RoLinC (Romance Linguistics Circle)
Universities of Cambridge and Newcastle
18.01. 22

Microvariation and statistical hypothesis testing: negation and negative concord in northern Italo-Romance

Diego Pescarini

CNRS – UMR 7320: Bases Corpus Langage

Context and general aim

- Problem: many claims/conjectures have not been tested against big datasets or following a standard methodology to obtain replicable results.
- Solution: now we can combine two streams of research:
  - a considerable effort in digitizing, reviving, and diffusing collections of linguistic data (e.g. atlases such as AIS/ALF or dialectological databases such as ASIt/Thesoc), which are now freely available thanks to the recent **open data revolution**;
  - a **paradigm shift** in syntactic research, where quantitative methods and experimental methodologies are becoming increasingly popular.
- By (re)approaching big collections of data through sound statistical methodologies we can finally establish empirical and theoretical firm points.

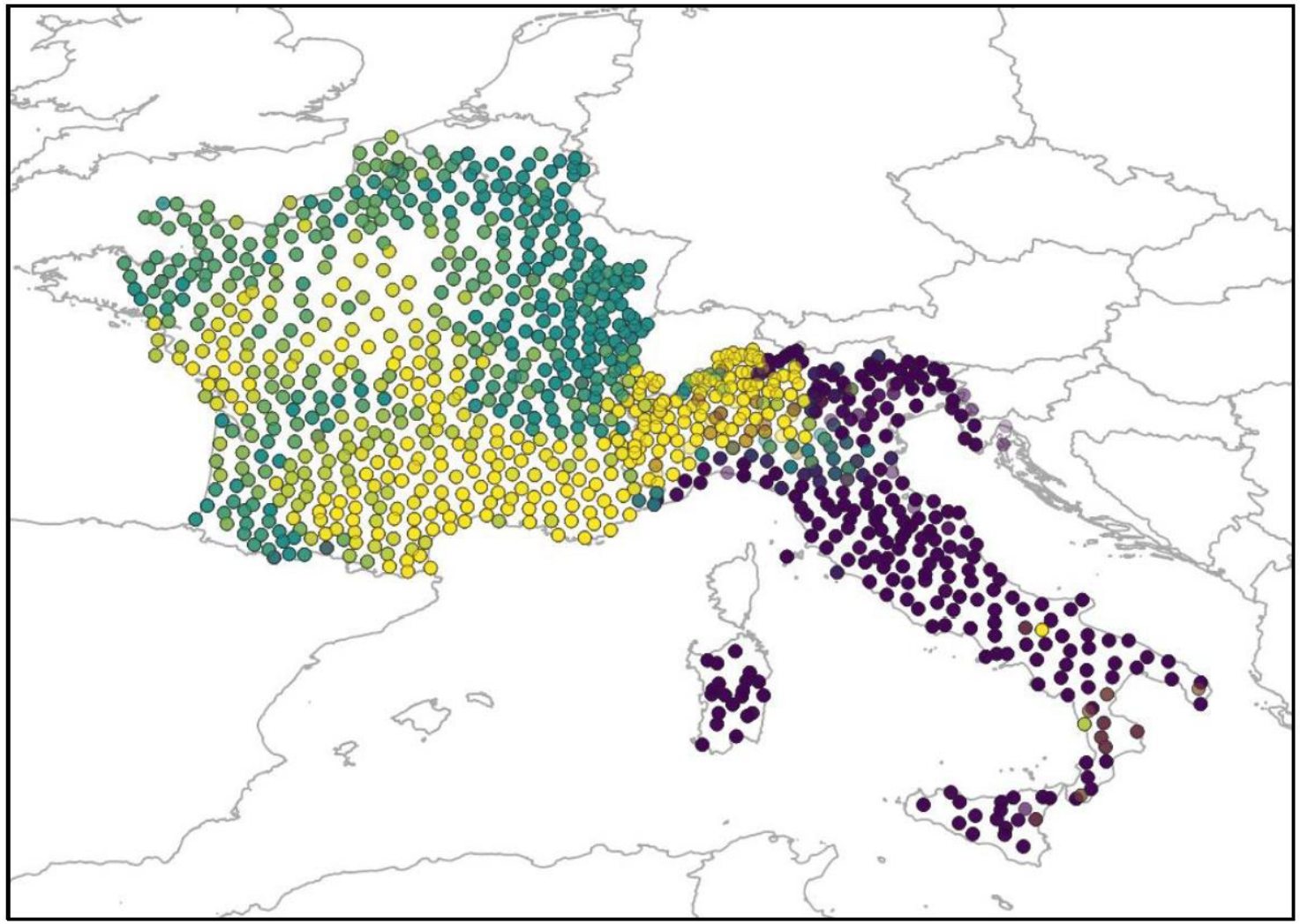
Specific aims of the talk

- Aim: testing hypotheses, conjectures, generalisations, implicational statements, etc. regarding the syntax of negation in northern Italian dialects.
- Datasets: Manzini & Savoia 2005; AIS
- Four topics/RQs:
  1. Negation and suppletive imperatives
  2. Emergence of postverbal negation markers
  3. Loss of preverbal negative markers
  4. Negative concord with postverbal negation
  5. Negative ‘tripling’ in languages with discontinuous negation

Terminology

- N1 usually derives from Latin NON (Fr. *ne*, It. *non*) and co-occurs with Negative Polarity Items (NPI), giving rise to negative concord configurations; it is usually placed preverbally and, in finite clauses, it is proclitic to the finite verb (only other clitics can occur between N1 and the verb);
- N2 derives from various kinds of elements such as nouns denoting a minimal quantity (dubbed *minimizers*, e.g. *pas* 'step', *point*, *mie/mica/brisa* 'crumb', etc.) or, to a lesser extent, negative quantifiers and polarity particles; it sometimes co-occurs with NPIs (cf. Fr. *je ne mange (*pas) rien* 'I eat nothing') and it normally occurs postverbally in finite clauses.

N1, N2 in central Romance (AIS and ALF datapoints)



1. Zanuttini's 1997 generalisation

Verbal forms that are **unique to the paradigm of the imperative**, cannot co-occur with preverbal negative markers, e.g.

	POSITIVE	NEGATIVE	
2sg	Parla!	Non parlare!	*Suppl*.
2pl	Parlate!	Non parlate!	

Additionally, negative imperatives differ from positive ones w.r.t. clitic placement:

2sg	Parlami!	Non (mi) parlare (mi)!
2pl	Parlatemi!	Non (mi) parlate (mi)!

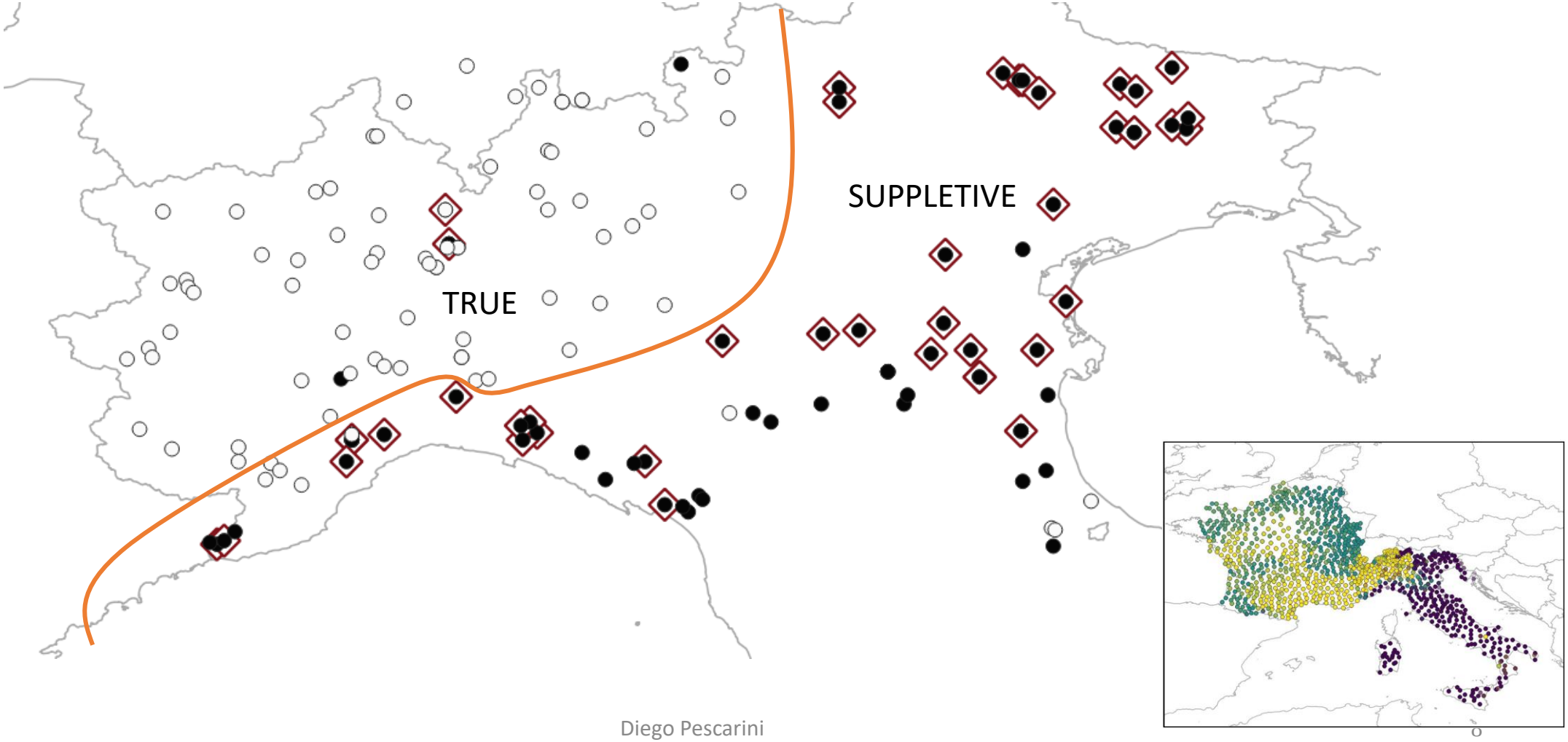
Counterexamples to Zanuttini's generalization

Manzini & Savoia (2005 III: 467) found several couterexamples:

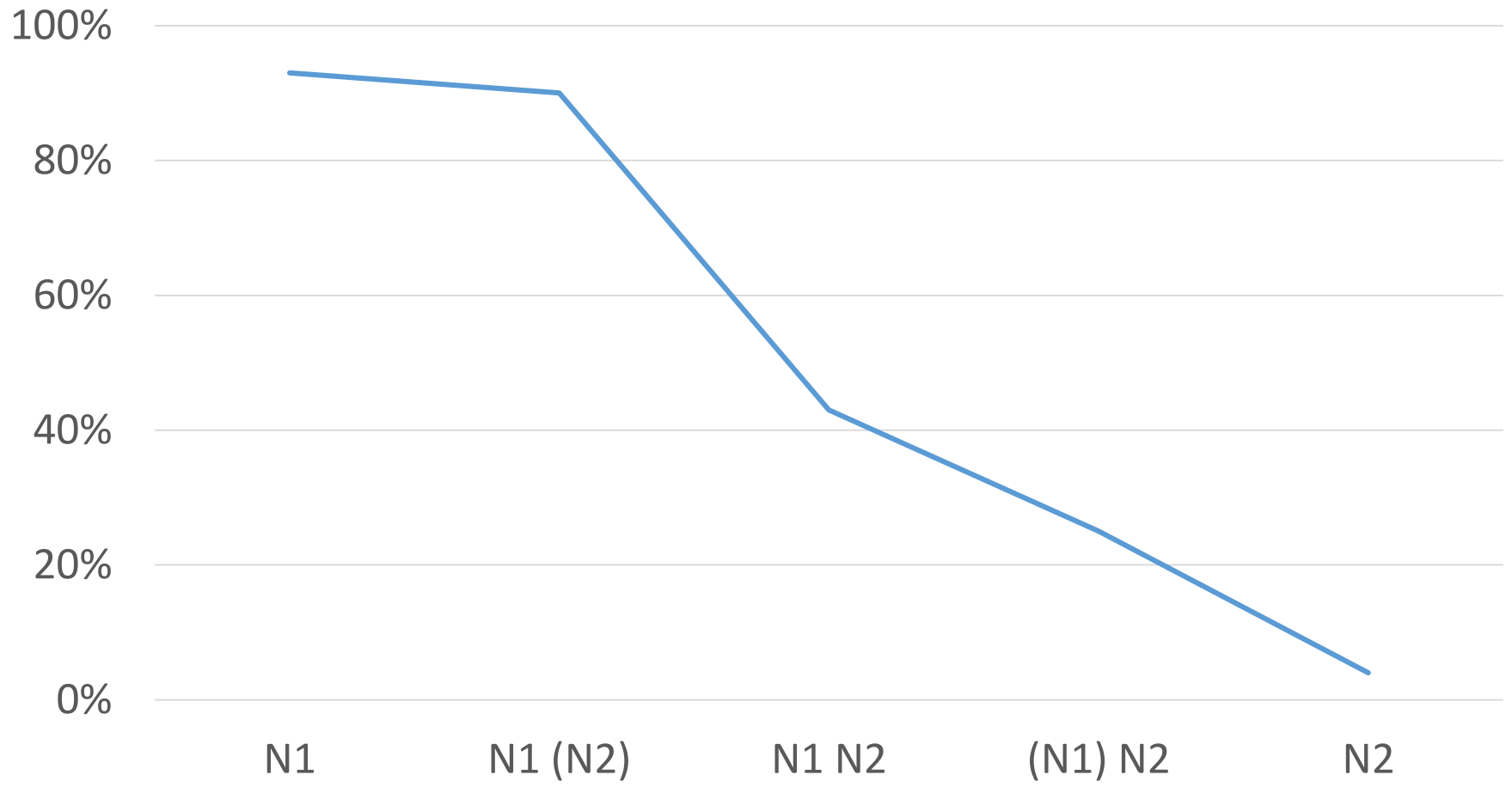
- true imperarives with N1 (cfr. *Mercato Saraceno*, *S.Mauro*, *Rontagnano*; Ampezzano e a.Mil. Vai 1995)
- true imperatives with N1 and N2 (cfr. *Modena*, *Finale Emilia*)
- suppletive imperative with N2 (Zanuttini 1997: 112 per l'Occitano; see also Benincà & Poletto 2007)

True vs suppletive imperatives

(diamonds = dialects with periphrastic negative imperatives, e.g. *no sta* + inf)



% of suppletive imperatives / negation type



1. Interim conclusion

- Zanuttini is right.

Open Qs:

- Can the association between suppletive imperatives and the syntax of sentential negation be accidental?
- What about plural imperatives?
- What about clitic placement?
- Is the type/etymolon of N2 relevant?

2. The emergence of N2

2a) the syntax of N2s is initially constrained by pragmatic conditions (Schwenter 2005; Larrivée 2008). Given its pragmatic value, is **N2 disfavoured in sentences containing presupposition triggers** (as proposed by Cinque 1991 [1976] for regional Italian)**?**

2b) Most N2s derive from NPIs (minimizers or quantifiers). Is **N2 favoured in non-veridical contexts** such as modals, *if-clauses*, questions, etc. (Giannakidou 1998)?

The AIS/ALF metadata

https://zenodo.org/record/5820466#.YeZqRPjjJc8

January 5, 2022

Dataset Open Access

Negation and negative concord in AIS and ALF examples

Pescarini, Diego

The file contains metadata on 22500 negative sentences from 1046 Italo-, Gallo-, and Rhaeto-Romance dialects. It is based on primary data from the *Atlas linguistique de la France* (ALF; Gilliéron and Edmont 1902-1910) and the *Sprach- und Sachatlas Italiens und der Südschweiz* (AIS; Jaberg and Jud 1928-1940).

This dataset is a research output of the project **Coding Syntactic Microvariation** (CoSMic), funded by the IDEX programme of the Université Côte d'Azur (CSI Recherche 2021).

languages MDPI

Article

A quantitative approach to microvariation: negative marking in central Romance

Diego Pescarini [1]

[1] CNRS, Université Côte d'Azur, BCL

Abstract: This work presents a multivariate analysis of the distribution of pre- and post-verbal negation in a corpus of data gathered from two linguistic atlases: the Linguistic Atlas of France (ALF) and the Italo-Swiss Atlas (AIS). Metadata concerning the distribution of pre- and post-verbal negation across dialects and across syntactic contexts is analyzed in order to ascertain whether (and to what extent) syntactic factors such as force, aspect, mood and negative concord affect the distribution of negators.

Keywords: syntax, negation, negative concord, microvariation, Romance

Data

Table 1. Overview of the dataset

Source	Negative sentences	Datapoints	Tokens
ALF	22	639	13185
AIS	24	407	9315
Total	46	1046	22500

Table 2. Counts of tokens per negation system (excluding clauses containing NPIs).

Source	N1	N1+N2	N2
ALF	64	4404	3416
AIS: 'northwestern'	559	190	1100
AIS: others	2050	5	18
Total	2673	4599	4534

Testing factivity

(hyp: presupposition triggers disfavour N2 Cinque 1991 [1976])

Comparing the incidence of N2 in presuppositional environments, e.g. the complement of factive verbs and non-presuppositional clauses in the AIS 'northwestern' dataset:

	N2	**no N2**
Baseline sentences *Non capisco* *Non voleva rimanere*	237	114
Factive verbs *mi meraviglio che non lo troviate* *mi rincresceva che non la trovassimo*	201	116

Interim conclusion: factivity has no significant effect ($p > .05$).

Testing veridicality /1
(hyp.: N2 is favoured in non-veridical contexts)

Comparing the incidence of N2 in non veridical environments, e.g. questions, *if* clauses and indicative clauses (baseline) in the AIS 'northwestern' dataset:

	N2	**no N2**
Indicative *Non capisco*	132	61
Non-veridical *Non vedi che...?* *Se non mangiamo...*	222	120

Interim conclusion: veridicality has no significant effect ($p > .05$).

Testing veridicality /2
(hyp.: N2 is favoured in non-veridical contexts)

- I tested whether factors *Region* and *Clause* are significant predictors of the distribution of N2

Table 9. Frequency/Probability of N2 w.r.t. Clause in a model with factors Region and Clause (AIS dataset, only 'northwestern' datapoints, excluding tokens containing NPIs and collinear values)

Clause	Tokens	Frequency of N2	Rbrul probability
Conditional	174	.75	.65
Future	187	.74	.63
Modal	161	.71	.53
Indicative	342	.68	.49
Question	168	.65	.42
If clause	174	.65	.41

- both factors are statistically significant (Region p=2.3e-249; Clause p=0.00053)…
- but no clear effect of veridicality on N2.

Testing veridicality /3

(hyp.: *certain* N2s are favoured in non-veridical contexts)

- N2 < minimizer: *mica*, *brisa*, *bec*, *point*, *pas*
- N2 < negative elements: *n(i)ent*, *no*

	minimizers	**negative elements**	***Marginal Row Totals***
nonveridical	463 (401.08) [9.56]	88 (149.92) [25.57]	551
indicative	262 (323.92) [11.84]	183 (121.08) [31.67]	445
Marginal Column Totals	725	271	996 (Grand Total)

The chi-square statistic is 78.6377. The *p*-value is < 0.00001.
Significant at *p* < .05.

2. Interim conclusions

- No effect of factivity on N2;
- Mild effect of non-veridicality on N2s deriving from minimizers.

- Linear regression found that Region is always a significant factor, as expected in a dataset containing data from tightly related languages.
- Region, however, is seldom sufficient to account for the distribution of syntactic variables. **Mixed models including syntactic factors usually fit the data best.**

3. The loss of N1

(hyp.: since most N2s derive from NPIs, **N1** is more likely to be omitted in nonveridical contexts)

In non-veridical contexts, N2 can be licensed by the non-veridical operator, e.g. INT:

- DECL: *T'**an** ved mia che…?* 'You don't see that…'
- INT: *%(**an**) vedat mia che…?* 'Don't you see that…?'

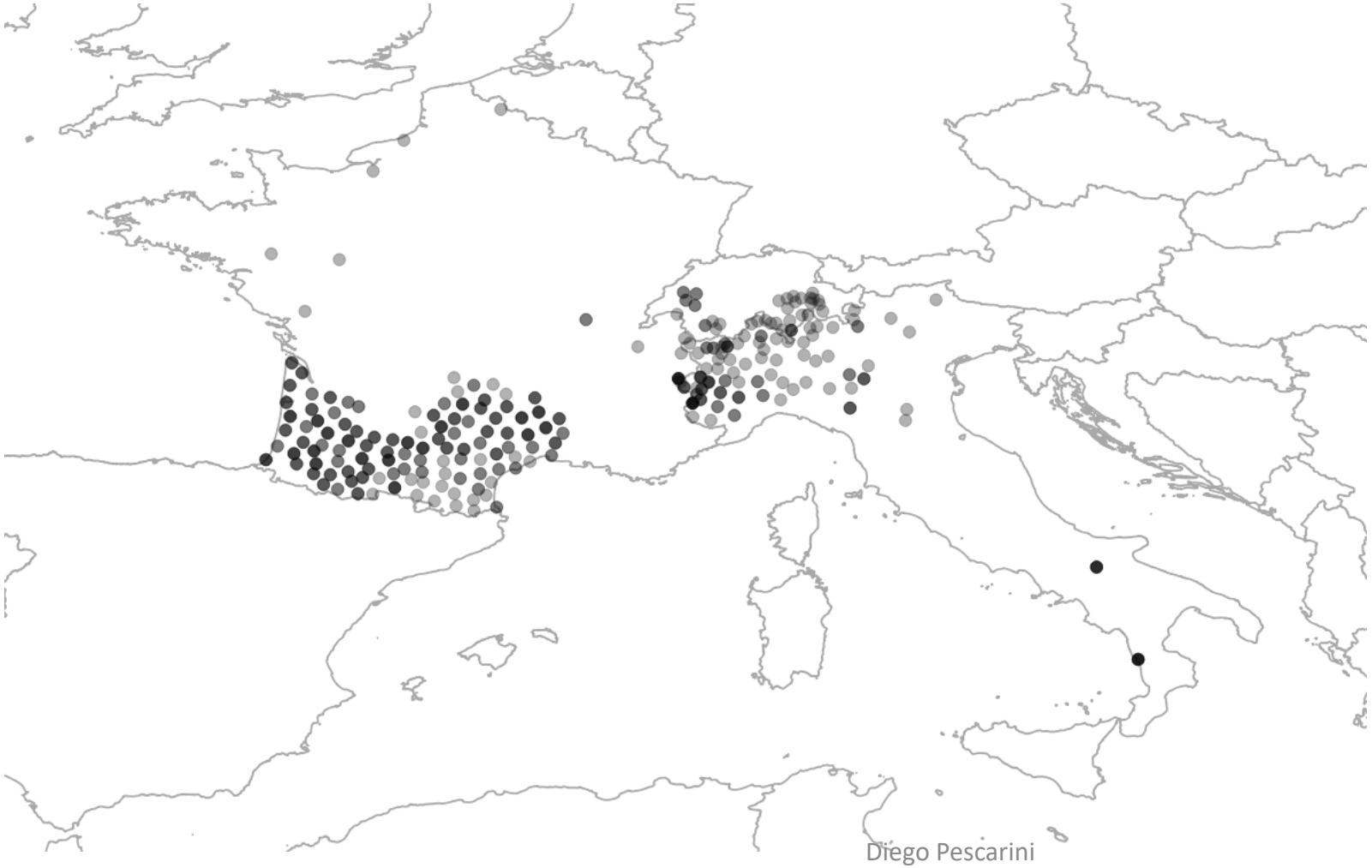
Clause	Tokens	Frequency of N1	Rbrul probability
Indicative	193	.44	.54
Question	168	.42	.50
Future	187	.42	.50
Conditional	174	.41	.49
Modal	161	.37	.42
Imperative	641	.35	.41

3. Interim conclusions

- N1 is more likely to be dropped in non-veridical contexts, where non-veridical operators (e.g. conditionals, modals) license N2 even if N1 is missing.

- As above, linear regression showed that Region and Clause type are both significant factors.

4. Negative concord with N2



NC with N2

Manzini & Savoia 2005 III: 258ff:

- Lexical variation: N2 + *mai* ‘never’ is rarer than N2 + *più* ‘anymore’;
- N2 + NegQ is favoured in compound tenses and in languages in whcih N2 precedes the past participle;
- NPIs embedded in a PP co-occur quite often with N2.

- Tentative hierarchies:
  - never > anymore
  - V N2 NPI > N2 Part NPI
  - DP > PP

NC with N2: AIS vs ALF

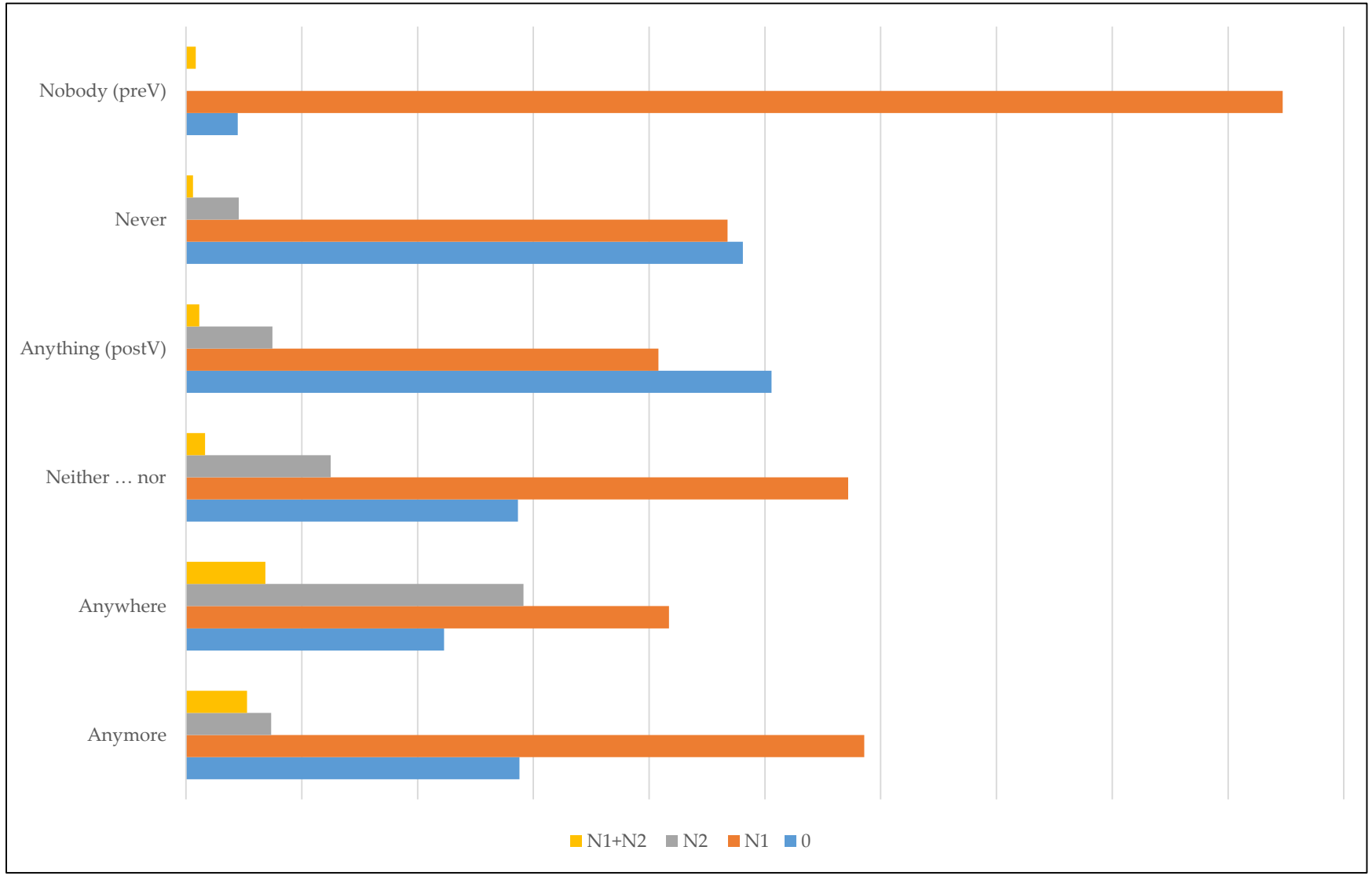
Table 10. Frequency/Probability of N2 w.r.t. NPIs in a model containing factors Region and NPI Type (AIS dataset, only 'northwestern' datapoints, excluding 'yet' and collinear values)

Clause	Tokens	Frequency of N2	Rbrul probability
Anywhere (PP)	175	.36	.79
Neither… nor	160	.36	.78
Anything (obj.)	172	.09	.27
Never	836	.05	.17

Table 18. Incidence of N2 w.r.t. NPI Type (ALF dataset)

Clause	Tokens	Frequency of N2
Yet	636	.99
Anymore	1846	.14
Neither … nor	632	.09
Nobody (subj.)	359	.01
Guère	343	.00

NC with N2: AIS + ALF



Comparison with Dagnac & Burnett 2016

	Montréal	Picard	AIS	ALF
Anywhere	.83	.90	.36 63/175	
Neither...nor			.36 57/160	.09 52/533
Anybody	.59	.33		.00* 3/359
Anything	.15	.13	.08 15/174	
Any	.11	-		
Never	.1	.10	.05 43/837	
Anymore	*	.01	.17 31/175	.06 116/1849

* All preverbal S.

4. Interim conclusions

- PPs meaning ‘anywhere’ and, to a lesser extent, *neither…nor* coordinations are the NPIs with which N2 co-occurs more frequently.
- Preverbal NPIs in *strong NC* languages (Gallo-Romance) favor N1 and strongly disfavor N2 (no data on Italo-Romance).
- Lexical variation w.r.t. to adverbs such as *anymore* and *never* is confirmed, although both can co-occur with N2.

Open questions:
- Asymmetries between adjuncts and arguments
- Asymmetries between pre and post-participial elements
- Strong *vs* weak NC
- Correlations with typology of N2

5. «Tripling»?

(hyp. N2+NPI is allowed iff N1 is missing)

Table 20. Incidence of N2-doubling w.r.t. NPI Types in clauses containing/not containing N1 (AIS 'western' and ALF dataset)

	Incidence of N2 in clauses containing a NPI	**Incidence of N2 in clauses containing a NPI and N1**
Anywhere	.57	.14
Neither … nor	.30	.03
Anymore	.23	.07
Anything	.12	.03
Never	.09	.01
guère/anybody	.00	.00

5. Interim conclusion

- Tripling is allowed

Open question:

- Is tripling allowed iff a language allows the co-occurrence of NPI and N2?

Conclusions

5) Tripling is allowed, but it is rarer than N2+NPI.

4) NPIs are not created equal:

- 4a) PPs meaning ‘anywhere’ and, to a lesser extent, *neither…nor* are the NPIs with which N2 co-occurs more frequently.
- 4b) preverbal NPIs in *strong NC* languages (Gallo-Romance) favor N1 and disfavor N2 (no data on Italo-Romance).
- 4c) “Lexical” variation w.r.t. to adverbs such as *anymore* and *never* is confirmed.

3/2b) the role of veridicality is confirmed:

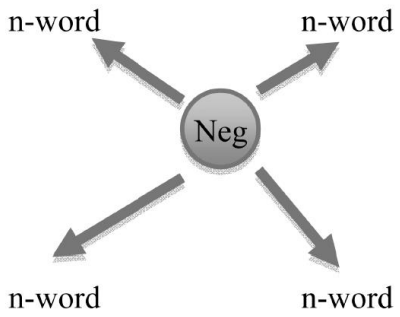
- 3) N1 is more likely to be dropped in non-veridical contexts.
- 2b) N2s deriving from minimizers are more frequent in non-veridical contexts.

2a) No effect of factivity on N2.

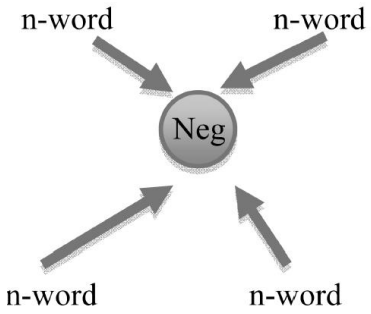
1) Zanuttini’s generalisation on suppletive imperatives is confirmed.

The debate

- Jespersen's classic approach: sentential negation is king.
- «Neo-classic» (Déprez's 2011 term) approaches, e.g. Zanuttini 1997 or Zeijlstra 2008
- Microparametric approach, which «takes the negative dependent expressions as key factors of variation in negative concord systems» (Déprez 2011: 222)



Inside-out perspectives



Outside-in perspectives

Thank you

diego.pescarini@gmail.com