

Language variety as a linguistic subsystem: typological implications

ALESSANDRO VIETTI¹, MASSIMO CERRUTI²

¹FREE UNIVERSITY OF BOZEN-BOLZANO, ²UNIVERSITY OF TURIN

Submitted: 14/11/2024 Revised version: 9/9/2025

Accepted: 10/9/2025 Published: 25/02/2026



Articles are published under a Creative Commons Attribution 4.0 International License (The authors remain the copyright holders and grant third parties the right to use, reproduce, and share the article).

Abstract

This paper emphasizes the importance of language variety in typological linguistics and its role in understanding inter- and intralingual variation. Adopting a corpus-based perspective, the paper explores the covariation structure that links linguistic phenomena and social information, revealing distinct linguistic subsystems within a language. Two case studies of sociolinguistic variation in Italian illustrate how language varieties emerging from the data can carry micro-typological meaning. The first study empirically tests the co-existence of ‘neo-standard Italian’ and the traditional literary standard in speakers’ usage, while the second is a phonological analysis of the allophony of /r/ in Italian spoken by a bilingual community in South Tyrol (Italy). The analysis indicates that the integration of language varieties into typological research allows for a more complex representation of the hierarchical structure of linguistic phenomena.

Keywords: language variety; corpus-based typology; multivariate statistical analysis; Italian.

1. Introduction

Central to this paper is the notion of language variety, which is a key concept in variationist sociolinguistics (especially in European sociolinguistics; cf. Auer 1997; Guy & Hinskens 2016; Berruto 2004, 2019). A language variety is characterized by a set of co-occurring linguistic features, which are associated with language-external

factors, such as speakers' geographical provenance and social identity, or situational characteristics (cf. Milroy & Milroy 1997; Berruto 2009, 2012; Guy 2013).¹ From this perspective, therefore, a language is conceived as a diasystem (in the sense of Weinreich 1954); that is, a linguistic system coherently organized in regional, social and situational varieties. Such “heterogeneous system is then viewed as a set of subsystems which alternate according to one set of co-occurring rules” (Weinreich et al. 1968: 108).

Each language variety is indeed characterized by its own set of grammar rules. Every linguistic feature of a given variety is a “variant of a linguistic variable” (in Labov's 1966b classic definition) and results from the application of a variable rule, which is determined by the interplay of language-external and language-internal factors (cf. Sankoff 2004). A set of variable rules (with associated probabilistic weights) makes a language variety a linguistic subsystem² – that is, a grammatical system operating within the overall grammar of a language. This concept aligns with the notions of *Varietätengrammatik*, or ‘variety grammar’ (Klein 2004) and ‘linguistic coherence’ (Hinskens & Guy 2016; Beaman & Guy 2022).

In what follows, we will discuss how sociolinguistic investigations based on the notion of language variety can complement the analysis of cross-linguistic variation in typological research. First (Section 2), we will briefly outline two overall approaches to linguistic typology, paying special attention to the one that relies on quantitative corpus-based analysis and takes into account language-internal variation. Attention will then turn to two quantitative studies on spoken Italian³ (Sections 3 and 4), both revolving around the detection of co-occurrence relationships between linguistic features and social variables that can imply typological relevance. In this paper, we refer to the analysis of these two published studies to propose a new interpretation of the results that emphasizes their typological implications. Finally (Section 5), based on the results of these studies, we will argue that the analysis of language-internal variation in a typological perspective cannot overlook the actual partitioning of a linguistic system into language varieties.

¹ From a different angle, a language variety or a social dialect can only be identified on the basis of a speech community (“the fundamental concept for identifying varieties is the language community”; Guy et al. 2022: 53), as defined by Labov (1966a; but see also Hymes 1967; Gumperz 1968).

² “Most of contemporary linguistics conceptualizes language varieties in terms of underlying linguistic systems – a grammar that is shared by the speakers of the variety” (Guy et al. 2022: 54).

³ ITA; Indo-European, Romance.

2. Linguistic typology and variationist sociolinguistics

As far as typological research is concerned, the traditional approach can be contrasted with a more recent one. The former can be referred to as “grammar-based” (cf. Schnell & Schiborr 2022), or even “type-based” (cf. Levshina 2019). This line of research mainly draws on data coming from reference grammars and other descriptive accounts, such as linguistic features collated in large typological databases (e.g., *WALS*, Dryer & Haspelmath 2013). Accordingly, studies in this perspective generally rely on qualitative methods and are mostly based on linguistic features with allegedly categorical presence (even though typological databases show dominant types rather than categorical features; see, e.g., Dryer & Haspelmath 2013). Language-internal variation is, therefore, generally disregarded (Wälchli 2009).

Conversely, a more recent approach to linguistic typology can be referred to as “corpus-based” (cf. Levshina 2022; Schnell & Schiborr 2022), or even “token-based” (cf. Levshina 2019; see also Haspelmath 2018). This strand of research mainly draws on data coming from parallel and comparable linguistic corpora, or better yet from “language use, as approximated by corpora” (Levshina 2019: 534). Studies in this perspective are conducted with quantitative methods and employ continuous measures, such as the frequency or probability of occurrence (see, e.g., Gerdes et al. 2021) and probability-derived measures (e.g. Shannon entropy; cf. Levshina 2019), to investigate typologically relevant gradient properties of languages (e.g. word order variability). Linguistic features with variable presence clearly reflect language-internal variation; it is no coincidence that an increasing number of corpus-based typological studies give an account of both cross-linguistic and intra-linguistic variability (cf. Schnell & Schiborr 2022).

In sociolinguistic research, the concept of language variety has always been rather controversial. From the very beginning, in fact, it has been considered by some as a mere theoretical construct, unable to represent the actual behavior of speakers⁴ (see, e.g., Hudson 1980). However, recent research has convincingly shown that the notion of language variety is firmly rooted in empirical evidence. Quantitative corpus-based studies conducted on actual speech data have indeed allowed for the detection of statistically robust sets of co-occurring features, each associated with socially-defined

⁴ This issue has then come to the forefront with the advent of ‘superdiverse’ and globalized societies, which has challenged the explanatory power of most of the traditional notions in sociolinguistics (see, e.g., Blommaert 2010).

groups or communicative situations (see, e.g., Ghyselen & De Vogelaer 2018; Vietti 2019; Villena-Ponsoda & Vida Castro 2020; Beaman & Sering 2022).

Studies like these have affinities with corpus-based typological studies, since they both draw on data coming from corpora⁵ and rely on quantitative methods. Nevertheless, quantitative measures employed by corpus-based typology tend to reduce language-internal variation to the presence of linguistic variants in different proportions. These measures can determine how much variation is found in a given dataset but disregard the arrangement of variants in a diasystem. Significantly, in typological studies that use corpus-based indices, each language (see, e.g., Levshina 2019, 2022; Schnell & Schiborr 2022: 176–177) or language variety (cf. Szmrecsanyi 2009) is assigned the value of an index of variability and regarded as a single data point. Languages or language varieties are then compared in terms of their respective index values.

From a sociolinguistic perspective, however, it is worth considering that every linguistic variant that comes into play in calculating the aforementioned indices is actually one of the co-occurring features that constitute a language variety; in other words, it is one of the structural units that form a grammatical subsystem (cf. Section 1). The simple use of indices of variability bears the risk of overlooking such an inner structure. A sociolinguistic approach can, therefore, complement corpus-based typology by providing interpretative tools for the analysis of the structure of language-internal variation. The following case studies provide us with an opportunity to elaborate on this issue.

3. Standard varieties in speakers' usage

The first study at hand (Cerruti & Vietti 2022) falls within the line of research dealing with the emergence of a new standard variety of Italian, the so-called 'neo-standard Italian'⁶ (cf. Berruto 1987, 2012; Cerruti et al. 2017; Ballarè 2020), and aims to

⁵ However, "it must be remarked that the use of corpora for typology is practically and methodologically more challenging than for sociolinguistics. The main issue is that cross-linguistic corpus-based studies need a reasonable number of corpora of different languages, which additionally must be sufficiently similar to allow a meaningful comparison" (Ballarè & Inglese 2023: 10; see also Levshina 2022).

⁶ In fact, the convergence of the standard usage towards sub-standard varieties has led to the emergence of a new standard. It consists of (i) a nationwide shared core of originally sub-standard features that have come to be used and accepted even in formal and educated speech, as well as in writing, and (ii)

empirically test the co-existence of this newly emerged standard and the traditional, literary standard variety of Italian in speakers' usage.

3.1. Communicative situations and communicative purposes

This study, which considers the speech of university students as a case in point, is based on a subset of data coming from KIP (Mauri et al. 2019); i.e., a speech corpus consisting of different types of communicative interactions recorded at the universities of Turin and Bologna (two major cities in Northern Italy). In particular, a comparison is made between three types of interactions collected in Bologna: (a) spontaneous conversations among students recorded by in-group members; (b) semi-structured interviews collected by students within their peer-group; and (c) student-professor interactions during oral examinations.

This investigation takes into account 36 morphological, syntactical and lexical features; 23 of which are associated in the research literature with either neo-standard Italian or literary standard Italian, while the other 13 are related to the situational characteristics and/or the social attributes of speakers. As for the whole set of investigated features, the reader is referred to Cerruti & Vietti (2022: 265–267). Here we will focus on the 23 features associated with either neo-standard Italian or literary standard Italian (see Table 1).

	Label	Feature	Standardness
1.	Cleft	cleft sentence	neo-standard
2.	R_disloc	right dislocation	neo-standard
3.	L_disloc	left dislocation	neo-standard
4.	RET/GAP	pronoun retention/gap strategy (IO, OBL, GEN)	neo-standard
5.	MFNC.che	multifunctional <i>che</i>	neo-standard
6.	Dem Loc	demonstrative + locative adverb	neo-standard
7.	Dem st*	aphaeretic forms of proximal demonstratives	neo-standard
8.	n_Dem qu*	distance-neutral demonstrative <i>quello(/questo)</i>	neo-standard
9.	Cl gli	third-person singular indirect object clitic <i>gli</i> used for female referents	neo-standard

and a number of regionally marked features that have become part of the standard usage in distinct geographic areas. This is actually part of a more general process, which in recent years has led to the emergence of new standard varieties in many European languages (see, e.g., Kristiansen & Coupland 2011).

	Label	Feature	Standardness
10.	Refl 3	third-person singular and third person plural pronouns (<i>lui/lei</i> and <i>loro</i>) used as reflexive pronouns	neo-standard
11.	Loc <i>ci</i>	existential construction with <i>ci</i> as pre-copular proform	neo-standard
12.	<i>quell* che</i>	routinized formula <i>quello/i/a/e che</i>	neo-standard
13.	<i>tipo</i>	non-nominal uses of <i>tipo</i>	neo-standard
14.	Phrasal.V	phrasal verbs	neo-standard
15.	V <i>ci</i>	idiomatic verb- <i>ci</i> constructions	neo-standard
16.	V <i>ne</i>	idiomatic verb- <i>ne</i> constructions	neo-standard
17.	Pass	passive construction	literary standard
18.	REL	relative pronoun (IO, OBL, GEN)	literary standard
19.	Dem <i>qu*</i>	<i>questo</i> -type proximal demonstratives, <i>quello</i> -type distal demonstratives	literary standard
20.	n_Dem <i>ciò</i>	distance-neutral demonstrative <i>ciò</i>	literary standard
21.	Cl <i>le</i>	third-person singular indirect object clitic <i>le</i> used for female referents	literary standard
22.	Refl <i>sé</i>	third person singular and third person plural reflexive pronoun <i>sé</i>	literary standard
23.	Loc <i>vi</i>	existential construction with <i>vi</i> as pre-copular proform	literary standard

Table 1: Neo-standard and literary standard features.

The main research question that the study seeks to answer is whether neo-standard Italian and literary standard Italian actually consist of two distinct sets of co-occurring linguistic features in the usage of speakers. Groups of features are detected by means of a Principal Component Analysis (PCA).⁷ PCA results in a geometric representation of elements on a plane. In Figure 1, each point on the plane stands for a linguistic feature under scrutiny (features are labelled as in Table 1). The distance between the points represents the degree of correlation; that is, the nearer two points are on the plane, the stronger the co-occurrence between the two corresponding features in speakers' usage. Two bundles of features can thus be identified: a particularly cohesive group of features on the left, and a set of more broadly distributed features on the right.

⁷ The use of PCA in the study of language varieties is well grounded in sociolinguistic research. The seminal application of this method can be found in Horvath & Sankoff (1987). Other highly significant works are cited in Walker, Hoffman & Meyerhoff (2022: 71–72) and Beaman & Sering (2022: 88–89).

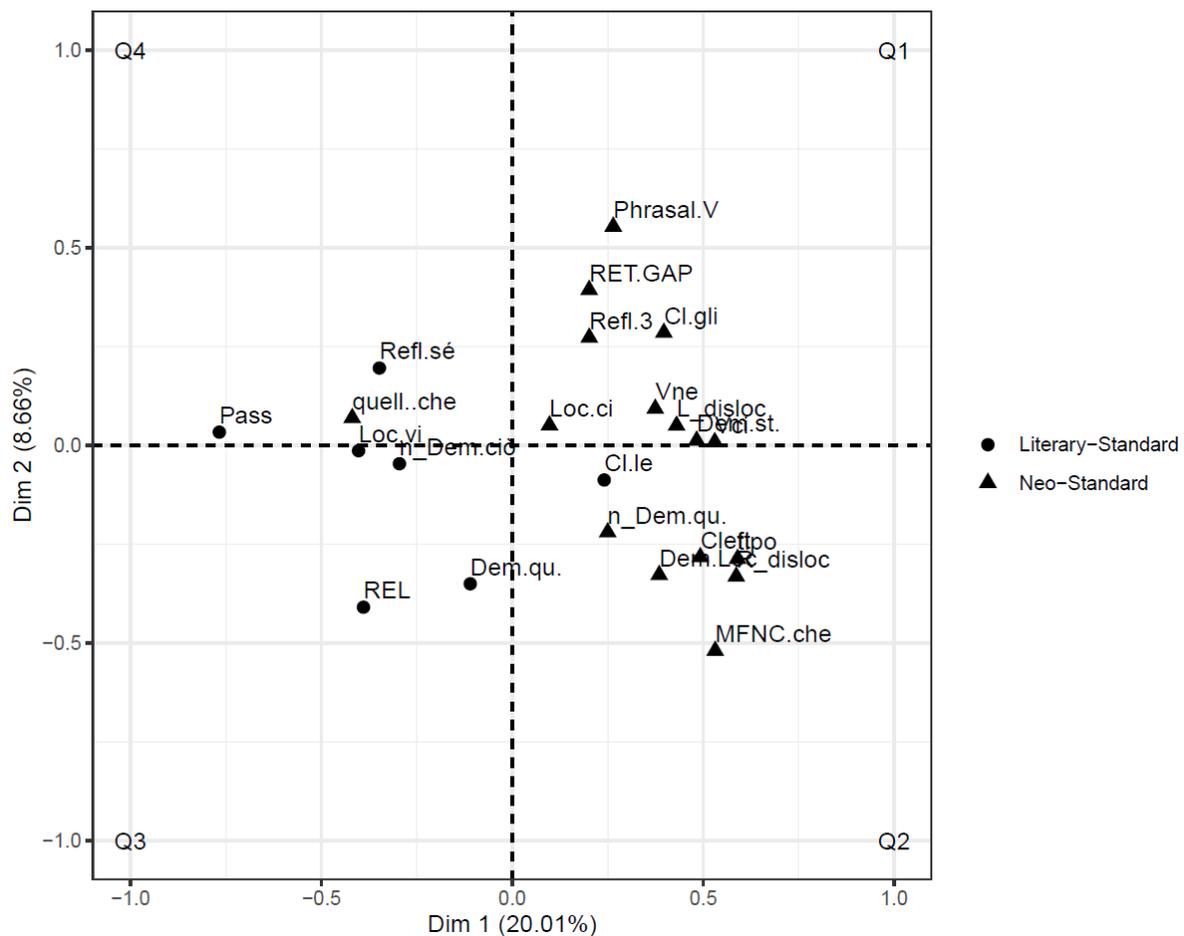


Figure 1: Plot of neo-standard and literary standard features in students' speech (Cerruti & Vietti 2022).

In Figure 2, each point represents a specific communicative interaction; the nearer two points are on the plane, the more features the corresponding communicative interactions have in common. In this respect, oral examinations stand apart from communications between peers (be they conversations or interviews). Figure 1 and Figure 2 can overlap. In fact, the linguistic features on the left half of Figure 1 tend to co-occur during oral examinations (left half of Figure 2), while the features on the right half of Figure 1 tend to co-occur during peer-to-peer communications (right half of Figure 2).

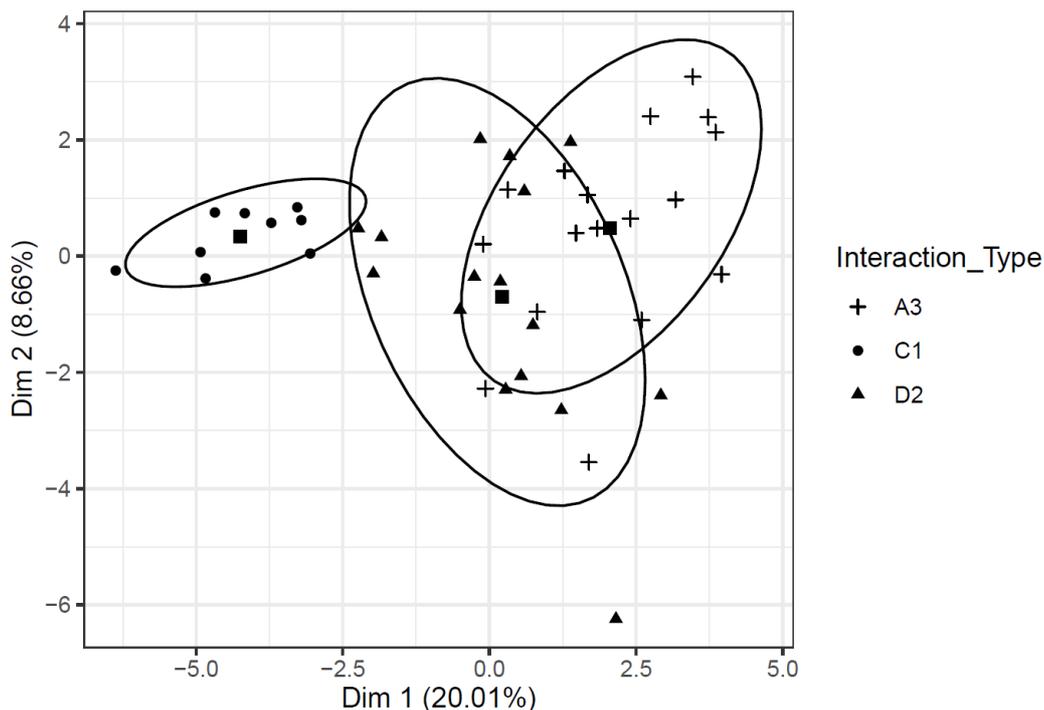


Figure 2: Plot of communicative interactions: spontaneous conversations (A3), semi-structured interviews (D2), oral examinations (C1) (Cerruti & Vietti 2022).

Moreover, in both Figure 1 and Figure 2 correlation patterns reflect the interplay of two underlying factors, represented by Dimension 1 (the horizontal axis) and Dimension 2 (the vertical axis). Dimension 1 can be associated with the opposition between formal and informal situations: the main difference between oral examinations (left half of Figure 2) and peer-to-peer communicative interactions (right half of Figure 2) relate to the asymmetrical vs. symmetrical relationship between the speakers; Dimension 2 can be associated with the opposition between interactional and informational purposes; the main differences between spontaneous conversations (upper end of Figure 2) and semi-structured interviews (lower end of Figure 2) indeed relate to the speakers’ communicative intentions: participants in spontaneous conversations are engaged in expressing feelings, opinions, and attitudes, as well as indexing social identities, while participants in interviews are more focused on conveying information (cf. Biber & Conrad 2009).

In conclusion, PCA allows for the detection of two distinct clusters of co-occurring features (see Figure 1): one consisting of neo-standard features (with the exception of feature 21, Table 1) and the other consisting of literary standard features (with the exception of feature 12, Table 1). Neo-standard features tend to cluster during peer-to-peer communicative interactions, while literary standard features do so during oral

examinations. Two distinct language varieties can thus be identified in students' speech according to the symmetrical vs. asymmetrical relationship between the participants and, therefore, in accordance with the opposition between informality and formality.⁸ At the same time, the co-occurrence of neo-standard features both in spontaneous conversations and in semi-structured interviews, namely in settings characterized by different communicative purposes, can reflect the multifunctional character of the newly emerged standard. Literary standard features, instead, seem to be more closely aligned with informational purposes.

3.2. Typological considerations

We now look at some new insights with these results. First of all, it is worth noting that most differences in grammar between literary standard Italian and neo-standard Italian are typologically relevant. In most cases, in fact, a given literary standard variant and its neo-standard counterpart represent two distinct values of “a structural property of language that describes one aspect of cross-linguistic diversity” (according to the definition of *feature* given in WALS, cf. Dryer & Haspelmath 2013); namely, two different values of a structural property of language that contribute in distinguishing between languages of different types (see also Comrie et al. 2013).

As regards, for example, relativization strategies, and in particular relativization on obliques, the literary standard use of a pronominal element signaling the syntactic role of the relativized item, as in utterance (1), can be contrasted with the neo-standard use of a clause-initial invariable complementizer, as in (2) (see also Cerruti 2017). This distinction clearly mirrors the opposition between two relativization strategies that can be found cross-linguistically (cf. Comrie & Kuteva 2013); that is, relative pronoun strategy and so-called ‘gap’ strategy (the latter involving cases in which the relative construction does not express the syntactic role of the relativized item).⁹

⁸ The co-occurrence of literary standard features during oral examinations appears to reflect the role that the old standard continues to play in school education (cf. Berruto 2017). The presence of neo-standard features in other formal contexts, such as spoken and, to a certain extent, written media, is, however, beyond dispute, as is widely documented by an extensive body of research (see Ballarè 2020 for an up-to-date review).

⁹ Another relativization strategy identified by Comrie & Kuteva (2013) is found to occur in our data, namely, pronoun retention; e.g., *il ragazzo che gli ho scritto una lettera* ‘the boy to whom I wrote a letter’, lit. ‘the boy that to him I wrote a letter’. Positions lower than direct object on Keenan & Comrie’s

(1) literary standard

la ragazza con cui lavoravo di solito
the girl with REL work.PST.IPFV.1SG usually
'the girl I used to work with'
(KIP Corpus, BOD1011)

(2) neo-standard

gente che non faccio conoscenza
people COMP NEG make.PRS.1SG acquaintance
'people I'm not acquainted with'
(KIP Corpus, BOA3004)

Further examples can be found in the domain of adnominal demonstratives. As for the traditional, literary standard, a norm 'without adjectives' has been replacing a Tuscan-based norm (cf. Da Milano 2015). The former reflects a distance-oriented system expressing a two-way contrast, in which a proximal demonstrative (*questo*) and a distal one (*quello*) indicate the relative distance between the referent and the speaker; see, e.g., excerpt (3). The Tuscan-based norm reflects a person-oriented system expressing a three-way contrast, in which a third demonstrative (*codesto*) denotes a referent near the hearer; see, e.g., example (4).¹⁰ The binary opposition between a proximal demonstrative (*questo*, sometimes in the apheretic form *sto*, as in excerpt (5)) and a distal one (*quello*) is also drawn on in neo-standard Italian. However, neo-standard is distinguished by the optional presence of a demonstrative preceding the noun and a locative adverb (*qui/qua*, *lì*, *là*) following the noun, as in example (6) (cf. Berruto 2012: 87–88). This is a distance-oriented system that can express a three-way contrast, in which *lì* indicates a referent nearer than *là* (cf. Benedetti & Ricca 2002). The order of noun and demonstrative and the distance contrasts of demonstratives are known to be sensitive to cross-linguistic diversity (cf. Dryer 2013a; Diessel 2013).

(1977) Accessibility Hierarchy are relativized by pronoun retention in 7.08% of cases (16/226), by gap in 9.73% (22/226) and by relative pronouns in 83.19% (188/226). Subject and object relativization exclusively employ the gap strategy. The indirect object is relativized by relative pronouns in 92% of cases (23/25) and by pronoun retention in 8% (2/25).

¹⁰ In our data, *codesto* is only found among speakers coming from Tuscany, as is the case with (4).

(3) literary standard

questo lavoro [...] quel testo

this work that text

'this work, that text'

(KIP Corpus, BOC1002, BOC1004)

(4) literary standard

codesto piatto

that dish

'that dish (near the hearer)'

(KIP Corpus, BOA3003)

(5) neo-standard

'sto rumore

this noise

'this noise'

(KIP Corpus, BOA3021)

(6) neo-standard

questo ragazzo qui [...] quel libro là [...]

this boy here that book there

quelle cose là

those things there

'this boy, that book, those things (over there)'

(KIP Corpus, BOA3013, BOA3011, BOA3012)

Another source of diversity in the world's languages is the grammatical expression of co-reference between the subject and a non-subject argument within the same clause (cf. Haspelmath 2023). In the third person, literary standard Italian makes a distinction between a series of independent non-reflexive personal pronouns (see, e.g., *lei* in utterance (7)) and the independent reflexive pronoun *sé* (see, e.g., (8)), the latter being used as both a singular and plural form that signals co-reference with the subject of the clause. Conversely, neo-standard Italian can also use the series of non-reflexive third-person pronouns to express co-reference with the subject (cf. Cordin 2001: 610–611), as in utterance (9).

(7) literary standard, neo-standard

lui tirava la palla a lei
he throw.PST.IPFV.3SG the ball to her
'he threw the ball to her'
(KIP Corpus, BOA3002)

(8) literary standard

i beni che la sposa porta con sé
the goods COMP the bride take.PRS.3SG with herself
'the goods that the bride brings with herself'
(KIP Corpus, BOD1011)

(9) neo-standard

non lavorava per gli altri,
NEG work.PST.IPFV.3SG for the others
lavorava per lei
work.PST.IPFV.3SG for her
'she didn't work for others, she worked for herself'
(KIP Corpus, BOA3013)

Moreover, literary standard Italian shows gender marking on independent non-reflexive third-person pronouns, in both singular and plural forms. As for the latter, *essi* is used for male referents and *esse* for female referents; see, e.g., excerpts (10) and (11). On the contrary, neo-standard Italian uses a single third-person plural form (*loro*), as in (12), and, therefore, has no overt gender distinction in independent third-person plural pronouns (cf. Berruto 2012: 83–84).¹¹ On a cross-linguistic level, different values of gender contrasts in independent personal pronouns contribute in distinguishing between languages of different types (cf. Siewierska 2013).¹²

¹¹ Gender marking on third-person plural pronouns, or lack thereof, is not among the linguistic features considered in Cerruti & Vietti (2022). However, in the subset of KIP data collected in Bologna, the use of *essi/esse* is found to occur only during oral examinations (cf. www.kiparla.it).

¹² In the world-wide context, gender contrast on third-person pronouns, in both singular and plural forms, is rarer than only on third-person plural pronouns (cf. Siewierska 2013). Similarly, relative pronoun strategy is rarer than gap strategy in both oblique and subject relativization (cf. Comrie & Kuteva 2013). Cases like these lead Grandi (2019) to hypothesize that literary standard Italian is more 'exotic' (in the sense of Dahl 1990) than neo-standard Italian.

(10) literary standard

la casa che essi fecero per noi
the house COMP they.M make.PST.PFV.3PL for us

‘the house they built for us’

(KIP Corpus, BOD1006)

(11) literary standard

qualsiasi esse siano
whatever they.F be.PRS.SBJV.3PL

‘whatever they may be’

(KIP Corpus, BOD1009)

(12) neo-standard

loro non capirebbero una parola
they NEG understand.PRS.COND.3PL a word

‘they wouldn’t understand a word’

(KIP Corpus, BOC1004)

Differences in grammar between literary standard Italian and neo-standard Italian can also be found in the coding of spatial relations in motion events. In fact, literary standard tends to express multiple semantic components of a motion event on a single verb. For example, the verb can encode (a) motion and path (see, e.g., *scendo* in (13)), with the possibility for manner to be expressed by phrases or clauses appearing in an adverbial function (e.g., the prepositional phrase *di corsa* ‘in a rush’, cf. Bernini 2010: 30–31, 33–35), as in verb-framed languages (cf., *inter alia*, Talmy 1985, 2000), or even (b) motion, path and manner (see, e.g., *ho infilato* in (14)), with the contribution of prefixes coding path (cf. Iacobini & Vergaro 2012), as in satellite-framed languages (cf. Beavers et al. 2010). On the contrary, neo-standard Italian is more prone to the use of phrasal verb constructions (cf. Berruto 2017: 42), in which some semantic components are coded on the verb and others on an accompanying particle, as in satellite-framed languages. This is, for instance, the case with phrasal verbs coding motion on the root and path on an adverb (see, e.g., *vado giù* in utterance (15)), or both motion and manner on the main verb and path on an adverb (see, e.g., *si lancia giù* in excerpt (16)).

(13) literary standard

scendo a prendere le brioche
descend.PRS.1SG to take the brioches
'I'll go down and buy the brioches'
(KIP Corpus, BOD2004)

(14) literary standard

ho infilato il telefono in tasca
have.PRS.1SG slip.PTCP the phone in pocket
'I slipped the phone in my pocket'
(KIP Corpus, BOA3021)

(15) neo-standard

ho detto vado giù
have.PRS.1SG say.PTCP go.PRS.1SG down
'I said I'll go down'
(KIP Corpus, BOA3003)

(16) neo-standard

così il gatto non si lancia giù
so the cat NEG self throw.PRS.1SG down
'so the cat can't jump down'
(KIP Corpus, BOA3013)

3.3 Language types and language varieties

The handful of structural properties exemplified above suffices to show that each set of standard features, i.e. that of literary standard Italian and that of neo-standard Italian, is related not only to extra-linguistic motivations, such as the degree of formality of the situation or communicative purposes (see Section 3.1), but also to linguistic factors. Two of these are the overt coding of grammatical oppositions and the analytic vs. synthetic marking of meaning. As for the former, literary standard appears to express more contrasting grammatical meanings than neo-standard. It is indicative that literary standard uses relative pronouns to provide overt indication about the syntactic role of a relativized item (feature 18, Table 1; cf. utterance (1)),

draws on a reflexive pronoun to signal co-reference with the subject of the clause (feature 22, Table 1; cf. (7), (8)), and shows gender distinction in independent third-person plural pronouns (cf. (10), (11)). On the contrary, neo-standard can encode all syntactic roles of a relativized item by an invariable complementizer (feature 4, Table 1; cf. (2)), resorts to non-reflexive personal pronouns also to convey a reflexive meaning (feature 10, Table 1; cf. (9)), and has no gender contrast in third-person plural pronouns (cf. (12)).

Moreover, literary standard appears to be more inclined to synthetic marking of meanings than neo-standard. For example, literary standard uses a single element, i.e. a relative pronoun, to mark subordination and overtly signal the syntactic role of a relativized item (feature 18, Table 1; cf. utterance (1))¹³ and tends to encode multiple semantic components of a motion event on a single verb (cf. (13), (14)). Analytic marking is more common in neo-standard. This is, for instance, the case with phrasal verb constructions, which encode some semantic components of a motion event on the verb and others on an accompanying particle (feature 14, Table 1; cf. (15), (16)). Similar considerations can even apply to the presence of a demonstrative simultaneously before and after a noun (feature 6, Table 1; cf. (6)), at least if we assume that the demonstrative preceding the noun is (or tends to be) used as a marker of definiteness or specificity (see, e.g., Parenti 2001; cf. Diessel 2013) and the locative adverb following the noun indicates a distance contrast; see, e.g., *'sta stampante lì ha sempre dato dei problemi* ('that printer has always malfunctioned', lit. 'this printer there has always malfunctioned'; Cerruti 2009: 90), in which the prenominal demonstrative appears to be distance-neutral.

It can, therefore, be argued that literary standard features tend to cluster (see Figure 1, Section 3.1) not only as a result of situational formality and informational purposes, but also because of common grammatical properties, such as the preference for overt coding of oppositions and synthetic marking of meanings. Overt coding and synthetic marking indeed reflect the characteristics of explicitness and 'integration' (in the sense of Chafe 1982) which are cross-linguistically associated with written, formal, and educated varieties (see, e.g., Biber & Conrad 2009). Similarly, neo-standard features can be argued to cluster (cf. Figure 1) not only under the impulse of informal and multifunctional communication, but also because of shared

¹³ Different is the case of pronoun retention strategy (e.g., *il ragazzo che gli ho scritto una lettera* 'the boy to whom I wrote a letter', lit. 'the boy that to him I wrote a letter'; footnote 9), which combines a subordinating conjunction with a case-marked resumptive element.

grammatical properties, such as the tendency to leave inferable distinctions uncoded and convey meanings analytically. In fact, these properties are typical of vernaculars and, more generally, spoken varieties (as well as contact varieties; cf. Szmrecsanyi 2009; Haspelmath & Michaelis 2017) and consistent with the nature of neo-standard features, which originate mainly from spoken, informal, and uneducated varieties (see, e.g., Berruto 2012: 73–75).¹⁴

Leaving aside generalizations about intra-linguistic variation (cf. Kortmann 2004), it is worth noting that different typological profiles can coexist in the same diasystem and, most importantly, can take the form of different language varieties.

As shown in Grandi (2019, 2022; see also Ballarè & Inglese 2023), there are, in fact, significant similarities between the notions of language type and language variety. First and foremost, both notions are defined in terms of a correlation between linguistic features, steered by one or more factors. A language type can indeed be understood as a correlation between ‘values of structural properties’ (to use WALS terminology, cf. Section 3.2), which is determined by language-internal factors generally reflecting basic underlying parameters of linguistic structuring (e.g. locus of marking, alignment, head-directionality, etc.). For example, the head-directionality parameter appears to determine some recurrent correlations across languages, such as that between OV (Object-Verb), RelN (Relative-Noun) and NAdp (Noun-Adposition), on the one hand, and that between VO, NRel and AdpN, on the other hand (cf. Dryer 2013b, 2013c).

Similarly, a language variety is defined as a correlation between variants of linguistic variables (cf. Section 1) that is associated with language-external factors, such as the geographical origin and the social identity of the speakers, or situational characteristics. For example (with special reference to our case study), the opposition between formal and informal situations can be seen as underpinning some subsystemic correlations, such as that between relative pronoun strategy, gender distinction in third-person plural pronouns and single verb constructions (as found in literary standard Italian), on the one hand, and that between gap strategy, the lack of gender marking in third-person plural pronouns and phrasal verb constructions (as found in neo-standard Italian), on the other hand.

¹⁴ Needless to say, this is not at odds with the status of neo-standard features. They came to be used and accepted even in formal and educated speech, and to a certain extent in formal and educated writing (cf. footnote 6), as a result of a bottom-up process of (re)standardization (cf. Berruto 1987, 2012; Cerruti et al. 2017).

At the same time, the correlation between features that defines a language variety can be fostered not only by language-external factors but also by language-internal motivations. As argued above, for example, relative pronoun strategy tends to cluster with single verb constructions also due to the preference (that formal varieties show) for the synthetic marking of meanings, while gap strategy tends to cluster with the lack of gender contrast in third-person plural pronouns also because of the tendency (of informal varieties) to leave inferable distinctions uncoded. Moreover, as shown in Section 3.2, some of the co-occurring features that define a language variety represent a bundle of values of typologically relevant properties of language. This further establishes the notion of language variety as a suitable unit of analysis for both variationist sociolinguistics and typological research.

Finally, even the correlation between features that defines a language type can possibly be associated with both language-external and language-internal factors. This is particularly tenable when referring to areal types (cf. Grandi 2022: 139–140); i.e., groups of languages (such as those of the *Balkan Sprachbund*) that share a common set of linguistic features as a result of geographical proximity and language contact, rather than genetic relationships. One cannot fail to notice, moreover, that the traditional approach to linguistic typology (cf. Section 2) tends to ascribe a standard language to the type representing the standard variety of that language; in fact, the so-called “grammar-based” (or “type-based”) typology mainly draws on data coming from reference grammars and thus focuses on the set of linguistic features that is codified as the reference model for the standard usage of a language. From this perspective, therefore, even a language type can be argued to have its own sociolinguistic, language-external, *côté*.

4. Phonetic variation in a bilingual community

The second study this paper considers is a phonological analysis of the realization of /r/ in Italian spoken by a bilingual community in South Tyrol (Italy). In this area, Italian and German, more specifically the Tyrolean dialect,¹⁵ have been in stable contact for almost a century, a process that has led to the development of a contact variety of Italian, strongly influenced in its phonology by German. Together with the contact variety, the linguistic repertoire includes a regional variety, largely similar to

¹⁵ Tyrolean is a German dialect belonging to the Southern Bavarian group (Wiesinger 1990).

the surrounding ones (such as Trentino and Veneto Italian; Vietti 2017; Vietti & Mereu 2023), spoken mainly by monolingual or Italian-dominant bilingual speakers.



Figure 3: Bolzano-Bozen (South Tyrol, Italy) and the two neighboring Italian regions (Trentino and Veneto).

Without partitioning the data into distinct varieties, the high variability in the production of /r/, in this context of language contact, would result in a redundant and unlikely system of allophonic relations among 11 variants distinguished by two constriction locations and four manners of articulation (see Table 2). The phonetic variability of /r/ poses, on a general level, a problem of phonetic-phonological characterization of this class of sounds (Lindau 1985; Wiese 2011), but in this particular case, the extreme mutability of realization raises a specific issue of mapping between allophones and contexts. On the one hand, the high variability in the production of /r/ (especially if uvular) does not easily allow for distinguishing between allophony and coarticulation (Iskarous et al. 2012); on the other hand, the frequent shift and contact between languages in a bilingual setting increases the entropy of the relationship between allophones and phonetic contexts: high entropy means many allophones for one phonetic context vs. low entropy where ideally the relationship is one allophone for one context (cf. Vietti & Spreafico 2018).

An analysis of how allophones co-occur in patterns associated with the social attributes of the speakers makes it possible to observe the self-organization of rhotic sounds into coherent subsystems traceable to distinct varieties of Italian. Instead of

looking at the frequency of the feature alone, this case study demonstrates the need to consider the distribution of a grammatical feature within a speech community to reveal its structuring into subsystems (Scobbie 2006; Docherty & Foulkes 2014).

4.1 The structure of /r/ allophonic variation

The dataset used in this study is based on reading and spontaneous speech tasks obtained from a sample of 14 speakers who were born and lived in Bolzano (South Tyrol).¹⁶ The speakers are aged 24 and 38 years, with 9 female participants and 5 males). Among these speakers, 4 are categorized as sequential bilinguals with Italian as L1 and 8 as sequential bilinguals with Tyrolean as L1, while the remaining two participants are simultaneous bilinguals¹⁷ (Paradis 2007).

In this corpus, we identified a set of 9 distinct realizations¹⁸ of rhotic sounds, plus r-vocalization and deletion (see Table 2, Vietti & Spreafico 2016) that total 2276 tokens.

	Alveolar	Retroflex	Uvular
Trill	r		ʀ
Tap/Flap	ɾ	ɽ	ʀ̥
Approximant	ɹ		ʁ
Fricative	ʀ		χ ʁ

Table 2: Variants of /r/ in Bolzano Italian.

The primary goal of this analysis is to determine whether all /r/ variants pattern together in a single system of allophonic relations, however redundant, or whether variants are distributed among groups of speakers yielding distinct subsystems.

To observe patterns of association between allophones of /r/, phonetic contexts, and speakers, we used Correspondence Analysis (CA) for bivariate analysis and Multiple Correspondence Analysis (MCA) for multivariate analysis, two statistical

¹⁶ The analysis is essentially similar to Vietti & Spreafico (2016), except for an additional correspondence analysis (Variants of /r/ - First Language) and, above all, the typological interpretation of the results.

¹⁷ The speakers were classified on the basis of a questionnaire in which they self-assessed their linguistic background (see Kaland et al. 2016, Kaland et al. 2019).

¹⁸ The variants of /r/ were identified through qualitative analysis of the spectrographic representations of the segments in Galatà, Vietti & Spreafico (2016).

techniques that are functionally similar to PCA (cf. Section 1) but differ in their mathematical treatment of variance and apply specifically to categorical variables¹⁹ (Levshina 2015; Vietti & Spreafico 2016). The associations among values of the variables are expressed as chi-square distances and are represented graphically as points on a plot, uncovering the underlying structure of the data. Although the method of calculation and the metric by which correlations are expressed are different, the interpretation is similar to a PCA plot: the closeness between points is an indication of the strength of the association between categories of a variable. Similarly, the order of dimensions in MCA corresponds to the capacity to describe the variability of the data (variance in PCA, inertia in MCA). In this sense the first dimension is therefore the most important one.

To better understand the distinctive effect of allophone distribution across speakers, we performed a two-step analysis. In a first step, using Correspondence Analysis we observe associative patterns only among two linguistic variables (*Variants of /r/* and *Phonetic contexts*), simulating the existence of a single rhotic system of allophonic relations for Italian in South Tyrol. Variants of /r/ are coded with a labeling scheme in which the place precedes the manner of articulation, such as U_Trill (uvular trill) or A_Fric (alveolar fricative). Retroflex, vocalized /r/ and deletions are labeled as Retro, Vow and Del. Phonetic contexts analyzed are: #RV (initial syllable and word position as in *rosso* ‘red’), CRV (consonant cluster in syllable onset as in *fretta* ‘hurry’), VR# (final syllable and word position as in *bar* ‘bar, café’), VR.C (syllable coda followed by a consonant as in *forchetta* ‘fork’), VRRV (geminate /r/ as in *carro* ‘cart’), VRV (syllable onset in intervocalic position as in *cuore* ‘heart’).

In the second step, we introduce a non-linguistic predictor into our analysis, i.e. *First language* in the acquisition process. This variable divides the possible paths of bilingual acquisition into three categories: Italian (sequential bilinguals:²⁰ Italian L1

¹⁹ PCA aims to find new uncorrelated axes (principal components) that capture the maximum variance in the data. It is based on Euclidean distance between data points: the core idea is that if two data points are far apart in terms of their Euclidean distance, they are considered very different. MCA analyzes the “profiles” of individuals or categories. A “profile” is the set of relative frequencies of an individual’s responses across all categories, or a category’s prevalence across all individuals. MCA aims to find dimensions that highlight the greatest differences in these profiles. To do this, MCA uses a chi-squared distance metric that is sensitive to the relative differences between categories, not their absolute counts.

²⁰ Sequential bilinguals in South Tyrol acquire their second language around age 4-5, often at a pre-school stage.

– Tyrolean L2), German (sequential bilinguals: Tyrolean German L1 – Italian L2), and Italian-German (simultaneous bilinguals: Italian L1 + Tyrolean L1) (Paradis 2007). The second analysis aims at testing the groupings of linguistic variables by type of bilingual speaker, thus bringing out distinct and mutually incompatible allophonic subsystems, one related to the German-Italian contact variety and one to the regional variety.

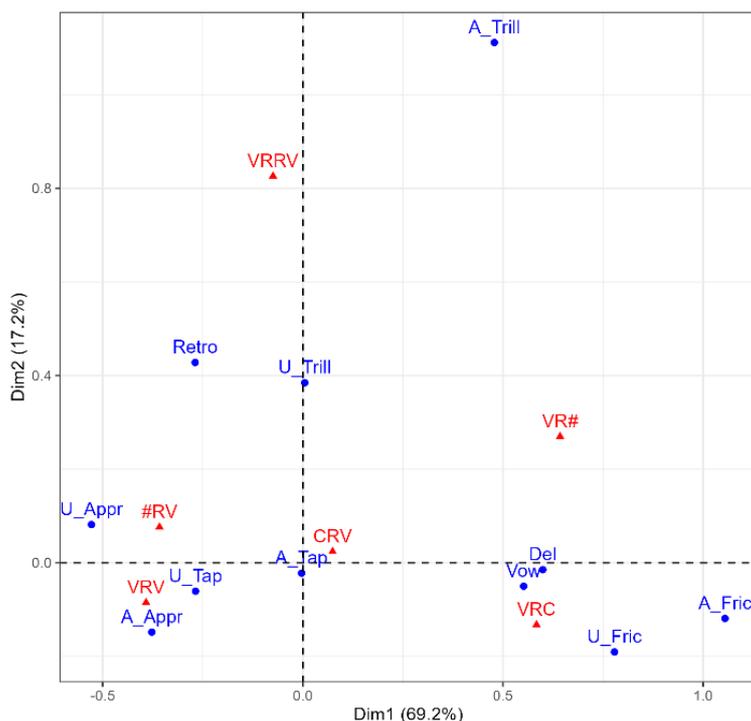


Figure 4: Plot of linguistic variants: /r/ variant (blue), Phonetic context (red).

The plot in Figure 4 reveals a general contextual distribution of allophones according to the manner of articulation (MoA). Dimension 1 distinguishes fricatives (on the right) from approximants (on the left) based on position within the syllable (coda vs. onset). Rhotic vocalization and deletions are found on the side of the syllabic coda contexts, while the taps (uvular and alveolar) are in a central position, closer to the origin of the plane, meaning that they are less context-dependent variants. Dimension 2 is mainly responsible for the distinction between the two trills and the retroflex and all other variants.²¹ The two trills are weakly associated with VRRV, a phonetic

²¹ The retroflex is a low-frequency phonetic realization in the corpus, the presence of which can be traced back to the influence of Veneto Italian on the Bolzano variety (Canepari 1986; Vietti 2017).

context in which the consonantal length is expressed as contrastive in Italian (Bertinetto & Loporcaro 2005; Krämer 2009). In the phonemic system of Italian spoken in Bolzano, the consonantal length distinction is not encoded simply as a gemination of the same segment (as, for example, for /l/ in *pala* ['pa:la] 'shovel' vs. *palla* ['palla] 'ball'), but through two distinct phonetic realizations: tap as the default variant and trill for word-internal geminate contexts.

In this first analysis, which simulates a single allophonic system, two variants similar in manner but distinct in place of articulation may occur in the same phonetic context. This results in a highly redundant system in which sounds are potentially contrastive because the two conditions of phonemicity seem to be met: both the distribution of allophones by manner is unpredictable, and the variants are sufficiently distinct from a perceptual point of view (Renwick & Ladd 2016; Kiparsky 2015). Although the phonological conditions for phonemicity are satisfied, the lack of lexical distinctiveness does not establish an opposition with place of articulation between MoA rhotic pairs. In addition, the hypothetical system that emerges from this distributional analysis would also be very unlikely from a typological point of view, as consonant inventories containing two rhotics rarely contrast by place of articulation²² (Maddieson 1988: 84–88). Rather, it seems, given the pressure of language contact, that we are dealing with a diasystem in which two subsystems are overlapping in phonetic and distributional properties (Pulgram 1964).

In the second plot (Figure 5), the introduction of correlations between linguistic variables and speakers' properties disentangles the apparent superposition of the subsystems by separating them into two groups of rhotic sounds. When the predictor *First Language* is added to the Multiple Correspondence Analysis, dimension 1 divides variants of /r/ by place of articulation: the uvular rhotics are found on the right side of the plane along with Tyrolean L1 sequential bilinguals (Ger) and simultaneous bilinguals (Ita-Ger), while on the left side, the alveolar rhotics are associated with Italian L1 sequential bilinguals (Ita).

²² In some southern Swedish dialects, for example, dorsal and coronal /r/ are present in the inventory, but they are in complementary distribution (Engstrand et al. 2007).

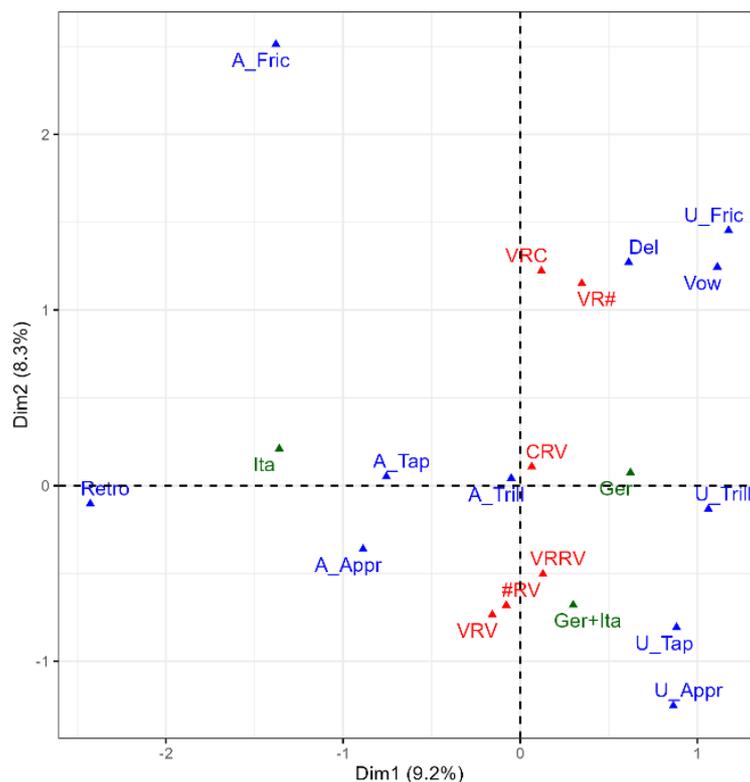


Figure 5: Plot of sociolinguistic variants: /r/ variant (blue), Phonetic context (red), First language of acquisition (green).

In this dimension, however, the two groups differ with respect to internal cohesion. The uvular rhotics, vocalizations and deletions are very close to each other and more clearly identify a German variety of Italian, due to the association with Tyrolean as L1 (Ger). In contrast, the alveolar rhotics show greater differences, e.g. retroflex takes on more extreme values as it correlates with one speaker's Veneto regional Italian,²³ while the alveolar trill has a central position with respect to the plot. Alveolar trill position is related to the fact that while L1 Italian speakers do not categorically use back variants of /r/, some simultaneous bilinguals (Ger+Ita) and Tyrolean L1 sequential bilinguals (Ger) may use front variants (e.g., the alveolar trill) when speaking Italian. The imperfect separation between variants by place of articulation and type of bilingual justifies the reduced ability of MCA to explain the variance in the data (9.2% and 8.3% for the first two dimensions). However, since no speaker

²³ As already noted in Vietti (2017), the Italian spoken in Bolzano is the result of a process of koineization between geographical varieties spoken in Italian regions bordering South Tyrol, such as Veneto (see Fig. 3). The Veneto component forms the basis of the Bolzano koine and today has taken on the social meaning of a broad local accent.

uses both variants of PoA in Italian, this means the subsystems likely cannot coexist in a single speaker's production and therefore that the integration of social information in the MCA is crucial to bring out the internal structure of the data.

If we consider Dimension 2, the variants seem to follow a rather weak pattern of association between phonetic contexts and manner of articulation. Pairs of /r/-sounds similar in manner are loosely mapped to the same phonetic contexts according to a pattern that seems to favor many-to-many relationships rather than a "one allophone-one context" relationship (Hall 2013). For example, the two fricative /r/ sounds (A_Fric and U_Fric) are located in the same upper half-plane and are weakly mapped to postvocalic /r/ settings (VRC and VR#), but the same is true for /r/ vocalizations and deletions. The introduction of the variable *First language* brings some order to the allophonic variation of /r/, but does not yet lead to the clear identification of coherent allophonic subsystems. One possible explanation for the increased variability and ambiguity of associations may be that the introduction of social information, on the one hand, allows for the emergence of two potential varieties of Italian, distinct by place of articulation, but, on the other hand, covers up incompatible allophone systems. For example, the alveolar trill is used by both Italian-L1 sequential bilinguals and simultaneous bilinguals (and also some Tyrolean-L1 sequential bilinguals) but arguably with different distributional rules. Differences between systems of allophonic relations are thus masked in a single flattened representation.²⁴ However, even with these limitations, the analysis reveals constraints on the high variability of /r/ that allow us to understand more clearly the nature of allophonic variation and the structure of the data.

The order manifests itself on two levels: the first concerns the partition of variation into two distinct subsystems on the basis of the social distribution of variants; the second shows the probabilistic nature of the relationships between allophones and phonetic contexts (Hall 2013), which is certainly amplified in a situation of language contact.

The analysis of the covariation bonds between linguistic and social features allows us to understand the relationship between the levels in which sociolinguistic reality

²⁴ Correspondence analysis on two variables (Variants of /r/ and First language), without considering the association with phonetic contexts, indicates a high variance explained by the first dimension (83.8%). This means that the social predictor is able to clearly order allophones by language variety, but also that within each language variety, the association between /r/ allophones and phonetic contexts does not generate fully coherent subsystems.

is organized into grammatical subsystems or varieties, which then is nothing more than the recognition of the orderly heterogeneity of language variation (Weinreich et al. 1968; Beaman & Guy 2022). The organization of a language into coexisting subsystems may or may not be reflective of variation at the (micro)typological level, but failure to recognize this multilevel, multidimensional articulation can lead to misinterpretation of the typological nature of a language system, whether under a grammar-based or frequency-based approach.

5. Conclusions

The aim of the paper has been to show the relevance of the notion of language variety for research in typological linguistics and, more broadly, to emphasize the continuity between variation across languages and their internal structuring. We have defined language varieties as minimally distinct grammatical subsystems within a larger language system, in which the links binding linguistic features emerge through co-occurrence with social units, such as the communicative situation or the speaker. As we have shown, the features of a language variety can also exhibit linguistic coherence not only for local communicative purpose, but also at the level of typological pattern in the expression of grammatical categories.

Underlying our analysis is the assumption that different approaches to typological research involve distinct interpretations of what should be considered a unit of linguistic observation. In type- or grammar-based typology, the unit of analysis is a descriptive account of a grammatical system, for which the linguistic feature to be analyzed will be of a discrete type and will take on a single categorical value for that grammar. For example, in the case of the typology of object and verb order, although internal variability is recognized, a dominant type will represent an entire grammar.

In corpus-based typology, the unit of analysis is a corpus and thus the (value of a) linguistic feature is represented as its frequency distribution. The typological variation of a linguistic feature can be analyzed in two ways, marginal and conditional, that is, univariate or bi- and multivariate. In the example of verb and object order, this means observing (a) the marginal frequency of the possible orders or (b) the conditional frequency, i.e. searching for dependencies between the frequency of orders and other relevant phenomena in the linguistic structure such as e.g. case marking or agreement. This second possibility can also be investigated more

abstractly with entropy measures that continuously capture the internal variability of a distribution (Koplenig et al. 2017; Levshina 2019).

According to this line of research, the reliability of inference relies on the representativeness of a corpus and comparability between corpora. The former ensures the generalizability of the inference, while the latter prevents uncontrolled variables from exerting undesirable effects on the variation of the observed linguistic feature. The study of any linguistic phenomenon from a corpus-based typology perspective is conducted on homogeneous datasets, i.e. based on the same type of text, so as to control for possible sources of covariation.

If comparing data from different languages requires a necessary selection of homogeneous datasets, what if corpora containing balanced sociolinguistic variation could be compared? What steps should be taken to make the correct crosslinguistic inferences from the data? In this paper, we aimed to focus on the covariation structure that links linguistic phenomena and social information and how the analysis of this structure can help identify distinct grammatical subsystems within an overall linguistic system. In particular, the diasystem of a language can be usefully decomposed into language varieties through the association between bundles of linguistic features and meaningful social units, such as the communicative situation and the social characteristics of the speaker. For this purpose, we employed statistical methods such as PCA and MCA that allow us to observe the emergence of patterns of association between variables.

The analysis conducted on these two cases of sociolinguistic variation in Italian show how the notion of language variety can have relevant implications for research in corpus-based typology. First, a corpus-based approach allows for the empirical identification of language varieties as emergent structures of covariation between linguistic and social elements. The first case study, for example, illustrates how the two identified standard varieties of Italian can coexist in speakers' competence to express stylistic differences along a scale of greater or lesser formality. Second, the varieties can also be linguistically coherent, that is, show how the socio-communicative purpose of usage can lead to the harmonic selection of categories that linguistic features can take on. As a result, the structuring of a language into relatively stable varieties makes micro-typological variation possible, as is the case with the different grammatical encoding strategies of the two standard varieties of Italian (Section 3).

Finally, varieties may also be subsystems related to groups of speakers and not necessarily to contexts of language use, which may be non-uniformly distributed in the speech community. This implies the presence of varieties that do not co-exist in the speakers' competence and therefore are potentially more dissimilar to each other. In the case of the study of allophonic variation of /r/ in contact varieties of Italian, failure to divide the micro-system of /r/ into two sets of variants would lead to a grammatical misdescription. In other words, a marginal analysis of the distribution of variants not conditioned by social variables would result in a phonemic category with anomalous allophonic variability, characterized by high contextual entropy, creating extreme difficulty in matching allophones and phonetic contexts.

Our study is limited in several respects, mainly because the two case studies are very specific in the linguistic features investigated and are also related to only one language, Italian. In addition, analyses of this kind are severely constrained methodologically by the nature of the source data and the difficulty in scaling crosslinguistically due to the lack of comparable and representative sociolinguistic corpora. However, we have illustrated how the concept of 'language variety', as a grammatical subsystem that can be identified through its social distribution, can be a useful heuristic in the corpus-based typology research program. Language variety plays a pivotal role within the hierarchical data structure that links individual linguistic features at one end to the entire population of sentence tokens (i.e., the language corpus) at the other end.

Abbreviations

1 = 1 st person	M = masculine	PTCP = participle
3 = 3 rd person	NEG = negative	REL = relative
COMP = complementizer	PFV = perfective	SBJV = subjunctive
COND = conditional	PL = plural	SG = singular
F = feminine	PRS = present	
IPFV = imperfective	PST = past	

References

Auer, Peter. 1997. Co-Occurrence Restrictions between Linguistic Variables: A Case for Social Dialectology, Phonological Theory and Variation Studies. In Hinskens,

- Frans L. & Van Hout, Roeland & Wetzels, W. Leo (eds.), *Current Issues in Linguistic Theory*, vol. 146, 69. Amsterdam: John Benjamins Publishing Company. (doi:[10.1075/cilt.146.05aue](https://doi.org/10.1075/cilt.146.05aue))
- Ballarè, Silvia. 2020. L'italiano neo-standard oggi: stato dell'arte. *Italiano LinguaDue* 12(2). 469–492. (doi:[10.13130/2037-3597/15013](https://doi.org/10.13130/2037-3597/15013))
- Ballarè, Silvia & Inglese, Guglielmo. 2023. Analyzing language variation: Where sociolinguistics and linguistic typology meet. In Ballarè, Silvia & Inglese, Guglielmo (eds.), *Sociolinguistic and Typological Perspectives on Language Variation*, 1–27. Berlin, Boston: De Gruyter.
- Beaman, Karen V. & Sering, Konstantin. 2022. Measuring change in lectal coherence across real- and apparent-time. In Beaman, Karen & Guy, Gregory R. (eds.), *The coherence of linguistic communities: orderly heterogeneity and social meaning*, 87–105. New York: Routledge.
- Beaman, Karen V. & Guy, Gregory R. (eds.). 2022. *The coherence of linguistic communities: orderly heterogeneity and social meaning* (Routledge Studies in Sociolinguistics). New York, NY: Routledge.
- Beavers, John & Levin, Beth & Tham, Shiao Wei. 2010. The typology of motion expression revisited. *Journal of Linguistics* 46(3). 331–377.
- Benedetti, Marina & Ricca, Davide. 2002. The systems of deictic place adverbs in the Mediterranean: Some general remarks. *Mediterranean languages. Papers from the MEDTYP workshop*, 13–32. Bochum: Brockmeyer.
- Bernini, Giuliano. 2010. Word classes and the coding of spatial relations in motion events: A contrastive typological approach. In Marotta, Giovanna & Lenci, Alessandro & Meini, Linda & Rovai, Francesco (eds.), *Space in Language. Proceedings of the Pisa International Conference*, 29–52. Pisa: ETS.
- Berruto, Gaetano. 1987. *Sociolinguistica dell'italiano contemporaneo*. Roma: La Nuova Italia Scientifica.
- Berruto, Gaetano. 2004. Sprachvarietät - Sprache (Gesamtsprache, historische Sprache). . In Ammon, Ulrich & Dittmar, Norbert & Mattheier, Klaus J. & Trudgill, Peter (eds.), *Sociolinguistics/Soziolinguistik. Ein internationales Handbuch zur Wissenschaft von Sprache und Gesellschaft*, vol. 1, 188–195. 2nd edn. Berlin, New York: De Gruyter.
- Berruto, Gaetano. 2009. Identifying dimensions of linguistic variation in a language space. In Auer, Peter & Schmidt, Jürgen Erich (eds.), *Language and Space. An*

- International Handbook of Linguistic Variation. Vol. 1. Theories and Methods*, 226–241. Berlin, New York: Mouton de Gruyter. (doi:[10.1515/9783110220278.226](https://doi.org/10.1515/9783110220278.226))
- Berruto, Gaetano. 2012. *Sociolinguistica dell'italiano contemporaneo*. Roma: Carocci.
- Berruto, Gaetano. 2017. What is changing in Italian today? Phenomena of restandardization in syntax and morphology: an overview. In Cerruti, Massimo & Crocco, Claudia & Marzo, Stefania (eds.), *Towards a New Standard*, 31–60. Berlin, New York: De Gruyter. (doi:[10.1515/9781614518839-002](https://doi.org/10.1515/9781614518839-002))
- Berruto, Gaetano. 2019. La nozione di ‘varietà di lingua’: una categoria obsoleta? In Bidese, Ermenegildo & Casalicchio, Jan & Moroni, Manuela Caterina (eds.), *La linguistica vista dalle Alpi: teoria, lessicografia e multilinguismo: studi in onore di Patrizia Cordin = Linguistic views from the Alps: language theory, lexicography and multilingualism: studies in honor of Patrizia Cordin* (Studia Romanica et Linguistica 57), 213–236. Berlin: Peter Lang.
- Bertinetto, PierMarco & Loporcaro, Michele. 2005. The sound pattern of Standard Italian, as compared with the varieties spoken in Florence, Milan and Rome. *Journal of the Phonetic Association* 35(2). 131–151.
- Biber, Douglas & Conrad, Susan. 2009. *Register, Genre, and Style*. Cambridge: Cambridge University Press. (doi:[10.1017/CBO9780511814358](https://doi.org/10.1017/CBO9780511814358))
- Blommaert, Jan. 2010. *The Sociolinguistics of Globalization*. 1st edn. Cambridge: Cambridge University Press. (doi:[10.1017/CBO9780511845307](https://doi.org/10.1017/CBO9780511845307))
- Canepari, Luciano. (1986). *Lingua italiana nel Veneto*. CLESP.
- Cerruti, Massimo. 2009. *Strutture dell'italiano regionale: morfosintassi di una varietà diatopica in prospettiva sociolinguistica* (Spazi comunicativi 7). Frankfurt am Main: Peter Lang.
- Cerruti, Massimo. 2017. Changes from below, changes from above: relative constructions in contemporary Italian. In Cerruti, Massimo & Crocco, Claudia & Marzo, Stefania (eds.), *Towards a New Standard*, 61–88. Berlin, New York: De Gruyter. (doi:[10.1515/9781614518839-003](https://doi.org/10.1515/9781614518839-003))
- Cerruti, Massimo & Crocco, Claudia & Marzo, Stefania (eds.). 2017. *Towards a New Standard: Theoretical and Empirical Studies on the Restandardization of Italian*. De Gruyter. (doi:[10.1515/9781614518839](https://doi.org/10.1515/9781614518839))
- Cerruti, Massimo & Vietti, Alessandro. 2022. Identifying language varieties: Coexisting standards in spoken Italian. In Beaman, Karen V. & Guy, Gregory R. (eds.), *The Coherence of Linguistic Communities. Orderly Heterogeneity and Social Meaning* (Routledge Studies in Sociolinguistics), 262–280. New York: Routledge.

- Chafe, Wallace L. 1982. Integration and Involvement in speaking, writing, and oral literature. In Tannen, Deborah (ed.), *Spoken and written language: Exploring orality and literacy*, 35–53. Norwood: Ablex.
- Comrie, Bernard & Dryer, Matthew S. & Gil, David & Haspelmath, Martin. 2013. Introduction (v2020.3). *The world atlas of language structures (online)*. Zenodo. (Ed. Dryer, Matthew S. & Haspelmath, Martin.) (doi:[10.5281/zenodo.7385533](https://doi.org/10.5281/zenodo.7385533))
- Comrie, Bernard & Kuteva, Tania. 2013. Relativization on Obliques. *The World Atlas of Language Structures Online (v2020.3)*. Zenodo. (Ed. Dryer, Matthew S. & Haspelmath, Martin.) (doi:[10.5281/zenodo.7385533](https://doi.org/10.5281/zenodo.7385533))
- Cordin, Patrizia. 2001. I pronomi riflessivi. In Renzi, Lorenzo & Salvi, Giampaolo & Cardinaletti, Anna (ed.), *Grande grammatica italiana di consultazione*, vol. I, 607–617. Bologna: Il Mulino.
- Da Milano, Federica. 2015. Italian. In Jungbluth, Konstanze & Da Milano, Federica (eds.), *Manual of Deixis in Romance Languages*, 59–74. De Gruyter. (doi:[10.1515/9783110317732-006](https://doi.org/10.1515/9783110317732-006))
- Dahl, Östen. 1990. Standard Average European as an Exotic Language. In Bechert, Johannes & Bernini, Giuliano & Buridant, Claude (eds.), *Toward a Typology of European Languages (Empirical Approaches to Language Typology)*, 3–8. Berlin: Mouton de Gruyter.
- Diessel, Holger. 2013. Distance Contrasts in Demonstratives. *The World Atlas of Language Structures Online (v2020.3)*. Zenodo. (Ed. Dryer, Matthew S. & Haspelmath, Martin.) (doi:[10.5281/zenodo.7385533](https://doi.org/10.5281/zenodo.7385533))
- Docherty, Gerard J. & Foulkes, Paul. 2014. An evaluation of usage-based approaches to the modelling of sociophonetic variability. *Lingua* 142. 42–56. (doi:[10.1016/j.lingua.2013.01.011](https://doi.org/10.1016/j.lingua.2013.01.011))
- Dryer, Matthew S. 2013a. Order of Demonstrative and Noun. *The World Atlas of Language Structures Online (v2020.3)*. Zenodo. (Ed. Dryer, Matthew S. & Haspelmath, Martin.) (doi:[10.5281/zenodo.7385533](https://doi.org/10.5281/zenodo.7385533))
- Dryer, Matthew S. 2013b. Relationship between the Order of Object and Verb and the Order of Relative Clause and Noun. *The World Atlas of Language Structures Online (v2020.3)*. Zenodo. (Ed. Dryer, Matthew S. & Haspelmath, Martin.) (doi:[10.5281/zenodo.7385533](https://doi.org/10.5281/zenodo.7385533))
- Dryer, Matthew S. 2013c. Relationship between the Order of Object and Verb and the Order of Adposition and Noun Phrase. *The World Atlas of Language Structures Online*

- (v2020.3). Zenodo. (Ed. Dryer, Matthew S. & Haspelmath, Martin.) (doi:[10.5281/zenodo.7385533](https://doi.org/10.5281/zenodo.7385533))
- Dryer, Matthew S. & Haspelmath, Martin (eds.). 2013. *WALS Online* (v2020.3). Zenodo. (doi:[10.5281/zenodo.7385533](https://doi.org/10.5281/zenodo.7385533))
- Engstrand, Olle & Frid, Johan & Lindblom, Björn. 2007. A Perceptual Bridge Between Coronal and Dorsal /r/. In Solé, Maria-Josep & Beddor, Patrice Speeter & Ohala, Manjari (eds.), *Experimental Approaches to Phonology*, 175–191. Oxford: Oxford University Press. (doi:[10.1093/oso/9780199296675.003.0012](https://doi.org/10.1093/oso/9780199296675.003.0012))
- Galatà, Vincenzo & Spreafico, Lorenzo & Vietti, Alessandro & Kaland, Constantijn. 2016. An acoustic analysis of /r/ in Tyrolean. *INTERSPEECH 2016*, 1002–1006. San Francisco, USA. (doi:[10.21437/Interspeech.2016-418](https://doi.org/10.21437/Interspeech.2016-418))
- Gerdes, Kim & Kahane, Sylvain & Chen, Xinying. 2021. Typometrics: From Implicational to Quantitative Universals in Word Order Typology. *Glossa: a journal of general linguistics*. Open Library of Humanities 6(1). (doi:[10.5334/gjgl.764](https://doi.org/10.5334/gjgl.764))
- Ghyselen, Anne-Sophie & De Vogelaer, Gunther. 2018. Seeking Systematicity in Variation: Theoretical and Methodological Considerations on the “Variety” Concept. *Frontiers in Psychology* 9. (doi:[10.3389/fpsyg.2018.00385](https://doi.org/10.3389/fpsyg.2018.00385))
- Grandi, Nicola. 2019. Che tipo, l’italiano neostandard! In Moretti, Bruno & Kunz, Aline & Natale, Silvia & Krakenberger, Etna (eds.), *Le tendenze dell’italiano contemporaneo rivisitate*, 59–74. Milano: Officinaventuno.
- Grandi, Nicola. 2022. La variazione tra le lingue e nelle lingue. In Grandi, Nicola & Mauri, Caterina (eds.), *La tipologia linguistica. Unità e diversità nelle lingue del mondo*, 121–173. Roma: Carocci.
- Gumperz, John. 1968. The Speech Community. In Sills, David L. (ed.), *International Encyclopedia of the Social Sciences*, 381–386. London: Macmillan.
- Guy, Gregory R. 2013. The Cognitive Coherence of Sociolects: How Do Speakers Handle Multiple Sociolinguistic Variables? *Journal of Pragmatics* 52. 63–71.
- Guy, Gregory R. & Hinskens, Frans. 2016. Linguistic coherence: Systems, repertoires and speech communities. *Lingua* 172–173. 1–9. (doi:[10.1016/j.lingua.2016.01.001](https://doi.org/10.1016/j.lingua.2016.01.001))
- Guy, Gregory R. & Oushiro, Livia & Mendes, Ronald Beline. 2022. Indexicality and coherence. In Beaman, Karen & Guy, Gregory R. (eds.), *The coherence of linguistic communities: orderly heterogeneity and social meaning*, 53–68. New York: Routledge.

- Hall, Kathleen Currie. 2013. A typology of intermediate phonological relationships. *The Linguistic Review*. De Gruyter Mouton 30(2). 215–275. (doi:[10.1515/tlr-2013-0008](https://doi.org/10.1515/tlr-2013-0008))
- Haspelmath, Martin. 2018. How comparative concepts and descriptive linguistic categories are different. In Van Olmen, Daniël & Mortelmans, Tanja & Brisard, Frank (eds.), *Aspects of Linguistic Variation* (Trend in Linguistics 324), 83–114. Berlin, New York: Mouton de Gruyter. (doi:[10.1515/9783110607963-004](https://doi.org/10.1515/9783110607963-004)) (Accessed May 29, 2023.)
- Haspelmath, Martin. 2023. Comparing reflexive constructions in the world's languages. In Janic, Katarzyna & Puddu, Nicoletta & Haspelmath, Martin (eds.), *Reflexive constructions in the world's languages* (Research on Comparative Grammar 3), 19–62. Berlin: Language Science Press.
- Haspelmath, Martin & Michaelis, Susanne. 2017. Analytic and synthetic: Typological change in varieties of European languages. In Buchstaller, Isabelle & Siebenhaar, Beat (eds.), *Language variation – European perspectives VI: Selected papers from the 8th International Conference on Language Variation in Europe (ICLaVE 8), Leipzig 2015*, 3–22. Amsterdam: Benjamins.
- Hinskens, Frans & Guy, Gregory R. (eds.). 2016. *Coherence, covariation and bricolage. Various approaches to the systematicity of language variation*. Thematic issue of *Lingua* 172/173.
- Horvath, Barbara & Sankoff, David. 1987. Delimiting the Sydney Speech Community. *Language in Society* 16. 179–204.
- Hudson, Richard Anthony. 1980. *Sociolinguistics* (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.
- Hymes, Dell. 1967. Models of the interaction of language and social setting. *Journal of Social Issues* 23(2). 8–28.
- Iacobini, Claudio & Vergaro, Carla. 2012. Manner of motion verbs in Italian: semantic distinctions and interlingual comparisons. In Ferreri, Silvana (ed.), *Lessico e lessicologia*, 71–87. Roma: Bulzoni.
- Iskarous, Khalil & McDonough, Joyce & Whalen, D. H. 2012. A gestural account of the velar fricative in Navajo. *Laboratory Phonology* 3(1). 195–210. (doi:[10.1515/lp-2012-0011](https://doi.org/10.1515/lp-2012-0011))
- Kaland, Constantijn & Galatà, Vincenzo & Spreafico, Lorenzo & Vietti, Alessandro. 2016. /r/ as language marker in bilingual speech production and perception.

- INTERSPEECH* 2016, 515–519. San Francisco, USA. (doi:[10.21437/Interspeech.2016-418](https://doi.org/10.21437/Interspeech.2016-418))
- Kaland, Constantijn & Galatà, Vincenzo & Spreafico, Lorenzo & Vietti, Alessandro. 2019. Which Language R You Speaking? /r/ as a Language Marker in Tyrolean and Italian Bilinguals. *Language and Speech* 62(1). 137–163. (doi:[10.1177/0023830917746551](https://doi.org/10.1177/0023830917746551))
- Keenan, Edward L. & Comrie, Bernard. 1977. Noun phrase accessibility and Universal Grammar. *Linguistic Inquiry* 8. 63–99.
- Kiparsky, Paul. 2015. New perspectives in historical linguistics. In Bower, Claire & Evans, Bethwyn (eds.), *The Routledge Handbook of Historical Linguistics*, 64–102. London, New York: Routledge.
- Klein, Wolfgang. 2004. Varietätengrammatik. In Ammon, Ulrich & Dittmar, Norbert & Mattheier, Klaus J. (eds.), *Soziolinguistik. Ein internationales Handbuch zur Wissenschaft von Sprache und Gesellschaft*, 1163–1172. Berlin, New York: Mouton de Gruyter.
- Koplenig, Alexander & Meyer, Peter & Wolfer, Sascha & Müller-Spitzer, Carolin. 2017. The statistical trade-off between word order and word structure – Large-scale evidence for the principle of least effort. *PLOS ONE* 12(3). e0173614. (Ed. Smith, Kenny.) (doi:[10.1371/journal.pone.0173614](https://doi.org/10.1371/journal.pone.0173614))
- Kortmann, Bernd (ed.). 2004. *Dialectology meets typology: dialect grammar from a cross-linguistic perspective* (Trends in Linguistics Studies and Monographs 153). Berlin: Mouton de Gruyter.
- Krämer, Martin. 2009. *The phonology of Italian* (Oxford Linguistics). Oxford, New York: Oxford University Press.
- Kristiansen, Tore & Coupland, Nikolas (eds.). 2011. *Standard languages and language standards in a changing Europe* (Standard Language Ideology in Contemporary Europe 1). Oslo: Novus Press.
- Labov, William. 1966a. *The Social Stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.
- Labov, William. 1966b. The linguistic variable as structural unit. *Washington Linguistics Review* 3. 4–22.
- Levshina, Natalia. 2015. *How to do Linguistics with R. Data exploration and statistical analysis*. Amsterdam, Philadelphia: Benjamins.

- Levshina, Natalia. 2019. Token-based typology and word order entropy: A study based on Universal Dependencies. *Linguistic Typology*. De Gruyter Mouton 23(3). 533–572. (doi:[10.1515/lingty-2019-0025](https://doi.org/10.1515/lingty-2019-0025))
- Levshina, Natalia. 2022. Corpus-based typology: applications, challenges and some solutions. *Linguistic Typology*. De Gruyter Mouton 26(1). 129–160. (doi:[10.1515/lingty-2020-0118](https://doi.org/10.1515/lingty-2020-0118))
- Lindau, Mona. 1985. The story of /r/. In Fromkin, Victoria (ed.), *Phonetic linguistics: essays in honor of Peter Ladefoged*, 157–168. Orlando: Academic Press.
- Maddieson, Ian. 1988. *Patterns of Sounds*. Cambridge: Cambridge University Press.
- Mauri, Caterina & Ballarè, Silvia & Gorla, Eugenio & Cerruti, Massimo & Suriano, Francesco. KIParla corpus: A new resource for spoken Italian. In Bernardi, Raffaella & Navigli, Roberto & Semeraro, Giovanni (eds.), *Proceedings of the Sixth Italian Conference on Computational Linguistics*, vol. 2841, 1–7. CEUR Workshop Proceedings.
- Milroy, James & Milroy, Lesley. 1997. Varieties and variation. In Coulmas, Florian (ed.), *The Handbook of Sociolinguistics*, 47–64. Oxford: Blackwell.
- Paradis, Johanne. 2007. Early bilingual and multilingual acquisition. In Auer, Peter & Wei, Li (eds.), *Handbook of Multilingualism and Multilingual Communication*, 15–44. Mouton de Gruyter. (doi:[10.1515/9783110198553.1.15](https://doi.org/10.1515/9783110198553.1.15)) (Accessed April 30, 2024.)
- Parenti, Alessandro. 2001. Sulla semantica dei dimostrativi. *Archivio Glottologico Italiano* 86(2). 174–193.
- Pulgram, Ernst. 1964. Structural comparison, diasystems, and dialectology. *Linguistics* 2(4). 66–82. (doi:[10.1515/ling.1964.2.4.66](https://doi.org/10.1515/ling.1964.2.4.66))
- Renwick, Margaret & Ladd, Robert D. 2016. Phonetic Distinctiveness vs. Lexical Contrastiveness in Non-Robust Phonemic Contrasts. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 7(1). 19. (doi:[10.5334/labphon.17](https://doi.org/10.5334/labphon.17))
- Sankoff, David. 2004. Variable Rules. In Ammon, Ulrich & Dittmar, Norbert & Mattheier, Klaus J. & Trudgill, Peter (eds.), *Soziolinguistik. Ein internationales Handbuch zur Wissenschaft von Sprache und Gesellschaft*, vol. 2, 1150–1163. Berlin, New York: De Gruyter. (doi:[10.1515/9783110171488.2.7.1150](https://doi.org/10.1515/9783110171488.2.7.1150))
- Schnell, Stefan & Schiborr, Nils. 2022. Crosslinguistic Corpus Studies in Linguistic Typology. *Annual Review of Linguistics* 8. 171–191. (doi:[10.1146/annurev-linguistics-031120-104629](https://doi.org/10.1146/annurev-linguistics-031120-104629))

- Scobbie, J. M. 2006. (R) as a Variable. In Brown, Keith (ed.), *Encyclopedia of Language & Linguistics (Second Edition)*, 337–344. Oxford: Elsevier. (doi:[10.1016/B0-08-044854-2/04711-8](https://doi.org/10.1016/B0-08-044854-2/04711-8))
- Siewierska, Anna. 2013. Gender Distinctions in Independent Personal Pronouns. *The World Atlas of Language Structures Online* (v2020.3). Zenodo. (Ed. Dryer, Matthew S. & Haspelmath, Martin.) (doi:[10.5281/zenodo.7385533](https://doi.org/10.5281/zenodo.7385533))
- Szmrecsanyi, Benedikt. 2009. Typological parameters of intralingual variability: Grammatical analyticity versus syntheticity in varieties of English. *Language Variation and Change* 21(3). 319–353. (doi:[10.1017/S0954394509990123](https://doi.org/10.1017/S0954394509990123))
- Talmy, Leonard. 1985. Lexicalisation patterns: semantic structure in lexical forms. In Shopen, Timothy (ed.), *Language typology and syntactic description*, 36–149. Cambridge: Cambridge University Press.
- Talmy, Leonard. 2000. *Toward a Cognitive Semantics*. MIT Press.
- Vietti, Alessandro. 2017. Italian in Bozen/Bolzano: the formation of a “new dialect.” In Cerruti, Massimo & Crocco, Claudia & Marzo, Stefania (eds.), *Towards a new standard: theoretical and empirical studies on the restandardization of Italian*, 176–212. Berlin: De Gruyter.
- Vietti, Alessandro. 2019. La varietà di lingua come insieme di tratti coerenti: verso una caratterizzazione empirica. *Rivista Italiana di Dialettologia* 43. 11–32.
- Vietti, Alessandro & Mereu, Daniela. 2023. Mid vowels at the crossroads between standard and regional Italian. *Sociolinguistica* 37(1). 17–39. (doi:[10.1515/soci-2022-0033](https://doi.org/10.1515/soci-2022-0033))
- Vietti, Alessandro & Spreafico, Lorenzo. 2016. Lo strano caso di /r/ a Bolzano: problemi di interfaccia. In Iacobini, Claudio & Voghera, Miriam & Savy, Renata (eds.), *Livelli di analisi e fenomeni di interfaccia*, 263–281. Roma: Bulzoni.
- Vietti, Alessandro & Spreafico, Lorenzo. 2018. Sprachkontakt in der Phonologie bilingualer Sprecher des Tirolischen. In Rabanus, Stefan (ed.), *Deutsch als Minderheitensprache in Italien. Theorie und Empirie kontaktinduzierten Sprachwandels* (Germanistische Linguistik), vol. 239–240, 49–78. Hildesheim: Georg Olms Verlag.
- Villena-Ponsoda, Juan-Andrés & Vida-Castro, Matilde. 2020. Variation, identity and indexicality in southern Spanish: On the emergence of a new variety in urban Andalusia. In Cerruti, Massimo & Tsiplakou, Stavroula (eds.), *Intermediate Language Varieties: Koinai and regional standards in Europe* (Studies in Language Variation), vol. 24, 149–182. Amsterdam: Benjamins. (doi:[10.1075/silv.24.07vil](https://doi.org/10.1075/silv.24.07vil))

- Wälchli, Bernhard. 2009. Data reduction typology and the bimodal distribution bias. *Linguistic Typology* 13(1). 77–94. (doi:[10.1515/LITY.2009.004](https://doi.org/10.1515/LITY.2009.004))
- Walker, James A. & Hoffman Michol F. & Meyerhoff, Miriam. 2022. What's in a Lect? Coherence in Phonetic and Grammatical Variation. In Beaman, Karen & Guy, Gregory R. (eds.), *The coherence of linguistic communities: orderly heterogeneity and social meaning*, 71–86. New York: Routledge.
- Weinreich, Uriel. 1954. Is a Structural Dialectology Possible? *WORD* 10(2–3). 388–400. (doi:[10.1080/00437956.1954.11659535](https://doi.org/10.1080/00437956.1954.11659535))
- Weinreich, Uriel & Labov, William & Herzog, Marvin I. 1968. Empirical Foundations for a Theory of Language Change. In Lehmann, Winfred P. & Malkiel, Yakov (eds.), *Directions for Historical Linguistics. A Symposium*, 97–195. Austin: University of Texas Press.
- Wiese, Richard. 2011. The Representation of Rhotics. In Oostendorp, Marc & Ewen, Colin J. & Hume, Elizabeth & Rice, Keren (eds.), *The Blackwell Companion to Phonology*, 1st edn, 1–19. Wiley. (doi:[10.1002/9781444335262.wbctp0030](https://doi.org/10.1002/9781444335262.wbctp0030))
- Wiesinger, Peter. 1990. The Central and Southern Bavarian dialects in Bavaria and Austria. In Russ, Charles (ed.), *The dialect of modern German: a linguistic survey*, 438–519. Stanford: Stanford University Press.

Contacts

alessandro.vietti@unibz.it

massimosimone.cerruti@unito.it