A corpus-based investigation of lexical uniformity in the standardization of Italian

Stefano De Pascale, Stefania Marzo, Dirk Speelman

RU Quantitative Lexicology and Variational Linguistics

Overview

1. Standardization history "from below"

2. Standardization in the Kingdom of Italy (1861)

3. Data & variables

4. Results

5. Conclusion

Standardization history «from below»

- Haugen's **top-down** model of standardization (1966)
  - Imposition of norm by elite and cultural institutions, government agencies etc.

- Kristiansen & Coupland **bottom-up** model (2011) :
  - In daily life, 'the people' make on-the-ground assessments of the social implications of using different ways of speaking (Auer & Spiekermann, 2011)

  - Sociolinguistic perspective: empirical data (corpora), impact of extralinguistic variables (statistical modeling)

Standardization history «from below»

- General goal: show empirical evidence for the (variable) impact of language policies on standard language change (~ Geeraerts 2003)

- Which general aspect of standardization?

- **Lexical uniformity (≈ reduction of polymorphy, reduction of free formal variants of a word)**

- Fundamental process in standardization (Geeraerts et. al, 1999, Milroy 2001, De Mauro 2014)

Standardization in the Kingdom of Italy

- ## Standard Italian (from 1525 to 1861)
  - Amended literary Florentine of the 14° century
  - Mainly a written language, learned and spoken by an elite
  - More or less cristallized during the next centuries

- ## 1861: Unification of Italy
  - No more than 9,5% of Italians speak Italian (Castellani 1982)
  - Everyday communication happens in one the base dialects

- ## 1868: *Commission for the unification of the language*

Standardization in the Kingdom of Italy

Three ideological stances on language policy in Italy

- «Traditionalists»
    - Mix of Tuscan purism (literary canon) and classicism (Greek and Latin authors)
    - Cultural preoccupation: trasmission of literary legacy

- «Florentinists»
    - (written) norm: Florentine spoken by contemporary elite
    - Political preoccupation: participation to democratic life

Standardization in the Kingdom of Italy

- «Pluricentrists»
  - Common language will develop from interregional contact
  - Scientific preoccupation: «standardology» avant-la-lettre

- Specific research questions:
  1. Which standard language ideology has had the most lasting impact on the standardization of Italian?
  2. Has that ideology been effective across the board (i.e. did the uniformization involve different lexical variables)?
  3. Has that ideology succeeded in unifying Italy: are some regions perhaps behind on others?

Data & variables

- DiaCORIS: diachronic reference corpus of Italian

  (25 million tokens; Onelli, Proietti, Seidenari, & Tamburini, 2006)

- From 1861 to 2001 (divided in 5 historical periods)

- 5 macrogenres:
  - Legal-administrative prose
  - Press (newspapers & periodicals)
  - Narrative prose
  - Essayistic prose
  - Miscellanea (lowbrow literature, translations, etc.)

Data & variables

Case studies: 3 types of popular alternations

1. Dental [ts] vs. palatal [tʃ] affricate alternation
    <Z>/<C> [e.g.: *pronunzia/pronuncia*]

*   Latin -NUNTI- root
    (e.g.: *denunciare* 'report', *annunciare* 'announce')

*   <Z> : literary-formal variant (latinism)
*   <C> : regular Tuscan phonetic outcome (Aski 2001)

Data & variables

Case studies: 3 types of popular alternations

2. Diphthong [wɔ] vs. monophtong [ɔ] alternation
   <UO>/<O> [e.g.: *spagnuolo/spagnolo*]

- Latin -Ŏ- in stressed, open syllables after palatals
  (e.g.: *fagi(u)olo* 'bean', *ai(u)ola* 'flower bed')

- <UO> : literary-formal variant

- <O> : contemporary variant (spoken Florentine)

Data & variables

Case studies: 3 types of popular alternations

3. Geminate vs. single bilabial stop alternation
    <B>/<BB> [e.g.: o*b*iettivo/o*bb*iettivo]

- Before <i> and <r>
  (e.g.: *ub(b)riaco* 'drunk'*, ob(b)iezione* 'objection')

- <B> : literary-formal variant (latinism)
- <BB> : regular Tuscan phonetic outcome

Results

	Variant preferred by «traditionalists»	Variant preferred by «Florentinists»	**Winning variant**
Affricate alternation	<Z>	<C>	**<C>**
Geminate alternation	<B>	<BB>	**<B>**
Diphthong alternation	<UO>	<O>	**<O>**

Results

	Variant preferred by «traditionalists»	Variant preferred by «Florentinists»	Winning variant
Affricate alternation	\<Z\>	\<C\>	\<C\>
Geminate alternation	\<B\>	\<BB\>	\<B\>
Diphthong alternation	\<UO\>	\<O\>	\<O\>

Results

|	Variant preferred by «traditionalists»	Variant preferred by «Florentinists»	Winning variant
Affricate alternation	<Z> 1861-1900: 67%	<C> 1861-1900: 33%	<C> 1968-2001: 90%
Geminate alternation	<B> 1861-1900: 36%	<BB> 1861-1900: 64%	<B> 1968-2001: 88%
Diphthong alternation	<UO> 1861-1900: 86%	<O> 1861-1900: 14%	<O> 1968-2001: 87%

Results: first summary

- Reduction of variation within alternation types
  - Achieved lexical uniformity in Standard Italian
  - Victory for proponents of (radical) unification

- Heterogeneity of sources across alternation types
  - 2x Florentine variant, 1x traditional variant
  - Victory for proponents of (relaxed) pluricentrism

- Winning variants start as minority variants

Affricate alternation (press) [C=0.76]
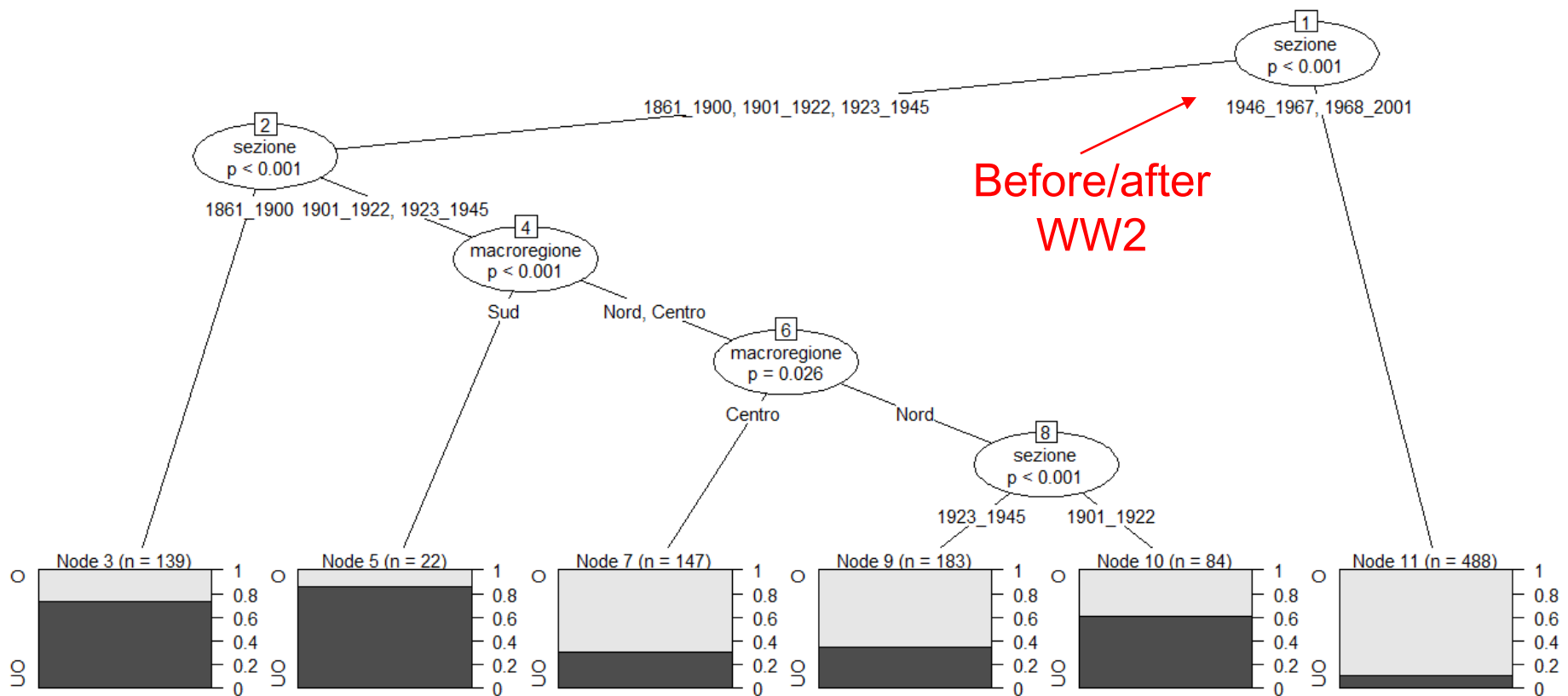

Early regional differentiation

Before/after Fascist era
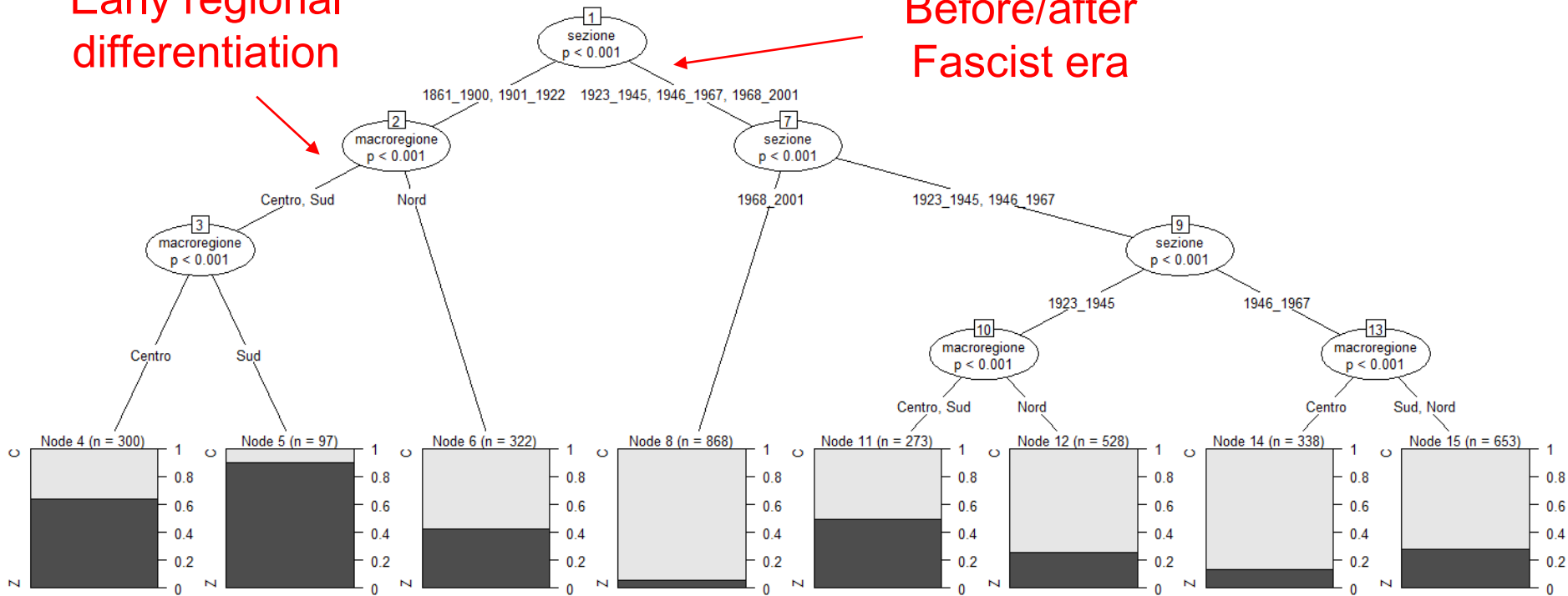
Almost complete disappearence of <Z>

Diphthong alternation (press) [C=0.78]
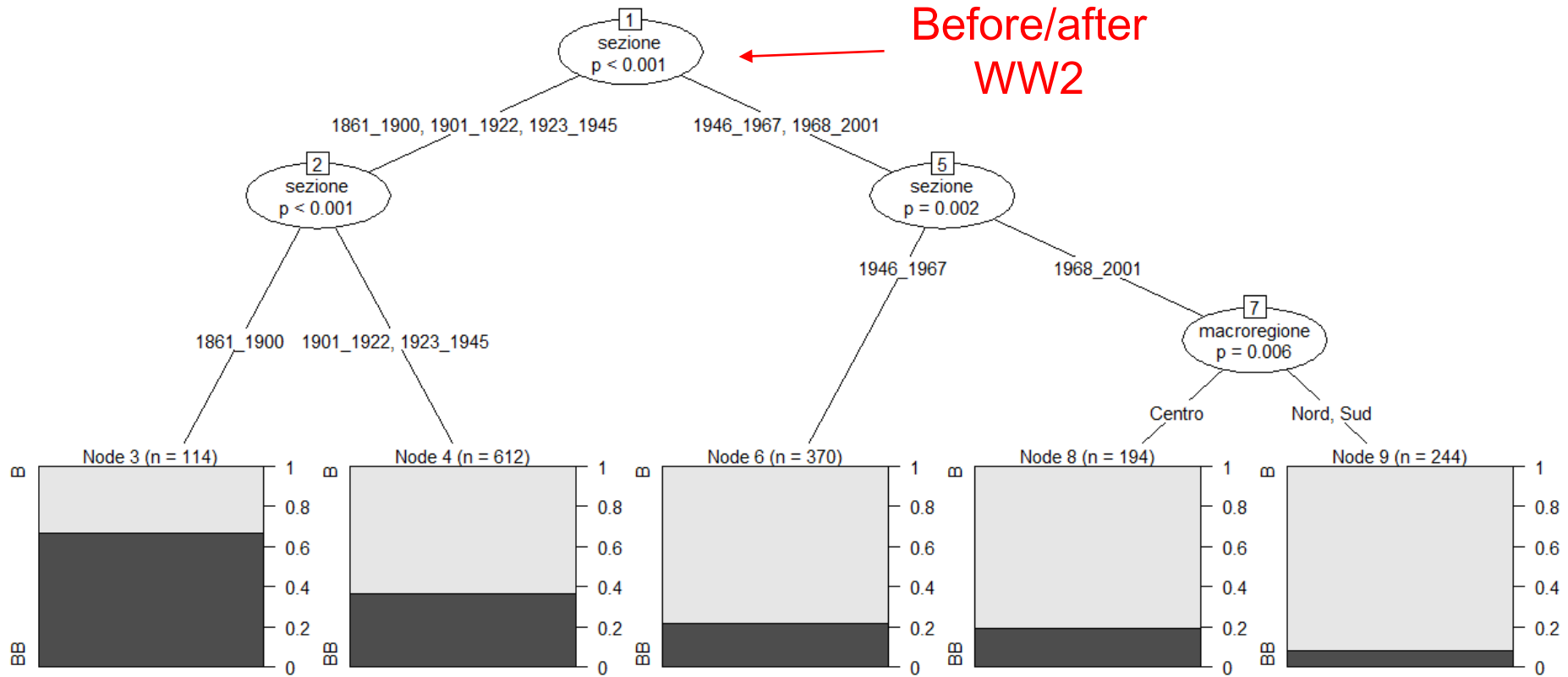

Before/after WW2

Conservative southern habits

Almost complete disappearence of <UO>

Geminate alternation (press) [C=0.74]

Before/after WW2

<BB> only majority variant in first period

Late (weak) regional differentiation

Majority variants per period per region

	North			Centre			South		
1861-1900	C	UO	BB	Z	UO	BB	Z	UO	BB
1901-1922	C	UO	B	Z	O	B	Z	UO	B
1923-1945	C	O	B	Z/C	O	B	Z/C	UO	B
1946-1967	C	O	B	C	O	B	C	O	B
1968-2001	C	O	B	C	O	B	C	O	B

- Red = final winning variant,
- Blue = final losing variant

Conclusion

- Lexical uniformity in Standard Italian only after WWII
  - «miracolo economico italiano»: increased interregional mobility, democratization of higher education etc.

  - Evolution in line with expectations of «pluricentrists»: usage-based standard language change

- In press, uniformity is first achieved in the North, during the fascist era (1923-1945), while the South catches up later

Conclusion

- Composite nature of the uniformization dynamics
  - (2X Florentine-backed variant, 1X tradition-backed variant)
  - Again, as predicted by the «pluricentrists», unity in language does not imply homogeneity of sources but rather «natural selection»

- Future work:
  - Inclusion of extra alternations
  - Further annotation of data (e.g.: provenance of authors)
  - Improve statistical models
  - Better understanding of lexical diffusion (paradigms)

Thank you!

for further information:
stefano.depascale@kuleuven.be
http://wwwling.arts.kuleuven.be/qlvl