insedicesimo
editoria di qualità in piccole quantità

AUTORI VARI

# Il patrimonio linguistico storico della Liguria 2

## ATTUALITÀ E FUTURO

*A cura di Erica Autelli*

Autori vari

# IL PATRIMONIO LINGUISTICO STORICO DELLA LIGURIA 2

## Attualità e futuro

Studi in onore di Fiorenzo Toso

A cura di Erica Autelli

insedicesimo
editoria di qualità in piccole quantità

# Il patrimonio linguistico storico della Liguria
Raccolta di studi

a cura di Erica Autelli

© Proprietà letteraria degli articoli riservata ai rispettivi autori.
© Impaginazione e grafica, *In*Sedicesimo, di Delfino & Enrile Editori snc

Stampa: *In*Sedicesimo
marchio editoriale di Delfino & Enrile Editori snc
via Giovanni Scarpa, 10 r. – 17100 Savona
ordini@insedicseimo.it – 019.811800

1ª edizione: Marzo 2024

Fotografia di copertina fornita da Gianni Giusto per l'Associazione *U Campanin Russu*.
Pubblicazione finanziata dall'Associazione *U Campanin Russu*.

Il presente volume è stato sottoposto a procedura di referaggio esterno in doppio cieco (double blind peer review).

ISBN 978-88-99866-79-2

**Ringraziamenti**
Si ringraziano i seguenti enti per aver contribuito allo svolgimento del convegno:

**Comune di Varazze Assessorato alla Cultura**

**Associazione culturale U Campanin Russu**

**Parco del Beigua**

**Regione Liguria**

**Corellimusica**

**Austrian Science Fund FWF**

**Universität Innsbruck**

**Univeristà degli Studi di Sassari**

**Università degli Studi di Genova**

# Indice

Saluto del Sindaco di Varazze Luigi Pierfederici.....9

Saluto del Consigliere regionale Avv. Alessandro Bozzano .... 11

Presentazione, di Giovanni Ghione.....13

Introduzione, di Erica Autelli.....17

Giorgio Toso.....25
*Gli usi linguistici dei Genovesi di Tunisi: alcuni riferimenti nella documentazione italiana e francese (secc. XVIII-XIX)*

Lorenzo Filipponio, Davide Garassino, Danila Dipino .....37
*Fine del genovese o un nuovo genovese?*

Philippe Boula de Mareüil, Antonio Romano, Marc Evrard, Alexandre François.....51
*Cartografia di innovazioni rispetto al latino attraverso un atlante sonoro dell'Europa*

Werner Forner .....63
*Truin zeneixi – Alcuni processi grammaticali al margine dei dialetti genovesi*

Erica Autelli, Riccardo Imperiale .....99
*Gli sviluppi della lessicografia genovese con GEPHRAS2*

Giuseppe Gandolfo ..... 113
*Il caso piscialandrea. Semiotica, tradizione, identità, contaminazione*

Giannino Balbis ..... 149
*Alta Val Bormida dialettale: l'ultimo decennio*

Alessandro Guasoni ..... 161
*Letteratura in genovese dal Novecento a oggi*

Davide Baglietto ..... 173
*La musa pastorale ligure e la sanfòrgna*

Philippe Boula de Mareüil (Université Paris-Saclay, CNRS, LISN),
Antonio Romano (Università degli Studi di Torino, LFSAG),
Marc Evrard (Université Paris-Saclay CNRS, LISN),
Alexandre François (LaTTiCe, CNRS, ENS-PSL, Université Sorbonne Nouvelle)

# Cartografia di innovazioni rispetto al latino attraverso un atlante sonoro dell'Europa

## *1. Introduzione*

L'Europa è caratterizzata da una grande diversità dialettale, che necessita ancora di essere documentata. Per più di un secolo, la favola esopica "Il vento di tramontana e il sole" è stata utilizzata dall'Associazione Fonetica Internazionale per illustrare molte lingue e dialetti parlati nel mondo. Su questa base, un atlante sonoro delle lingue regionali della Francia è stato progettato, prima di essere esteso ad altri paesi europei (Boula de Mareüil *et al.* 2019). Questo atlante linguistico, disponibile su https://atlas.limsi.fr, permette di ascoltare e leggere questa storia di un minuto in centinaia di versioni, in lingue minoritarie o dialetti. La maggior parte dei parlanti dell'atlante, registrati sul campo, hanno anche tradotto un elenco di un centinaio di parole (che si riferiscono in particolare alla fauna e alla flora) nelle loro varietà.

Sulla base di questi dati digitali e seguendo i principi della dialettometria (Goebl 2002; Gooskens 2005; Nerbonne *et al.* 2007; Patriarca *et al.* 2020), del metodo comparativo e soprattutto della glottometria storica (François 2014; Kalyan/François 2018), proponiamo strumenti computazionali per rispondere alle seguenti domande. C'è una maggiore variazione tra i dialetti romanzi settentrionali e meridionali, o tra l'ovest e l'est del dominio (dialetti ibero-romanzi da un lato, dialetti gallo- e italo-romanzi dall'altro)? Come possiamo quantificarlo? In che misura i raggruppamenti dipendono dai livelli considerati (fonetico, morfosintattico o lessicale)?

Questo studio si basa su un sottocorpus che illustra decine di dialetti romanzi provenienti da Francia, Belgio, Svizzera, Italia, Spagna e Portogallo, per i quali 145 innovazioni relative al latino sono state codificate sotto forma di una matrice di 1 (presenza) e 0 (assenza). Abbiamo codificato:

- innovazioni fonetiche, come il cambiamento E > [wa];
- innovazioni morfologiche, come la fusione degli imperfetti latini in -ABA- e -EBA-;
- innovazioni sintattiche, come la narrazione al passato prossimo;
- innovazioni lessicali, come la sostituzione di CUM 'con' da APUD HOQUE > *avec*.

Abbiamo applicato diverse tecniche di classificazione per visualizzare i *cluster* emergenti (sotto forma di alberi, proiezioni su un piano bidimensionale, ecc.) e disegnare le isoglosse principali. I risultati dei diversi metodi di analisi e calcolo verranno confrontati per proporne una sintesi che consenta di rivalutare l'ubicazione delle principali linee di divisione tra gruppi dialettali. Questi risultati gettano nuova luce sulla dialettologia: contribuiscono infatti a modellare le dinamiche di espansione territoriale dalla disgregazione del latino.

## *2. Corpus e metodi*

Il nostro atlante sonoro conta più di 200 punti di indagine, con dati raccolti principalmente tra il 2014 e il 2020. Fra questi, abbiamo selezionato 61 punti in diverse varietà romanze. Abbiamo privilegiato le varietà "medie", poco marcate o centrali all'interno delle loro aree linguistiche. Un altro criterio è stato quello di scegliere i casi dove eravamo rimasti in contatto con i parlanti, per poter porre loro domande in caso di dubbio o mancanze di informazioni. Gli informatori erano ben radicati nella loro co-

munità, provenienti da diversi background socio-professionali, e molto spesso in pensione. Abbiamo considerato un parlante di riferimento per ogni area dialettale romanza: 21 in Francia, 4 in Belgio, 4 in Svizzera, 6 in Spagna, 1 in Portogallo, 25 in Italia. Alcune varietà, condivise da diversi paesi, erano rappresentate diverse volte: per esempio il francoprovenzale, tra Francia, Svizzera e Italia; il catalano, tra Francia, Spagna e Italia (l'algherese parlato in Sardegna). La tabella 1 elenca le varietà o gruppi di varietà illustrati.

| Gallo-romanzo | Italo-romanzo | Reto-romanzo | Ibero-romanzo |
|---|---|---|---|
| Settentrionale (*Oïl*)<br>Meridionale (*Oc*)<br>Francoprovenzale<br>*Croissant* | Settentrionale<br>Mediano<br>Meridionale<br>Toscano<br>Gallurese-sassarese<br>Sardo | Ladino<br>Friulano<br>Romancio | Aragonese<br>Asturiano<br>Castigliano<br>Galiziano<br>Portoghese |

Tabella 1: Varietà o gruppi di varietà illustrati

La tabella 2 riporta alcuni esempi di innovazioni. Ad esempio, se l'/r/ apicale del latino è diventato un fonema dorsale (/ʁ/ francese), almeno all'inizio delle parole (#_), l'innovazione è indicata come '1'. Questo esempio mostra l'importanza di specificare il contesto nella codifica, perché questa innovazione può dipendere della posizione del fonema nella parola. La caratteristica seguente della tabella, che si riferisce al plurale dei sostantivi e degli aggettivi in *-i/-e*, è considerata come innovazione rispetto al caso accusativo plurale del latino in -S: è ciò che si osserva nelle varietà italo-romanze – ma non in sardo. L'ultima innovazione della tabella 2 si riferisce alla sostituzione degli imperfetti latini in ERA- con STABA-, che si osserva in quasi tutte le varietà di *oïl*: STABAT > *èsteût* in vallone, *était* 'era' in francese ecc.

| Innovazione | | | Fonetica | Morfologica | Lessicale |
|---|---|---|---|---|---|
| Paese | cat | Città | R > [ʁ] / # _ | Plurali in *-i/-e* | ERA- sostituito da STABA- |
| Belgio | wal | Liegi | 1 | 0 | 1 |
| Italia | lig | Varazze | 1 | 1 | 0 |
| Portogallo | por | Lisbona | 1 | 0 | 0 |
| Svizzera | rum | Zuoz | 0 | 0 | 0 |

Tabella 2: Estratto della tabella di annotazione

Altre innovazioni sono legate alla linea La Spezia-Rimini, che distingue le varietà gallo-italiche (a nord di essa, per semplificare) dalle varietà dell'Italia centro-meridionale (Chambers/Trudgill 2004): la sonorizzazione, l'indebolimento o la perdita delle consonanti sorde (principalmente le occlusive intervocaliche) caratterizzano le località sopra la linea. Esempio: latino SECURU(M) > ligure *segûo*, provenzale *segur*, francese *sûr*.

Abbiamo utilizzato diversi algoritmi di classificazione e di selezione di attributi, tra i quali il *clustering* gerarchico agglomerativo (noto anche come metodo di Ward) e gli alberi di decisione, che forniscono rappresentazioni facilmente leggibili. Sono state utilizzate le librerie Python Scikit-learn e Plotly, in particolare per visualizzare i risultati sotto forma di coropleti (cioè superfici colorate, che possono corrispondere alle nostre aree dialettali).

## *3. Risultati*

Dall'intero insieme di caratteristiche, i punti di indagine considerati possono essere classificati in base al numero di innovazioni che mostrano, e questo numero può essere associato a un colore più o meno scuro su una mappa di base in formato GeoJSON poi vettorializzata.

La Francia d'*oïl* (gallo-romanza settentrionale) è la più innovativa (cfr. Figura 1). È il caso, in particolare, del dialetto angioino, nel nord-ovest del dominio (con più di 65 innovazioni, tra cui

la palatalizzazione CL > [Cj], sconosciuta in francese), mentre in Francia la Corsica (affiliata al gruppo toscano) è la regione con il minor numero di innovazioni rispetto al latino. A livello europeo, il toscano è il più conservatore, secondo le nostre misurazioni (meno di 30 innovazioni). Tra questi due estremi, le varietà ibero-romanze e le varietà italo-romanze meridionali presentano circa 35 innovazioni.

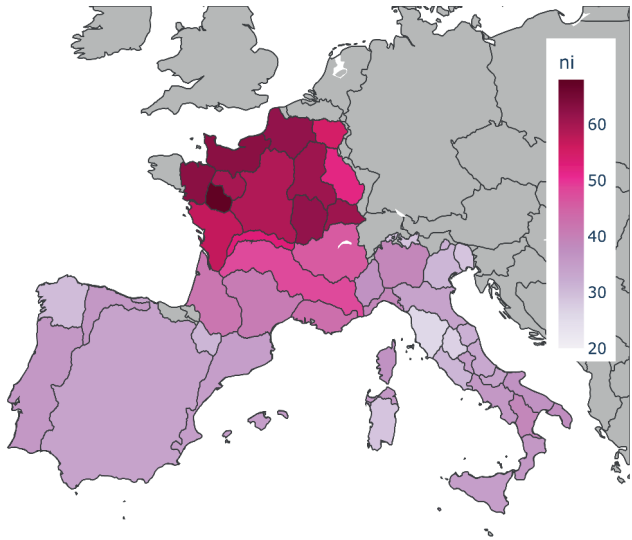


Figura 1: Mappa linguistica delle aree romanze tanto più innovative in quanto il colore è scuro

Questi punteggi sono in parte in accordo con Pei (1949), il quale, sulla base di modificazioni fonologiche delle vocali accentate rispetto al latino, concluse che il sardo e l'italiano sono i più conservatori, il francese il più innovativo. Sono in disaccordo con Walter (1994: 119), la quale, sulla base del vocabolario, rileva che le lingue più periferiche come lo spagnolo e il portoghese tendono ad essere più conservatrici rispetto a lingue più centrali come

l'italiano e il francese: per fare un esempio dal nostro elenco di parole isolate, citiamo lo spagnolo *yegua* o il portoghese *égua* < EQUA(M) 'cavalla', mentre l'italiano e il francese utilizzano forme riflesse di altri etimi. Con gli stessi principi presentati da Cugno (2023), il presente lavoro fornisce un quadro quantitativo che permette di distinguere tra il toscano (che servì come norma di prestigio per la lingua italiana) e gli altri dialetti italiani. Inoltre, permette di distinguere tra diversi livelli linguistici.

I risultati del *clustering* gerarchico, ottenuti utilizzando tutte le innovazioni o solo un sottoinsieme per garantire la robustezza e la parsimonia dell'approccio, forniscono risposte euristiche alle domande sollevate in introduzione. Nel dendrogramma semplificato della Figura 2, potato con 7 foglie, l'analisi dei *cluster* conferma una divisione Nord/Sud, con una divisione che passa attraverso il centro dell'area occitana (gallo-romanza meridionale) con l'*oïl* e i dialetti intermedi che si raggruppano al Nord e il resto al Sud. Nel secondo ramo del dendrogramma, la divisione principale è tra le varietà sud-occitane, catalane e ibero-romanze da un lato, le varietà reto- e italo-romanze dall'altro. La divisione successiva oppone il Nord e il Sud della linea La Spezia-Rimini: le varietà reto-romanze e italo-romanze settentrionali al Nord, le varietà italo-romanze centrali (+ il sardo) da un lato, le varietà italo-romanze meridionali dall'altro, al Sud.

Diversi colori sono stati associati alle 7 classi risultanti dal dendrogramma della Figura 2 e proiettati sulla mappa della Figura 3. È interessante notare che lingue come lo spagnolo e il portoghese appaiono più vicine tra loro rispetto, ad esempio, al lombardo milanese e al campano napoletano. Inoltre, il ligure (gruppo gallo-italico), l'abruzzese, il sardo, il siciliano, il calabrese e il salentino sono raggruppati con i dialetti italo-romanzi mediani – con il gallurese-sassarese, il corso e il toscano. Nei nostri dati, il ligure si differenzia dal suo vicino lombardo più di quanto lo faccia il lombardo rispetto ai suoi vicini piemontese

ed emiliano-romagnolo. Ad esempio, il ligure ha mantenuto le vocali posttoniche più dei suoi vicini: confronta il ligure *forte* < FORTE(M) con il piemontese, il lombardo o l'emiliano-romagnolo *fòrt/fort*. Nel Mezzogiorno d'Italia, occorre distinguere tra i dialetti alto-meridionali (il pugliese, il lucano, il campano, il molisano), e i dialetti estremo-meridionali, imparentati con il siciliano (Pellegrini 1977). Per inciso, i campioni liguri si distanziano anche sul piano delle metriche ritmiche da quelli piemontesi analizzati con gli stessi metodi da Romano/Mairano/Pollifrone (2010), ricadendo nell'area in cui si collocano abitualmente campioni di parlato riconducibili a lingue *syllable-timed*.

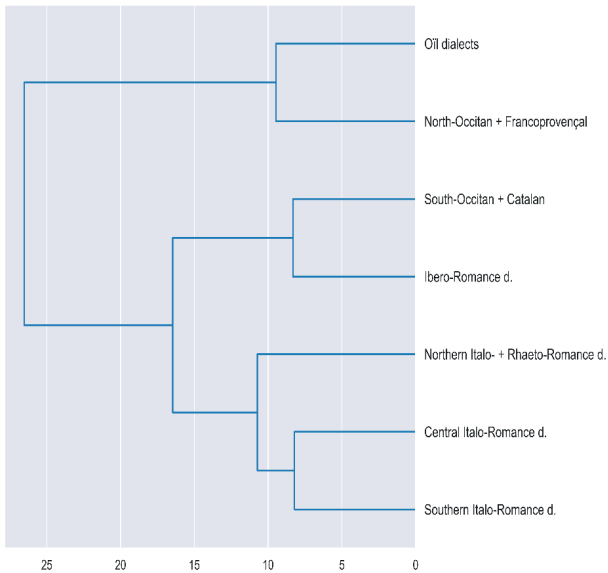


Figura 2: Dendrogramma semplificato risultante da un *clustering* con 7 classi

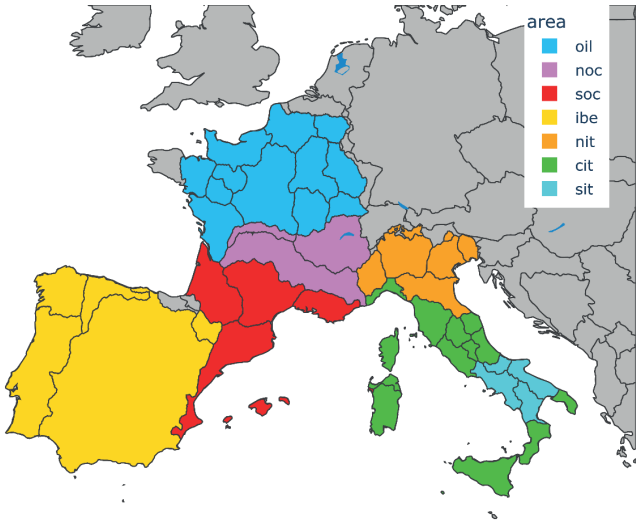


Figura 3: Mappa corrispondente al dendrogramma con 7 *cluster*

Infine, abbiamo cercato di determinare le caratteristiche più importanti o discriminanti per ottenere il *clustering* con 7 classi riportato nella Figura 2, utilizzando diversi algoritmi come gli alberi di decisione. Tra queste caratteristiche più importanti vi sono la palatalizzazione della sillaba latina CA, che caratterizza la maggior parte dei dialetti gallo-romanzi settentrionali, e la semplificazione delle geminate al nord della linea La Spezia-Rimini. La maggior parte delle innovazioni riguardano tratti fonetici e fonologici, ma abbiamo anche caratteristiche morfosintattiche e lessicali, come il soggetto non nullo nelle varietà gallo-romanze settentrionali, e la sostituzione di CUM 'con' da AMPHI > *amb* nelle varietà occitano-romanze. Conservando soltanto le innovazioni morfosintattiche (che nella nostra annotazione sono 35), troviamo ancora una tripartizione Nord, Sud-Est, Sud-Ovest.

## *4. Conclusioni*

Questo studio ha evidenziato la diversità delle lingue minoritarie nello spazio romanzo. Nonostante le dimensioni di questo spazio, la raccolta sul campo di dati a grana fine – anche da un piccolo corpus – si è rivelata utile per svelare con precisione la distribuzione delle varietà romanze. L'uso di metodi computazionali di apprendimento supervisionato o non supervisionato offre nuove informazioni sulla classificazione delle varietà romanze.

I risultati confermano la robustezza della divisione Nord/Sud, con l'area d'*oïl* che è (di gran lunga) la più innovativa e, secondariamente, un'opposizione tra il Sud-Ovest (soprattutto ibero-romanzo) e il Sud-Est (soprattutto italo-romanzo). Non pretendiamo che ciò dia un'immagine della ricostruzione storica, separando innanzitutto il gallo-romanzo dal latino. Al contrario, la Sardegna e la Lucania, le prime regioni conquistate dai Romani, si separarono prima (Adams 2007; Goldstein 2023). L'argomentazione recente non può non poggiare naturalmente sui capisaldi posti da Lausberg (1939). Tuttavia, la stima del tempo di divergenza richiederebbe la datazione di un centinaio di innovazioni se volessimo perseguire il nostro approccio quantitativo con tecniche di apprendimento automatico.

Non pretendiamo nemmeno che l'approccio seguito identifichi le isoglosse percepite come le più importanti. Un'innovazione che investe l'intero lessico, come il passaggio dal plurale in -s a plurali in *-i/-e*, è senza dubbio di maggiore impatto di qualsiasi cambiamento fonetico irregolare o innovazione lessicale. La ponderazione delle innovazioni rispetto al latino va oltre lo scopo di questo articolo e richiederebbe grandi quantità di dati dialettali. Stiamo attualmente valutando modalità semi-automatiche per estendere l'attuale approccio a più punti di indagine, tra cui l'Istria e la Romania. In tutti i casi, i risultati dovranno essere confrontati con quelli forniti da valutazioni basate su misure di distanza (di tipo Levenshtein o acustiche): questa è un'altra area di ricerca che ha applicazioni di geolocalizzazione.

## *5. Ringraziamenti*

Questo lavoro fa parte del progetto “Computational Language Documentation by 2025” (CLD2025) finanziato dall’Agenzia francese per la ricerca (ANR) e dalla Fondazione tedesca per la ricerca (DFG). Siamo grati a Frédéric Vernier per il suo aiuto con la cartografia e a Cédric Jr Tonga per l’elaborazione dei dati. Ringraziamo inoltre Lorenza Brasile, Lucia Molinu, Michela Russo e Paolo Roseano per i loro consigli su alcune varietà italiane.

## Bibliografia


Adams, J. N. *The Regional Riversification of Latin 200 BC–AD 600*, CUP, Cambridge 2007.

Boula de Mareüil, P./De Iacovo, V./Romano, A./Vernier, F. *Un atlante sonoro delle lingue di Francia e d'Italia: focus sulle parlate liguri*, in *Il patrimonio linguistico storico della Liguria. Attualità e futuro*, a cura di F. Toso, Insedicesimo, Savona 2019, pp. 33-46.

Chambers, J./Trudgill, P. *Dialectology*, CUP, Cambridge 2004.

Cugno, F. *Italian Dialect Classifications*, in *Dialectologia* 10, 2023, pp. 197-230.

François, A. *Trees, Waves and Linkages: Models of Language Diversification*, in *The Routledge Handbook of Historical Linguistics* a cura di C. Bowern, B. Evans, Routledge, London 2014, pp. 161-189.

Goebl, H. *Analyse dialectométrique des structures de profondeur de l'ALF*, in *Revue de linguistique romane,* 66/261–262, 2002, pp. 5-63.

Goldstein, D. *Divergence-Time Estimation in Indo-European The case of Latin*, in *Diachronica*, 40/3, pp. 1-53.

Gooskens, C. *Travel Time as a Predictor of Linguistic Distance*, in *Dialectologia et Geolinguistica,* 13, 2005, pp. 38-62.

Kalyan, S./François, A. *Freeing the Comparative Method from the Tree Model: A Framework for Historical Glottometry*, in *Let's Talk about Trees: Tackling Problems in Representing Phylogenic Relationships among Languages*, a cura di R. Kikusawa, L. Reid, National Museum of Ethnology, Osaka 2018, pp. 59-89.

Lausberg, H. *Die Mundarten Südlukaniens*, Niemeyer, Halle 1939.

Nerbonne, J./Kleiweg, P./Heeringa, W./Manni, F. *Projecting Dialect Differences to Geography: Bootstrap Clustering vs.*

*Noisy Clustering*, in *Data Analysis, Machine Learning, and Applicationsy*, a cura di C. Preisach, L. Schmidt-Thieme, H. Burkhardt, R. Decker, Springer, Berlin 2007, pp. 647-654.

Patriarca, M./Heinsalu, E./Léonard, J. L. *Language in Space and Time*, CUP, Cambridge 2020.

Pei, M. A. *A New Methodology for Romance Classification*, in *Word* 5/2, 1949, pp. 135-146.

Pellegrini, G. B. *Carta dei dialetti d'Italia*, Pacini, Pisa 1977.

Romano, A./Mairano, P./Pollifrone, B. *Variabilità ritmica di varietà dialettali del Piemonte*, in *La dimensione temporale del parlato*, a cura di S. Schmid, M. Schwarzenbach, D. Studer, EDK, Torriana 2010, pp. 101-112.

Walter, H. *L'aventure des langues en Occident*, Éditions Robert Laffont, Paris 1994.