See the following sentences:

(4) I woaß, dass-st du kumm-st
I know that-2SG you come-2SG
'I know that you are coming'

(5) I woaß, dass-ts ihr kumm-ts
I know that-2PL you.PL come-2PL
'I know that you (pl.) are coming'

In Gardenese Ladin, there is no such agreement between C and Vfin, being C generally [-pronominal] in Romance Languages. It makes sense though to postulate that the same (superficial) phenomenon, that is 2SG/PL-drop, is triggered by different (structural/internal) reasons, which are consistent with the very system of the languages in contact (by which we mean the C-dominant vs. Infl-dominant nature proposed in Hulk & van Kemenade 1995). In this context, the intense areal contact situation with Bavarian has probably reinforced the systematization of this phenomenon in Gardenese, but it is not its main cause.

In this context, the intense areal contact situation with Bavarian has probably reinforced the systematization of this phenomenon in Gardenese, but it is not its main cause.

4. Clitic Syntax: The realization of the 3SG expletive clitic across Italo-romance varieties

4.1. Introduction and *AlpiLinK* data

As far as clitic expletives are considered, a first theoretical correlation needs to be pointed out concerning the nature and typology of the expletives we are dealing with. Since the first seminal works on the Null Subject Parameter (hence NSP: see Rizzi 1982, 1994), one of the descriptive generalizations relates the negative setting of NSP to the presence of lexical expletives. That is indeed true for [-NSP] (or [-prodrop]) Languages such as English or Std. German:

(15) German

- a. Johann/Er ist zu-r Prüfung angekommen
John/he is to-the.DAT.F.SG exam arrived
'John/he arrived to the exam'
- b. *[pro_{ref}] ist zu-r Prüfung angekommen
[pro_{ref.3SG}] is to-the.DAT.F.SG exam arrived
- c. Es/*[pro_{EXPL}] regnet heute
3SG.EXPL/[pro_{EXPL}] rains today
'It rains today'

As the examples in (15) show, in non-argumental/impersonal constructions (for instance with weather verbs) a 3SG-form expletive needs to be overtly realized for the sentence to be well-formed.¹⁵

¹⁵ Contrastively, expletives are further classified according to their role within the syntactic structure. In this sense, we can distinguish between CP-expletives and TP-expletives: take these sentences in Std. German as example:

- a. Es kamen nur drei Studenten zu-r Prüfung
3SG.EXPL come.PAST.3PL only three students to-the exam
heute today
- b. Heute kamen Ø/*es nur drei Studenten zu-r
today come.PAST.3PL Ø/3SG.EXPL only three students to-the
Prüfung exam
'Only three students came to the exam today'
- c. Es regnet heute
3SG.EXPL rain.3SG today
- d. Heute regnet es/*[pro]
today rain.3SG 3SG.EXPL/*[pro]
'It rains today'

The opposite is confirmed in [+NSP]-languages, such as Italian or Spanish, where the possibility to license a referential [pro] correlates with the absence of a class of expletives:

(16) Italian

- a. Gianni/lui è venuto a-ll' esame
John/he is come to-the exam
'John arrived to the exam'
- b. [pro_{ref}] è venuto a-ll' esame
[pro_{ref.3SG}] is come to-the exam
'(He) arrived to the exam'
- c. [pro_{EXPL}]/*esso piove
[pro_{EXPL}]/3SG.EXPL rains
'(It) rains'

If, however, we observe the Italo-Romance varieties in northeastern Italy, the data show a more complex situation:

(17) Trentino, Casalicchio & Cordin (2020: 109)

L=/*[pro.3SG] à nevegà tuta la nòt
3SG=/[pro.3SG] has snowed all.F.SG the night
'It has been snowing all night long'

(18) Furlan (Gemona), AlpiLinK (U0175)

Vuêi al=/*[pro.3SG] à neveât
today 3SG=/[pro.3SG] has snowed
'Today it has snowed'

(19) Venetan (Verona), AlpiLinK (U0946)

L=/[pro.3SG] à nevegà¹⁶
3SG=/[pro.3SG] has snowed
'It has snowed'

(20) Venetan (Padua), AlpiLinK (U0918)

[pro.3SG] ga nevegà
[pro.3SG] has snowed
'It has snowed'

These two structures contrast in terms of sentence-internal expl-drop: as for (a), the expletive *es* is a CP-expletive, in that it is generated directly in SpecCP as its main function is to satisfy V2 constraint (i.e. to lexicalize SpecCP): once V2 is "checked" by another element (in b the adverb *heute*), the expletive is null. On the other hand, the expletive *es* in (c, d) is a TP-expletive, that is, it lexicalizes the structural subject position (via NOM-assignment): when clause-initial (c), it is generated in SpecTP, then moves to SpecCP to satisfy the V2-costraint; if SpecCP is already lexicalized (d), it stays in SpecTP and cannot be dropped, Std. German being a [-prodprop] language.

¹⁶ As for Veronese and neighboring varieties, both structures are well attested and equally spread in our Corpus, suggesting an individual level of variation.

(21) Julian (Trieste), AlpiLinK (U0701)

[pro.3SG] ga nevegado
 [pro.3SG] has snowed
 'It has snowed'

Differently from Std. Italian, such constructions in Italo-Romance varieties appear to pattern similarly to German or, in general, to [-NSP] languages. That is generally true for impersonal construction, in that a 3SG clitic is lexically realized mandatorily. Interestingly, the presence of a subject clitic is mandatory also with referential subjects (DPSSubject as well as pronominal ones),¹⁷ implying the so-called Clitic-Doubling (henceforth: CD) phenomenon (see also Poletto 1993, 2000 for a more detailed analysis):

(22) a. Trentino, Brandi & Cordin (1989: 113)

El Mario *(el=) parla
 the Mario 3SG= speak.3SG
 'Mario speaks'

b. Trentino, Brandi & Cordin (1989: 113)

Ti *(te=) parli
 2SG 2SG= speak.2SG
 'You speak'

4.2. Discussion

The mandatory realization of 3SG-clitics in NE Italo-Romance has been extensively debated, above all their nature within the syntactic structure as well as the correlation with the NSP. In terms of theoretical assumptions, we follow Brandi & Cordin's (1989) statement – also supported in Tomaselli & Bidese's works (2019, 2023) – that

¹⁷ As shown in Brandi & Cordin (1989) and further analysed in Poletto's works (1993, 2000), this pattern shows an essential distinction between Italo-romance varieties and other [-NSP] Romance languages, such as French:

- a. Il= parle
 3SG= speak.3SG
- b. Jean parle
 John speak.3SG
- c. *Jean il= parle
 John 3SG= speak.3SG
- d. Jean, il= parle
 John 3SG= speak.3SG

As the examples clearly show, clitic doubling works fine as a left-dislocation strategy for DPSSubjects (d), but non-dislocated DPSSubjects do not require CD, as for the sentence in (c) which is considered ill-formed in French and well-formed in Trentino varieties. This element supports the hypothesis that subject clitics in NE Italo-Romance varieties differ structurally from clitics in [-prodrop] languages.

such clitics are part of verbal morphology, thus not interfering with a positive setting of [NSP] for these varieties, which we consider [+prodrop] languages.¹⁸

As Tomaselli & Bidese (2023) pointed out, an interesting aspect concerns the distribution of 3SG-clitics with weather verbs, which is somehow influenced by the geographical location of the varieties considered. In particular, the phenomenon seems to ‘weaken’ as it moves from north to south, as shown by the examples drawn above for the Veronese area (where there is greater variation in the presence/absence of the phenomenon) to the Paduan/Venice area. A similar pattern is confirmed in Pescarini’s

¹⁸ Subject clitics differ structurally in NE Italo-Romance, as they do not project independently nor share the same status of German expletives or French clitics. In fact, in addition to CD-structures with DPSubjects, a fundamental difference concerns CD-structures with (bare) subj-quantifiers. See the following examples in French and Trentino:

- a. French

Personne	(*il=)	n'=a	rien	dit
no_one	3SG=	not=has	anything	said

- b. Trentino

Nisun	*(l=)ha	dit	niente	
no_one	3SG=has	said	anything	

‘No one said anything’

Contrastively, quantifiers cannot undergo topicalization. For this reason, they are to be considered in a subject position within the structure (i.e. SpecInflP in Romance); this explains why they cannot be doubled by a subject clitic in French, as the quantifier already lexicalizes the position. The opposite applies for Trentino, where clitics do not occupy any specific position but are considered part of verbal morphology, so they do not interfere with the quantifiers (*nisun*) being in SpecInflP. As interestingly suggested by one of the anonymous reviewers though, subj-pronouns in French have been hypothesized to be fully grammaticalized as fixed markers, also co-occurring with not-dislocated DPSubjects. Nevertheless, these facts do not invalidate the idea that French and Trentino clitics do not share the same status, given that in Trentino dialects (as well as in Friulian) they are mandatorily realized, which is not the case for French ones.

Further confirmation can be found in the same structures in Gardenese Ladin, which has lately developed V2-related phenomena typical of (standard) Germanic (data from ASIt).

- c. Selva di Val Gardena, ASIt

Zachei	(*l=)	cherderá	sú	1	prufessëur
someone	3SG=	call.FUT.3SG	PRT	the	professor

‘Someone will call the professor’

- d. Spere [che *(l=) vénie zachei]
hope.1SG [that 3SG= come.SUBJ.3SG someone]
‘I hope that someone will come’

In (c), the fronted quantifier cannot be doubled by the clitic. On the contrary, in free-inverted subject (post-VP) structures, CD is mandatory. This aspect may suggest that clitics in Gardenese Ladin do not share the same nature as other NE Italo-Romance varieties but rather are more structurally similar to the French (and perhaps German) ones. Further research is needed on this topic; however, it may be speculated that V2-related patterns in Gardenese Ladin (which can be traced back to contact with German) also influence somehow the status of subject clitics and, consequently, NSP-related phenomena.

analysis (2022: 6) on the distribution of the phenomenon in a wider set of northern Italo-romance varieties.

This correlation between the gradual weakening of the phenomenon and the geographical location of the varieties opens the way for possible analysis in terms of areal Germanic-Romance language contact. From this point of view, the presence of an Alpine *Sprachbund* has already been speculated (Gaeta & Seiler 2021) and, specifically, variation in NSP-related phenomena seems to provide further evidence in this regard, as pointed out in Bidese & Tomaselli (2021) and Madaro (2023). In these works, in particular, the key-point is to conceive NSP – to be intended as a cluster of microparameters (see Rizzi 1994, then the works in Biberauer *et al.* 2010 and in Cognola & Casalicchio 2018) – in its “granularity” as a mean to derive variational phenomena due to language contact. In this sense, patterns which contrast with a specific setting of the NSP in each language or language variety are to be intended as superficial manifestations (i.e. “Superficial Overlapping”, Bidese & Tomaselli 2021) which do not interfere with the structural setting of the (macro)parameter.

4.3. How areal contact works in Triveneto

In this direction, the large set of data available on the AlpiLinK-Corpus allows us to conduct a fine-grained analysis of the distribution of 3SG clitics across Triveneto. As we have already pointed out, crowdsourcing enables the acquisition of a lot of substantial data from different varieties. In the study presented here, we draw upon data from 155 speakers from Trentino, 76 from Veneto, and 25 from Friuli – thus giving us a broad picture of the linguistic situation in Triveneto.

In order to observe the distribution of 3SG clitics, we checked the target sentence “È nevicato in montagna ieri?” (‘Did it snow on the mountains yesterday?’). AlpiLinK’s results confirm the well-known in the literature. Indeed, in Trentino, «metereological verbs require a 3rd singular subject clitic», as Casalicchio and Cordin 2020: 109 states (see also Brandi and Cordin 1989). Data from AlpiLinK exhibits a similar trend: 3rd subject clitic is found in 57 varieties (83.55% of the surveyed speakers), while in 10 locations it is absent. Thus, most results align with the type exemplified in (23a), while only a few speakers responded with the type exemplified in (23b):

(23) a. Spormaggiore, AlpiLinK (U0716)

A=1	floccia	ieri	n	montagna?	[+clitic]
Has=3SG	snowed	yesterday	in	mountain	

‘Did it snow in the mountains yesterday?’

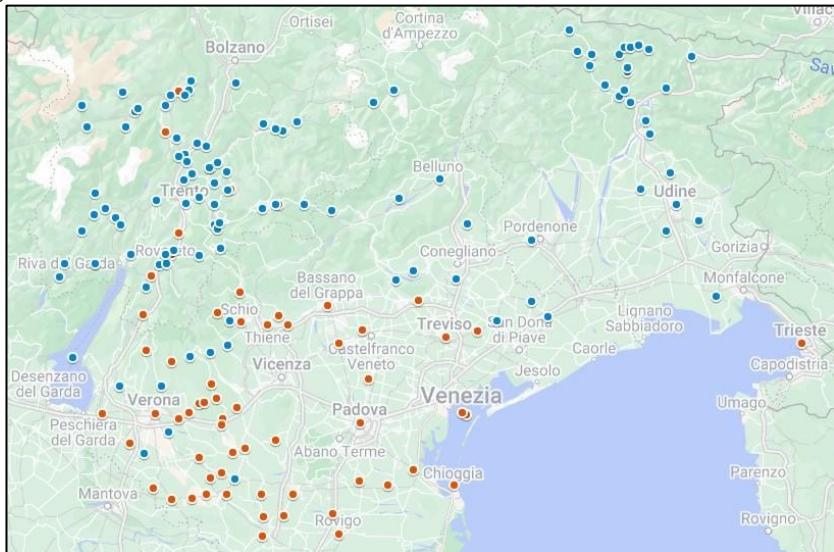
b. Avio, AlpiLinK (U0285)

A	nevegà	n	montagna	ieri?	[-clitic]
Has	snowed	in	mountain	yesterday	

‘Did it snow in the mountains yesterday?’

The following map presents a snapshot of the phenomenon. Based on the presence of the 3SG-clitic, we assigned value 1 to the variety (blue dots in the map) if the sentence contained the subject clitic, and value 0 to the variety if there was no clitic (red dots in the map):

Figure 3. 3SG clitics distribution with weather verbs across NE-Italo-romance



Our results align with Tomaselli and Bidese (2023) conclusion that the phenomenon appears to follow a geographical distribution in the Triveneto, with its occurrence diminishing progressively from north to south. We see that the two varieties reported in the examples in (23) are varieties from Trentino; however, Avio (example 22b), where 3SG subject clitic was not reported, is in the southern part of Trentino. Even in Veneto, according to AlpiLinK's current findings, the phenomenon is tenuous and spread mainly in the northern part of the region. Indeed, AlpiLinK's Veneto data reports the occurrence of the phenomenon in only 22 locations (all in the southwestern and northeastern areas), while it fails to appear in 53 locations. Although AlpiLinK gives us little data for the moment, in Friuli the phenomenon is present, eventually defining a north-eastern area in which meteorological verbs are paired with expletive clitics.

However, the situation is not so well defined. Looking at the map, there are some areas characterized by a high rate of variation in the distribution of the phenomenon, particularly in the contact area between Veneto and Trentino. If the southernmost Trentino varieties seem to show an opposite setting from the default (see Casalicchio & Cordin 2020), the same can be said for the more “outer” Veneto varieties, that is, in contact with Trentino varieties to the north and west (see Veronesese/Bellunese) and Friulian varieties to the east (see Trevisan varieties). Similarly, there is a notable internal variation: in Mezzocorona, for instance, 3 out of 7 speakers did not include the expletive in their production. The Furlan picture is clearer; there, despite the limited data, all 26 mapped points exhibit the phenomenon.

To sum up, mapping the data extracted from AlpiLinK can confirm the findings made by Tomaselli and Bidese (2023) on the VinKo data. In both corpora, «the maintenance of lexical expletives in the dialects spoken in Trentino is more robust in comparison with the Venetian region» within which there is «a clear decrease from the southwestern area (the province of Verona) toward the southeastern provinces of Vicenza, Padova, and Venice» (Tomaselli and Bidese 2023: 12).

5. Conclusions

In this paper we have presented data on two phenomena of clitic syntax in Gardenese Ladin and in Romance varieties of northeastern Italy, namely: (i) the drop of the 2SG clitic subject in correspondence with the interrogative particle *pa*, and (ii) the areal distribution of 3SG clitic subjects in impersonal constructions (with weather verbs). The approach used for the analysis of these two phenomena is opposite: on the one hand, we have analyzed the phenomenon in a very limited area and with respect to a single variety (Gardenese Ladin); on the other hand, we have studied the distribution in a much wider area encompassing about 250 different varieties.

This leads us to reflect on the potential of a tool like *AlpiLinK* and crowdsourcing in linguistic research, specifically when dealing with the dynamics behind language contact and variation at the level of areal linguistics. As already mentioned in the introductory section of this paper, this type of corpora makes it possible to collect data on single varieties and phenomena that are poorly documented in the literature. In particular, this helps to assess the extent to which a given phenomenon is widespread within a given community. A crucial example in this direction is represented by the 2SG drop in Gardenese Ladin together with the encliticization of the particle *-pa*. The combination of these two aspects makes the 2SG loss a probable internal development due to phonetic/phonological reasons, which, however, seems to be somewhat reinforced by contact with Southern Bavarian varieties of the region where the 2SG drop is also attested and theoretically explained due to phi-features on the C-head. This leads to an identical (superficial) representation of the phenomenon due to different causes, which are still consistent with the internal structure as well as core features of the variety.

Moreover, while considering some critical issues due to the physical absence of the interviewer, it is evident that this tool allows the coverage of entire areas, thus providing a much richer picture of variation phenomena in areal linguistics. A natural consequence of the latter aspect is that more punctual isoglosses can be drawn for specific morphosyntactic phenomena. As Figure 3 above shows, if we focus on the dynamics of areal language contact, we have to consider the more or less gradual expansion of certain phenomena (such as the emergence of clitics in impersonal constructions) as the propagation of microparameters: two opposite manifestations (one negative and one positive of the parameter) create a situation of ‘instability’, i.e. a higher rate of variation. That is the case for the area between Trentino and Veneto, where there is no clear-cut situation. For Friulan instead the pattern appears to be generally more solid in all its varieties (from the northernmost varieties of Carnico to Aquileia Friulan) and the setting switch only in the southernmost area where Giulian (Venetan) varieties of Triestino and Gradese are spoken, and subject clitics are not found with impersonal verbs. It should also be noted that once the central Veneto block is passed, the 3SG subject clitics reappear in the Emilian varieties, then disappear in the central and southern Italo-Romance varieties completely.

As far as contact-induced language variation is concerned, this leads us to postulate contact as the interplay between two (or more) epicenters of propagation of a given manifestation. The farther the varieties are from the epicenter, the greater the instability, thus creating areas of morphosyntactic ‘noise’. In our case, the 3SG clitic

drop (represented by the central Venetan area) contrasts with the mandatory realization 3SG clitics in Trentino, Upper Venetan, and Friulian varieties: the different degrees of instability on the border varieties (Trentino-Veneto vs. Friuli-Veneto) may be due to historical as well as geographical reasons.

Acknowledgments

We are grateful to the members of the scientific and organizational committee of the 17th Cambridge Italian Dialect Syntax-Morphology Meeting (CIDS 17) held at the University of Zurich for the opportunity given to us to present our research results.

Special thanks go to Jan Casalicchio, who provided us with his precious support regarding transcriptions and proofreading of the Gardenese Ladino sentences.

We are also grateful to the three anonymous reviewers for their thoughtful observations and comments, which surely helped us improve the quality of this paper. As required by the Italian academic system, we declare that this contribution was written in full collaboration and that Romano Madaro takes responsibility for Sections 4.1 and 4.2, and Section 5, Michele Cosentino for Sections 1, 3.1 and 3.3, Ermenegildo Bidese for Section 2 and 3.4, and Simone Barco for Sections 3.2 and 4.3.

References

ALD II = *Atlante linguistico del latino dolomitico e dei dialetti limitrofi*, II. <https://www.ald.gwi.uni-muenchen.de/it/?db=ald2>.

Aikhenvald, Alexandra Y. 2007. Grammars in contact: A cross-linguistic perspective. In Alexandra Y. Aikhenvald & Robert M. W. Dixon (eds), *Grammars in contact. A cross-linguistic typology*, 1-66. Oxford: Oxford University Press.

Barco, Simone. 2024. Dialect contact and koineization in Italy. In Barbara Hans-Bianchi, Chiara Truppi and Barbara Vogt (eds), *Speakers and Structures in Language Contact: Pluralistic Approaches to Change and Variation*, 225-254. Berlin, Boston: De Gruyter Mouton. <https://doi.org/10.1515/978311188348-009>

Bayer, Josef. 1984. COMP in Bavarian Syntax. *The Linguistic Review* 3: 209-274.

Bayer, Josef. 2012. From Modal Particle to Interrogative Marker: A Study of German *denn*. In L. Brugé, A. Cardinaletti, G. Giusti, N. Munaro, & C. Poletto (eds), *Functional Heads. The Cartography of Syntactic Structures*, vol. 7, 13-28 Oxford-New York: Oxford University Press.

Bayer, Josef. 2013. Klitisierung, Reanalyse und die Lizensierung von Nullformen: zwei Beispiele aus dem Bairischen. In W. Abraham & E. Leiss (eds), *Dialektologie in neuem Gewand. Zu Mikro-/Varietätenlinguistik, Sprachenvergleich und Universalgrammatik*, 29-45 Hamburg: Helmut Buske.

Belardi, Walter. 1984. Studi Gardenesi . In W. Belardi, P. Cipriano, P. Di Giovine, & M. Mancini (eds), *Studi latini e romanzi in memoria di Antonio Pagliaro*, 269-349. Roma: Dipartimento di Studi Glottoantropologici – Università «La Sapienza».

Benincà, Paola. 1994. L’interferenza sintattica: di un aspetto della sintassi ladina considerato di origine tedesca. In *La variazione sintattica*, 89-104. Bologna: Il Mulino.

Bernardi, Rut. 2002. *Curs de gherdëina: trëdesc lezions per mparé la rujeneda de Gherdëina. Tredici lezioni per imparare la lingua gardenese*. San Martin de Tor: Istitut Ladin «Micurà de Rü».

Berruto, Gaetano. 2007. Situazioni sociolinguistiche e tutela delle lingue minoritarie. Considerazioni alla luce della “Survey Ladins”. *Mondo Ladino* 31: 37-63.

Biberauer, Theresa, Holmberg, Anders, Roberts, Ian, & Michelle Sheehan. 2010. *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge University Press.

Bidese, Ermenegildo 2023. *Sprachkontakt generativ. Eine Untersuchung kontaktbedingten syntaktischen Wandels im Zimbrischen*. Berlin: De Gruyter.

Bidese, Ermenegildo & Alessandra Tomaselli. 2021. Language synchronization north and south of the Brenner Pass: modeling the continuum. *STUF – Language Typology Universals* 74: 185-216. <https://dx.doi.org/10.1515/stuf-2021-1028>

Bidese, Ermenegildo, Cordin, Patrizia, & Michele Cosentino. Limits and advantages of modern technologies for documenting the Italian Alps linguistic varieties: from *VinKo* to *AlpiLink*. Paper presented at the meeting *Documenting Languages, Documenting Cultures*, University of Naples, 5-6 October 2023.

Brandi, Luciana & Patrizia Cordin. 1989. Two Italian Dialects and the Null Subject Parameter. In O.A. Jaeggli & K.J. Safir (eds), *The Null Subject Parameter*, 111-142. Dordrecht: Springer.

Braune, Wilhelm. 2004. *Althochdeutsche Grammatik I: Laut- und Formenlehre*, revised by Ingo Reiffenstein. Tübingen: Max Niemeyer Verlag.

Casalicchio, Jan. 2020. Il ladino e i suoi idiom. In P. Videsott, R. Videsott, & J. Casalicchio (eds), *Manuale di linguistica ladina*, 144-201. Berlin-Boston: de Gruyter.

Casalicchio, Jan. 2021. Partial pro-drop in Gardenese Ladin: Syntax and acquisition. Paper presented at 15. *Cambridge Italian Dialects Syntax-Morphology Meeting*, University of Helsinki, 8-10 September 2021.

Casalicchio, Jan & Federica Cognola. 2018. V2 and (micro)-variation in two Rhaeto-Romance varieties of Northern Italy. In R. D’Alessandro & D. Pescarini (eds), *Advances in Italian Dialectology. Sketches of Italo-Romance grammars*, 72-105. Leiden-Boston: Brill.

Casalicchio, Jan & Patrizia Cordin. 2020. *Grammar of Central Trentino. A Romance Dialect from North-East Italy*. Leiden-Boston: Brill.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, Noam. 1982. *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1986. *Knowledge of Language. Its Nature, Origin, and Use*. New York: Praeger Publishers.

Cognola, Federica & Jan Casalicchio (eds), 2018. Null Subjects in Generative Grammar: A Synchronic and Diachronic Perspective. Oxford: Oxford University Press.

Cordin, Patrizia. 2023. La vitalità delle lingue di minoranza nel Trentino. Esiti della ricerca ‘CLaM (Cimbro Ladino Mòcheno) 2021’. *Studi trentini. Storia* 102/1: 211–228.

Cordin, Patrizia & Atsushi Dohi. 2020. Particelle modali. Un confronto tra dialetti trentini e ladino fassano. *Mondo ladino* XLIV: 43-68.

Dell’Aquila, Vittorio. 2010. Comunità Ladina. In *Enciclopedia dell’Italiano*. Roma: Istituto della Enciclopedia Italiana Treccani. Available at https://www.treccani.it/enciclopedia/comunita-ladina_%28Enciclopedia-dell%27Italiano%29/ (accessed on January 19, 2025).

Dohi Atsushi. 2017. La grammaticalizzazione della particella pa in alcune varietà ladine dolomitiche. *Ladinia* 41: 181-199.

Dohi, Atsushi. 2017-18. *La particella pa nelle varietà del ladino dolomitico con particolare attenzione al fassano*. Ph.D. thesis, Università degli Studi di Trento.

Fiorentini, Ilaria. 2020. Il pluringuismo dei ladini: aspetti sociolinguistici. In P. Videsott, R. Videsott, & J. Casalicchio (eds), *Manuale di linguistica ladina*, 480-502. Berlin-Boston: de Gruyter.

Friedman, Victor A. & Brian D. Joseph. 2020. Reassessing Sprachbunds: A view from the Balkans. In Raymond Hickey (ed.), *The Cambridge handbook of areal linguistics*. Cambridge: Cambridge University Press, 55–87.

Fuß, Eric. 2004. Diachronic clues to synchronic syntax: pro-drop and complementizer agreement in Bavarian. In E. Fuß & C. Trips (eds), *Diachronic Clues to Synchronic Grammar*, 59-100. Amsterdam-Philadelphia: John Benjamins.

van Gelderen, Elly. 2011. *The linguistic cycle: Language change and the language faculty*. Oxford: Oxford University Press.

Gsell, Otto. 1990. Beiträge und Materialien zur Etymologie des Dolomitenladinischen (M-P). *Ladinia* 14: 121-160.

Hack, Franziska M. 2014. The particle *po* in the varieties of dolomitic ladin – grammaticalisation from a temporal adverb into an interrogative marker. *Studia linguistica* 68, 1: 49-76.

Hulk, Aafke & Ans van Kemenade. 1995. Verb Second, Pro-Drop, Functional Projections and Language Change. In A. Battye & I.G. Roberts (eds), *Clause Structure and Language Change*, 227–256. New York: Oxford University Press.

Kruijt, Anne, Cordin, Patrizia, & Stefan Rabanus. 2023. On the validity of crowdsourced data. In E. Pustka, C. Quijada Van den Berghe, & V. Weiland (eds), *Corpus Dialectology*, 10–33. Amsterdam-Philadelphia: Benjamins.

Madaro, Romano. L'area di convergenza romano-germanica nelle Alpi e la posizione peculiare del Timavese. In S. Dal Negro & D. Mereu (eds), *Confini nelle lingue e tra le lingue. Atti del LV Congresso della Società di Linguistica Italiana*, 111-126. Milano: Officinaventuno.

McMahon, April M. S. 1994. *Understanding language change*. Cambridge: Cambridge University Press.

Munaro, Nicola & Cecilia Poletto. 2002. Distribuzione ed uso delle particelle frasali in alcune varietà venete. *Quaderni Patavini di Linguistica* 18: 107-126.

Muysken, Peter. 2010. Scenarios for Language Contact. In Raymond Hickey (ed.), *The Handbook of Language Contact*, 265–281. Chichester: Wiley-Blackwell.

Pescarini, Diego. 2022. Expletive Subject Clitics in Northern Italo-Romance. *Languages* 7: 265. <https://doi.org/10.3390/languages7040265>

Poletto, Cecilia. 1993. Subject clitic/verb inversion in north eastern Italian dialects. *Venice Working Papers in Linguistics* 3: 95–137.

Poletto, Cecilia. 2000. *The Higher Functional Field: Evidence from Northern Italian Dialects*. Oxford: Oxford University Press.

Poletto, Cecilia. 2002. The left periphery of a V2-Rhaetoromance dialect. A new perspective on V2 and V3. In S. Barbiers, L. Cornips, & S. van der Kleij (eds), *Microvariation. Proceedings of the Workshop on Syntactic Microvariation (Amsterdam, August 2000)*, 214-242. Amsterdam: Meertens Institute.

Rabanus, Stefan & Alessandra Tomaselli (2017), Räume, Grenzen und Übergänge: Subjektrealisierung im Sprachkontaktraum Deutsch-Italienisch. In: H. Christen, P. Gilles, & Ch. Purschke (eds), *Räume, Grenzen, Übergänge. Akten des 5. Kongresses*

der Internationalen Gesellschaft für Dialektologie des Deutschen (IGDD). Stuttgart: Steiner, 283-303.

Rabanus, Stefan, Kruijt, Anne, Tagliani, Marta, Tomaselli, Alessandra, Padovan, Andrea, Alber, Birgit, Cordin, Patrizia, Zamparelli, Roberto, & Barbara Maria Vogt. 2023. *VinKo (Varieties in Contact) Corpus v1.2*. Bolzano: ERCC. <https://clarin.eurac.edu/repository/xmlui/handle/20.500.12124/74>

Rabanus, Stefan, Kruijt, Anne, Alber, Birgit, Bidese, Ermenegildo, Gaeta, Livio, & Gianmario Raimondi. 2024. *AlpiLinK Corpus 1.1.3*. <https://zenodo.org/doi/10.5281/zenodo.8360169>

Renzi, Lorenzo & Laura Vanelli. 1983. I pronomi soggetto in alcune varietà romanzo. In AA.VV. *Scritti in onore di Giovan Battista Pellegrini*, 121-45. Pisa: Pacini.

Rizzi, Luigi. 1982. *Issues in Italian syntax*. Dordrecht: Foris.

Rizzi, Luigi. 1986. On the status of subject clitics in Romance. In O. Jaeggli & C. Silva-Corvalan (eds), *Studies in Romance linguistics*, 391-419. Dordrecht: Foris.

Rizzi, Luigi. 1994. Early Null Subjects and Root Null Subjects. In T. Hoekstra & B.D. Schwartz (eds), *Language Acquisition Studies in Generative Grammar, Language Acquisition and Language Disorders*, 151-176. John Benjamins Publishing Company.

Salvi, Giampaolo. 2016. Ladin. In A. Ledgeway & M. Maiden (eds), *The Oxford Guide to the Romance Languages*, 154-168. Oxford: Oxford University Press.

Salvi, Giampaolo. 2020. Il ladino e le sue caratteristiche. In P. Videsott, R. Videsott & J. Casalicchio (eds), *Manuale di linguistica ladina*, 67-108. Berlin-Boston: de Gruyter.

Schleicher, August. 1863. *Die Darwinsche Theorie und die Sprachwissenschaft*. Weimar: Böhlau.

Schmidt, Johannes. 1872. *Die Verwandtschaftsverhältnisse der indogermanischen Sprachen*. Weimar: Böhlau.

Tomaselli, Alessandra & Ermenegildo Bidese. 2019. Subject clitic languages in comparison. Subject clitics, finite verb movement, and nominative case assignment in Germanic (Bavarian, Cimbrian) and Romance (French, North Italian) varieties. In: Ermenegildo Bidese, Jan Casalicchio and Manuela Caterina Moroni (a cura di), *La linguistica vista dalle Alpi. Teoria, lessicografia e multilinguismo—Linguistic views from the Alps. Language Theory, Lexicography and Multilingualism*. Berlin: Peter Lang, pp. 43-61.

Tomaselli, Alessandra & Ermenegildo Bidese. 2023. Fortune and Decay of Lexical Expletives in Germanic and Romance along the Adige River. *Languages* 8: 44. <https://doi.org/10.3390/languages8010044>

Trubetzkoy, Nikolai S. 1923. Vavilonskaja ba š nja i sme š enie jazykov. *Evrazijskij vremennik* 3, 107 – 124.

Vian, Josef Anton. 1864. *Gröden, der Grödner und seine Sprache. Von einem Einheimischen*. Bozen: Wohlgemuth.

Weiβ, Helmut. 1998. *Syntax des Bairischen. Studien zur Grammatik einer natürlichen Sprache*. Tübingen: Max Niemeyer.