



HAL
open science

Syntactic microvariation, hypothesis testing, and the role of counterexamples

Diego Pescarini

► **To cite this version:**

Diego Pescarini. Syntactic microvariation, hypothesis testing, and the role of counterexamples. 16th Cambridge Italian Dialect Syntax-Morphology Meeting (CIDSM 16), Sep 2022, Napoli, Italy. hal-03906755

HAL Id: hal-03906755

<https://hal.science/hal-03906755v1>

Submitted on 19 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Syntactic microvariation, hypothesis testing, and the role of counterexamples

Diego Pescarini (CNRS, Nice)

1. Introduction

State of the art:

	(Micro)comparative Syntax		(Micro)parametric Syntax	
GOAL	Fine-tuning syntactic analyses		Modelling syntactic variation	
e.g.	Kayne; Poletto	Manzini & Savoia	Ledgeway	Guardiano & Longobardi
SAMPLING	Narrow (theory-driven)	Extensive (not necessarily statistically significant)		Extensive (statistically significant)
	Non-exhaustive	Exhaustive (geographically and/or typologically)		Non-exhaustive; across families/groups
MODELLING	Lexicalization/ Externalization of Functional elements		Yes/no parametric choices	
			Bottom up	Top down

On parameters

- 80s: the Faculty of Language comprises a) a series of invariable Principles and b) a built-in switchboard of Parameters, allowing languages to vary according to a finite and predetermined number of choices.
- Alternative approach (dubbed *Borer-Chomsky Conjecture* by Baker 2008): variation results from a) the extreme simplicity of syntactic principles, which do not impede the emergence of variation, and b) from the properties (technically: *features*) of functional elements. “[a]ssociating parameter values with lexical entries reduces them to the one part of a language which clearly must be learned anyway: the lexicon” (Borer 1984: 29), whereas basic syntactic principles (e.g. *merge*) remain inert to change and variation (Longobardi 2001).

Chomsky 2005:6: “three factors that enter into the growth of language in the individual:

- 1) Genetic endowment, apparently nearly uniform for the species [...]
- 2) Experience, *which leads to variation*, within a fairly narrow range, as in the case of other subsystems of the human capacity and the organism generally.
- 3) Principles not specific to the faculty of language.” (emphasis mine)

Typical RQs:

- Do Parameters exist?
- What is their format/nature?
- How are they organized?

(Macro) Parameter ??

Parameter		Parameter		
Variables	Variables	Variables	Variables	Variables

Problem: We don't see Parameters (I-language). We see language-specific instantiations of Parameters (E-language).

Syntactic variable: A co-occurs/precedes/agrees with B (in the domain D)

Hypothesis: statistically significant associations between syntactic variables are key to understanding parameters.

My method:

- i) I work on descriptive variables/simple phenomena.
- ii) I focus on associations between variables that *a priori* are not interlocked.
- iii) I compare variables pairwise).
- iv) No dedicated survey ("How to make your best with incomplete data" Garzonio & Poletto 2018).
- v) Sample: approx. 360 northern Italian dialects contained in the [ASIt](#) database (Syntactic Atlas of Italy) and published in Manzini & Savoia 2005.

2. A case study: SCIs in NIDs (Pescarini in press a/b; cf. CIDSMS 2019)

Cross-dialectal distribution of subject clitics in a sample of northern Italo-Romance dialects, focusing on syntactic variables that have been examined in previous qualitative analyses (Brandi & Cordin 1981/1989; Rizzi 1986; Poletto 2000; Manzini & Savoia 2005; Roberts 2014):

Some properties of subject clitics that are subject to variation:

- 1) Doubling of an operator-like subject;
- 2) Expletives, i.e. whether weather verbs occur with a subject clitic or not;
- 3) Person-driven gaps;
- 4) Invariable SCIs across persons;
- 5) Order w.r.t. clitic negation;
- 6) Inversion;
- 7) Relationship with "rich" Infl.

Variable 1 (Doubling). Most northern Italian dialects allow the co-occurrence of subject clitics and phrasal subjects. However, not all subjects can be doubled by clitic formatives. Poletto (2000) shows that, whereas pronominal and DP subjects are doubled quite systematically (see (1a)), operator-like subjects such as *wh* elements and bare quantifiers are doubled less frequently, see (1b).

- (1) a. 'marjo **el** 'riva do'maŋ (Ver.)
Mario 3sg.m.nom= arrive.3sg Tomorrow
'Mario (he) will arrive tomorrow.'

- b. ʃi 'riv(a) *elo
 Who arrive.3sg =3sg.m.nom
 'Who will arrive?'

Variable 2 (Expletives). NIDs differ with respect to the occurrence of expletive subject clitics in impersonal constructions, i.e. sentences containing non-thematic verbs or featuring a postverbal/clausal subject.

- (2) a. L 'rua (Monno)
 3sg.m.nom= arrive.3sg
 'He/she is coming.'
 b. el 'plof
 3sg.m.nom= rain.3sg
 'It rains.'
- (3) a. el 'riva (Ver.)
 3sg.m.nom= arrive.3sg
 'He/she is coming.'
 b. *el 'piove
 3sg.m.nom= rain.3sg
 'It rains'

Testing Expletives (with weather verbs) / Doubling (of the subject who). $p = .002565$

Table 1. Number of dialects exhibiting expletives (with weather verbs) and doubling of the wh element *who*. Dataset: Manzini & Savoia 2005 + selected datapoints from the ASIIt database.

	No doubling	Doubling
<i>No expletive</i>	25	10
<i>Expletive</i>	67	88

Variable 3 (Gaps). SCIs are often missing, cf. Table 2.

- (4) a. _ magno el pomo (Ver.)
 b. a mangio er pomo (La Spezia)
 'I am eating an apple.'

Testing Gaps (1sg person) / Expletives. ($p = .000141$)

Table 2. Number of dialects exhibiting expletives (with weather verbs) and gaps (at the 1st person singular) in Manzini & Savoia's 2005 sample.

	No Gaps	Gaps
<i>No expletive</i>	2	20
<i>Expletive</i>	74	90

The incidence of gaps is higher in dialects without expletives.

Testing Doubling / gaps. (not significant at $p < .05$.)

Table 3. Number of dialects exhibiting gaps (at the 1st person singular) and doubling of the wh element *who*. Dataset: Manzini & Savoia 2005 + selected datapoints from the ASIIt database.

	No Gaps	Gaps
No doubling	75	44
Doubling	52	40

Variable 4 (Syncretism). SCIs are often *syncretic*, i.e. different bundles of person and number features are externalized by the same exponent, e.g. *a, i* in Table 4.

Table 4. Examples of paradigms of subject clitics in Italo-Romance dialects

	Verona	Fornero	Piverone	Olivone
1sg		i	i	a
2sg	te	ti	at	tu
3sg (m/f)	(e)l/la	al/la	al/la	u/ra
1pl		i	i	a
2pl		i	i	a
3pl (m/f)	i/le	i	a	i

If the 1sg SCI is present, it is likely to be an invariable clitic, a *dummy*:

Table 5. Number of dialects exhibiting syncretism involving the 1st person singular SCI. Dataset: Manzini & Savoia 2005.

	No gap
No syncretism	8
Syncretism	89

Since dummies correlate with expletives, can the former be analyzed as a particular kind of expletives?

Variable 5 (NegPos). SCIs can either precede or follow preverbal negation (in the NIDs that still display a preverbal negator)

- (5) a. **al** (**na**) 'dɔrma 'mia (Agazzano, PC)
 he= not= sleep.3SG neg
 b. **no** **l** 'dɔrme 'mia (Ver.)
 not= he= sleep.3SG neg

Testing Gaps (1sg SCIs) vs NegPosition. ($p < 0.00001$).

Table 6. Number of dialects where the 1sg is missing and other SCIs precede/follow the preverbal negation. Dataset: Manzini & Savoia 2005.

	Neg > SCIs	SCIs > Neg
Gap	11	16
No Gap	0	49

Testing NegPos / Expletives. ($p = 0.0475$).

Table 7. Number of dialects in which SCIs occur with weather verbs and SCIs precede/follow preverbal negation. Dataset: Manzini & Savoia 2005.

	Neg > SCIs	SCIs > Neg
<i>No Expletive</i>	4	7
<i>Expletive</i>	7	58

Testing NegPos / Doubling. (Not significant at $p < .05$)

Table 8. Number of dialects in which SCIs double *who* and precede/follow preverbal negation. Dataset: Manzini & Savoia 2005.

	Neg > SCIs	SCIs > Neg
<i>No Doubling</i>	1	11
<i>Doubling</i>	7	21

Testing Inversion / Negation (Not significant at $p < .05$)

Table 9. Number of dialects exhibiting inversion (in positive interrogatives) and/or SCIs preceding preverbal negation. Dataset: Manzini & Savoia 2005.

	No Inversion	Inversion
<i>Neg > SCIs</i>	4	7
<i>SCIs > Neg</i>	17	49

[From now on, I will report only significant associations]

An aside on an orthogonal Variable: true/fake imperatives

- (6) a. mangia!
b. non mangiare!

Testing Fake imperatives / presence/absence of a preverbal (clitic) negation, cf. Zanuttini's 1997 ($p < 0.00001$).

Table 10. Number of dialects exhibiting fake/true imperatives with/without preverbal negation. Dataset: Manzini & Savoia 2005.

	True Imp	Fake Imp
<i>No PreV Neg</i>	77	20
<i>PreV Neg</i>	7	83

Testing fake/true imperatives / NegPos. (Not significant at $p < .05$).

Table 11. Number of dialects exhibiting fake/true imperatives and in which SCLs precede/follow negation. Dataset: Manzini & Savoia 2005.

	True Imp	Fake Imp
<i>Neg > SCLs</i>	2	6
<i>SCLs > Neg</i>	18	47

Variable 6 (Inversion). In main interrogatives, subject clitics are placed postverbally (i.e. in *enclisis*, see (4a)), whereas object clitics remain preverbal (in *proclisis*, see (4b)):

- (7) a. 'riv- ela do'maŋ (Veronese)
 arrive.3sg =3sg.f.nom= tomorrow
 'Does she/it arrive tomorrow?'
 b. te riv- ela do'maŋ
 2sg.dat= arrive3sg =3sg.f.nom Tomorrow
 'Will you receive it tomorrow?' (lit. 'Will it/she arrive to you
 tomorrow?')

Testing inversion / expletives (Not significant at $p < .05$)

Table 12. Number of dialects exhibiting expletive SCLs and/or inversion. Dataset: Manzini & Savoia 2005.

	No Inversion	Inversion
<i>No Expletives</i>	3	14
<i>Expletives</i>	52	84

Variable 7 (Rich Inflection).

The relationship between the richness of verbal inflection, the syntax of subject clitics and the licensing of null subjects has been debated since the 19th century (Meyer-Lübke 1895; within the generative framework, see Perlmutter 1971; Taraldsen 1980; on northern Italo-Romance, see Roberts 2014).

Rich Infl:

- number of distinctive endings (including zeroes)
- present indicative of regular verbs.
- excluding 1/2pl verbs, where Infl triggers stem allomorphy/suppletion
- as for the remaining 4 verb forms (1-3sg, 3pl),
 - o 0-1 contrasts → poor
 - o 2-3 contrasts → rich

Testing RichInfl / Expletives. (Almost significant; $p = .0502$)

Table 13. Number of dialects exhibiting expletive SCIs in which SCIs precede/follow preverbal negation Dataset: Manzini & Savoia 2005.

	Poor Infl	Rich Infl
No Expletive	2	19
Expletive	50	118

Testing RichInfl / Expletives & Gaps. ($p < 0.00001$)

Table 14. Presence of expletives and dummies in dialects with poor and rich verbal inflection (dataset: Manzini & Savoia 2005).

	Poor infl	Rich infl
both expl and dummy	52	38
either expl or dummy	13	63
neither expl nor dummy	1	18

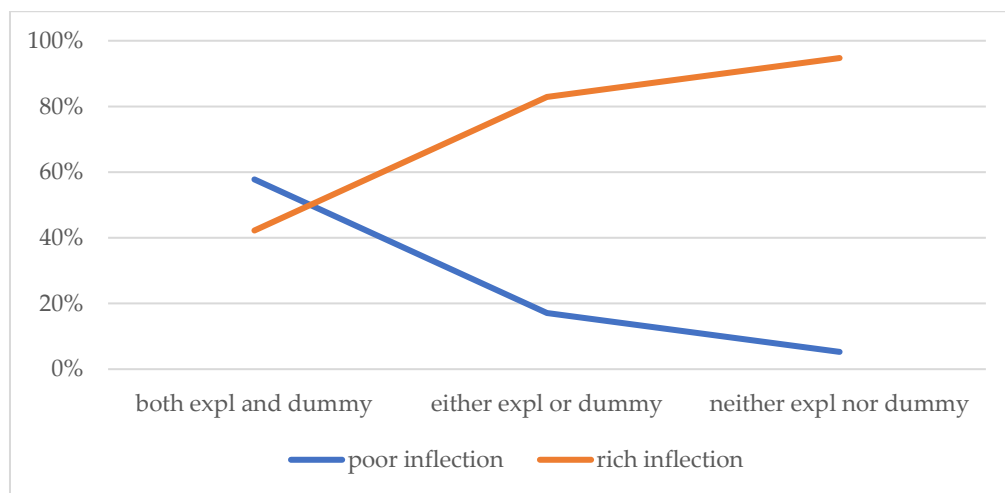


Fig. 1

3. Discussion

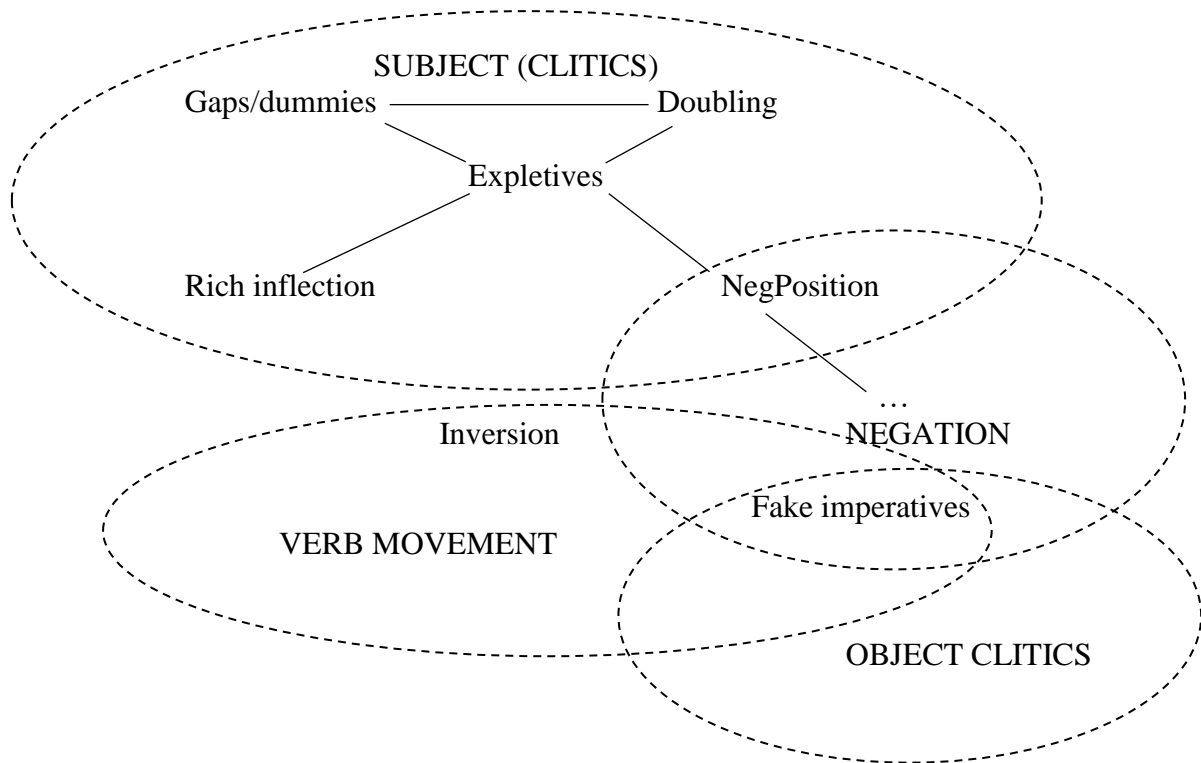
Summary:

	Doubling	Expletives	Gaps	NegPos
Doubling	-	yes	-	-
Gaps/dummies	*	yes	-	-
Negation	*	yes	yes	-
Rich Inflection	*	yes		*
Inversion	*	*	*	*

Remarks:

- i. Having a sample of 180-200 dialects, we can test associations between categorical variables;

- ii. Some of these associations proved to be statistically significant;
- iii. Associations are not *transitive*: if <a b> & <b c> do not imply <a c>. Therefore, no hierarchy.
- iv. The explanandum of Parametric Theory is a network of variables:



- v. If associations were transitive, dialects would tend to fall into two (proto)types (e.g, {a1, b1, c1} vs {a0, b0, c0}). In fact, no bimodal distribution.
- vi. No prediction can be made regarding the overall structure of E-languages. E-languages are affected by other 2nd/3rd factors.
- vii. As for SCIs, we have a constellation of phenomena revolving around a primary variable, namely “Expletive”.
- viii. Other phenomena (“Gaps”, “Doubling”, “Rich inflection”) are arguably related to this primary variable.
- ix. The associated Parameter seems to be EPP-related.
- x. Ways to satisfy the EPP:
 - DP
 - X → (Stylistic Fronting)
 - D → (Subject clitics)
 - (Rich) Infl