

Modeling Alphabet Diffusion and Community Interactions in Ancient Italy with Spatiotemporal Point Processes

Tom Francis, Muhammad Rehan
tdfrancis@ucla.edu, rehanmuh@ucla.edu

University of California, Los Angeles
Connected Past 2026

September 23, 2026



Scan the QR code for slides

Overview

1 Introduction

2 Data

3 Model

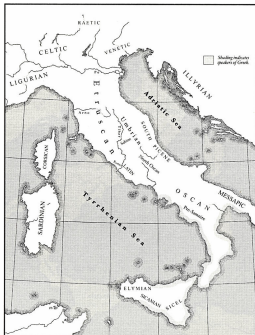
4 Results and Interpretation

Introduction



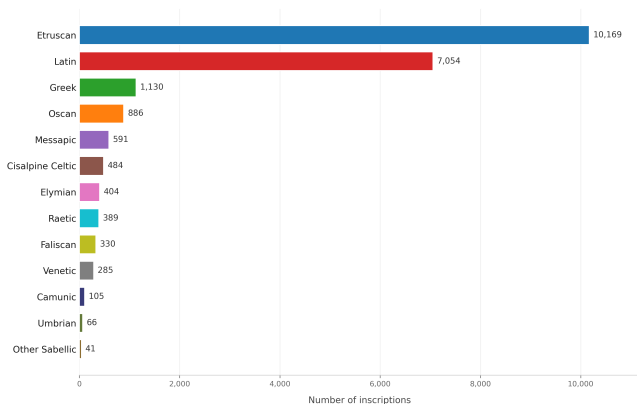
Languages attested in the inscriptional record in 1st millennium BCE Italy

A. Language map of ancient Italy

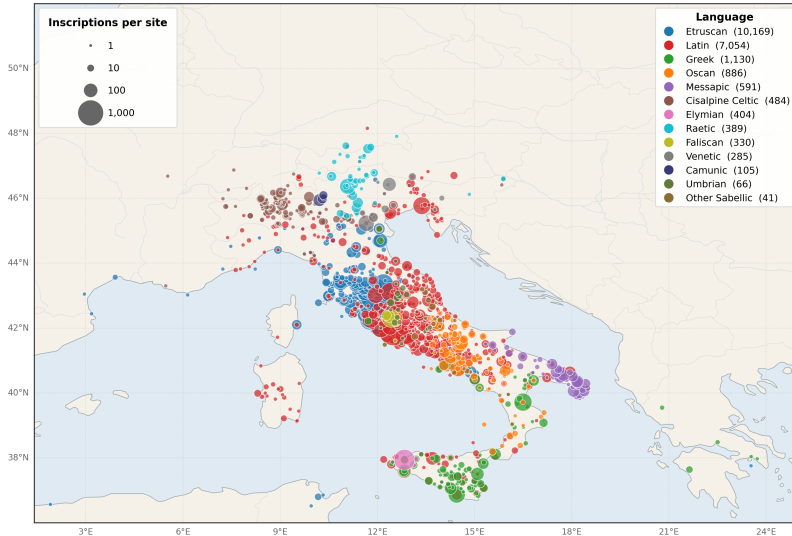


Wallace 2007:x

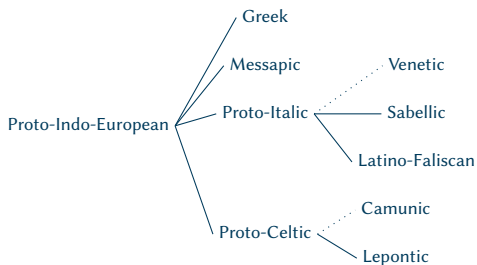
B. Language distribution in the dataset



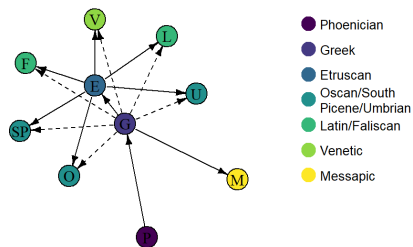
Spatial distribution of languages in the epigraphic record



Linguistic relationships vs. alphabet relationships

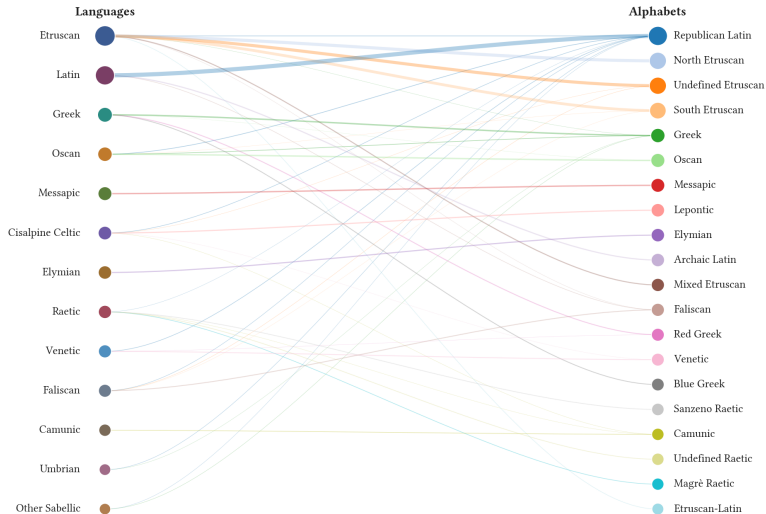


Phylogeny of Indo-European languages in Ancient Italy



Generally accepted direction of alphabet spread in Italy

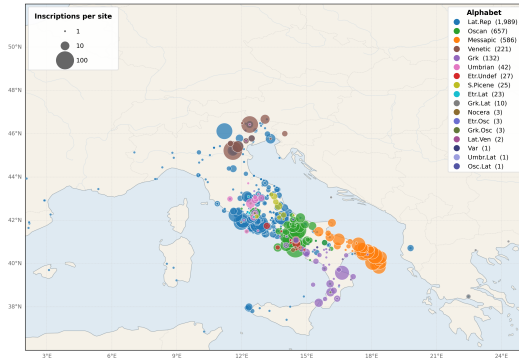
Alphabet relationships vs. linguistic relationships



Data

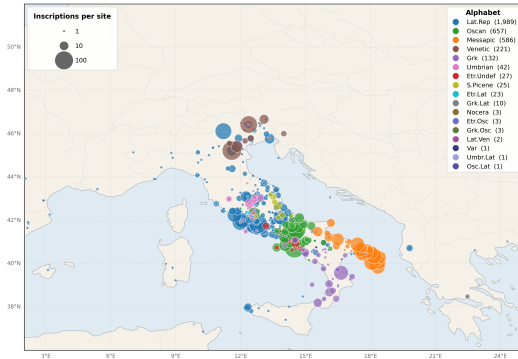


Comparison of alphabets in the epigraphic record between databases

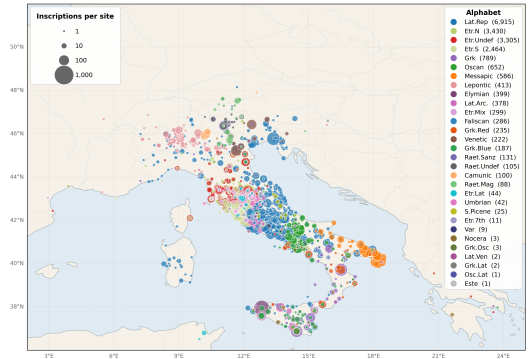


CEIPOM's 3,868-inscription dataset

Comparison of alphabets in the epigraphic record between databases



CEIPOM's 3,868-inscription dataset

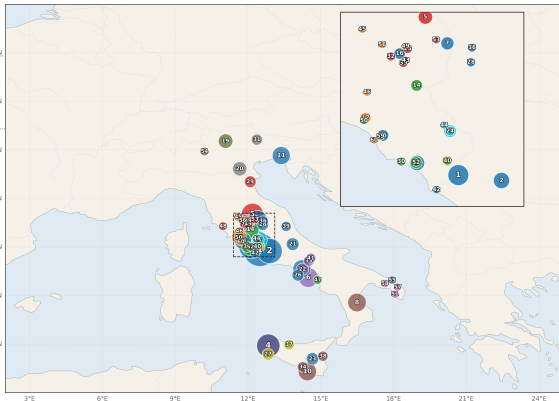


Current 21,937 inscription dataset

Alphabet distribution: major epigraphic centers (N ≥ 50)

Sites by inscription count

1. Roma
2. Castrubrium (Cadore)
3. Palestrina / Praeneste
4. Cerveteri / Agrippa
5. Segesta
6. Aliphan
7. Veii
8. Cosa
9. Ostia Antica
10. Capua
11. Todi
12. Perugia / Perusia
13. Spina
14. Cosa
15. Cosa
16. Cosa
17. Cosa
18. Cosa
19. Cosa
20. Cosa
21. Cosa
22. Cosa
23. Cosa
24. Cosa
25. Cosa
26. Cosa
27. Cosa
28. Cosa
29. Cosa
30. Cosa
31. Cosa
32. Cosa
33. Cosa
34. Cosa
35. Cosa
36. Cosa
37. Cosa
38. Cosa
39. Cosa
40. Cosa
41. Cosa
42. Cosa
43. Cosa
44. Cosa
45. Cosa
46. Cosa
47. Cosa
48. Cosa
49. Cosa
50. Cosa



Dominant alphabet

- Lat. Rep. (16 sites)
- Etr. N. (8 sites)
- Etr. S. (8 sites)
- Etr. Undef. (6 sites)
- Oscan (4 sites)
- Griek (4 sites)
- Messapic (3 sites)
- Faliscan (3 sites)
- Lat. Arc. (3 sites)
- Venetian (2 sites)
- Elymian (1 site)
- Raet. Sanz. (1 site)
- Camunic (1 site)

Inscriptions per site

- 50
- 100
- 250
- 500
- 1,000

Research questions

- When moving outside of a Greco-Roman conception of ancient Italy, what new insights into diffusion and community interactions in 1st millennium BCE can be inferred from the inscriptional record?

Research questions

- When moving outside of a Greco-Roman conception of ancient Italy, what new insights into diffusion and community interactions in 1st millennium BCE can be inferred from the inscriptional record?
- Does a quantitative analysis support the oft-made claim about the monolithic influence of Latin on changes in alphabet distribution in the final 2 centuries BCE (e.g., De Simone and Marchesini 2002; Wallace 2007; Wallace 2008), or does it suggest that narratives of alphabetic resistance are lost in this explanation?

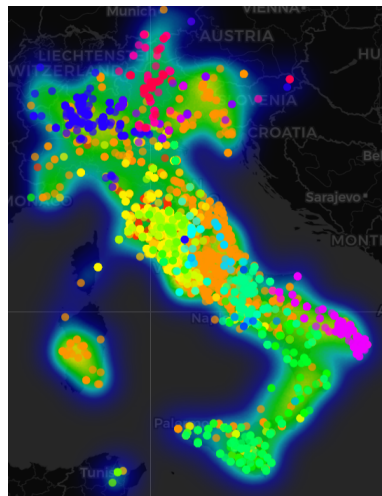
Research questions

- When moving outside of a Greco-Roman conception of ancient Italy, what new insights into diffusion and community interactions in 1st millennium BCE can be inferred from the inscriptional record?
- Does a quantitative analysis support the oft-made claim about the monolithic influence of Latin on changes in alphabet distribution in the final 2 centuries BCE (e.g., De Simone and Marchesini 2002; Wallace 2007; Wallace 2008), or does it suggest that narratives of alphabetic resistance are lost in this explanation?
- Can community interaction dynamics be inferred at scale using spatiotemporal point processes over features of material culture?

Data collection for the model: variables

- We utilize six recorded attributes for every inscription:

- Alphabet
- Object type
- Dating interval of an inscription
- Geographic coordinates
- Findspot name
- Region of findspot

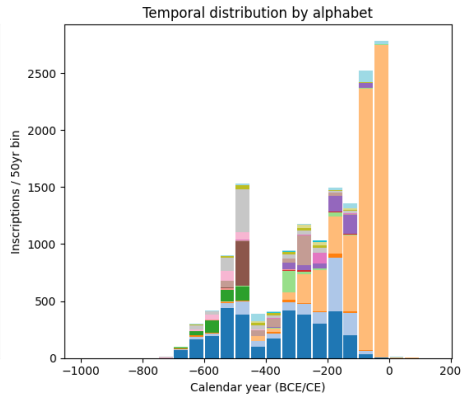
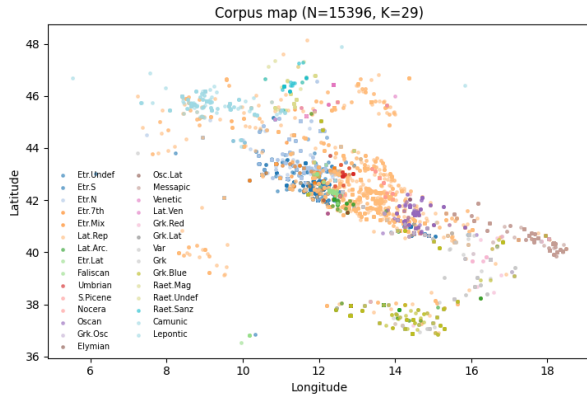


Spatial density of all inscriptions

Data collection for the model: major editions and databases consulted

Language	Primary Edition(s) Used	Database(s) Used
Etruscan	Rix and Meiser 2014	N/A
Greek	Dubois 1989, 1995, 2002, 2008	I.Sicily
Latin	<i>CIL</i> I ² , III, VI, XI	<i>CEIPoM</i> , <i>EDCS</i>
Elymian	Agostiniano 1977, 2021	I.Sicily
Raetic	Volumes cited in <i>TIR</i>	<i>TIR</i>
Lepontic	Volumes cited in <i>Lexicon Leponticum</i>	<i>Lexicon Leponticum</i>
Cisalpine Gaulish	Volumes cited in <i>Lexicon Leponticum</i>	<i>Lexicon Leponticum</i>
Sabellic Languages	Crawford et al. 2011, Rix 2002	<i>CEIPoM</i>
Faliscan	Bakkum 2009	<i>CEIPoM</i>
Messapic	De Simone and Marchesini 2002	<i>CEIPoM</i>
Camunic	Tibiletti Bruno 1990, Prosdocimi 1965	N/A
Venetic	Lejeune 1974, Prosdocimi and Pellegrini 1967	<i>CEIPoM</i>

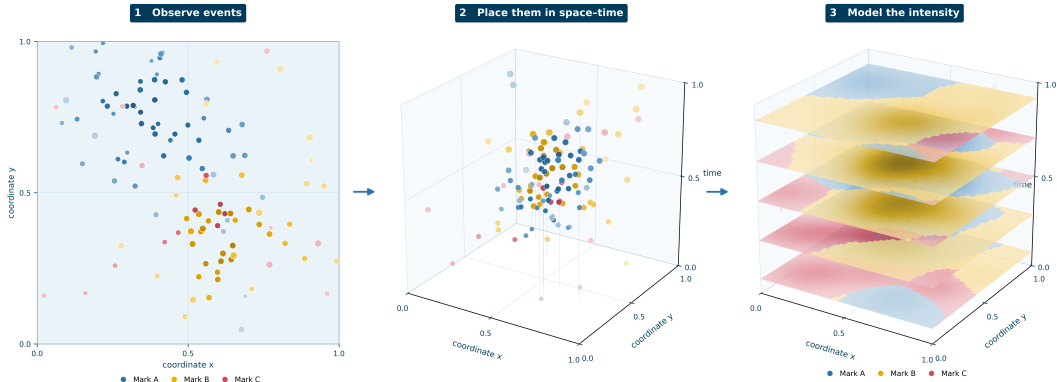
Corpus data for modeling



Model



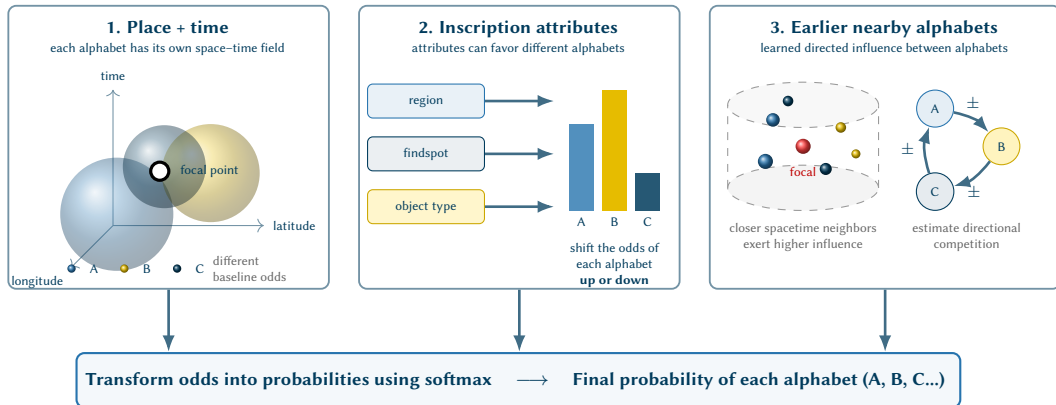
What is a spatiotemporal point process?



Spatiotemporal point process estimate when and where particular event types are likely to occur

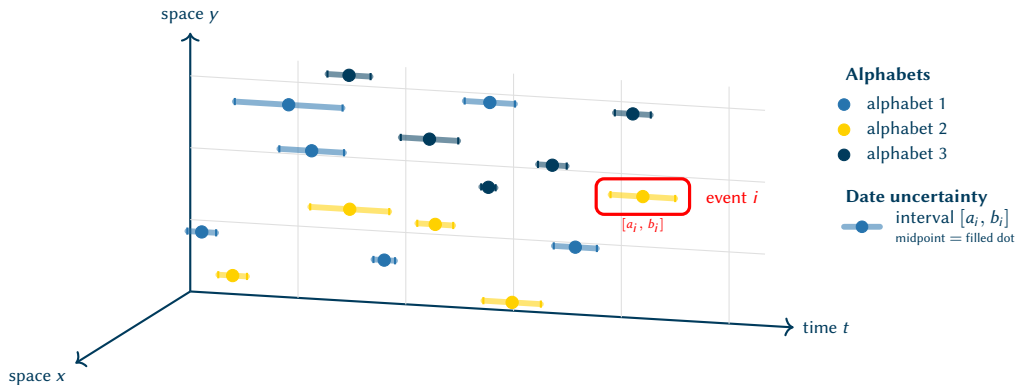
Applying spatiotemporal point processes to alphabet data

For an inscription at a given **place and date interval** we calculate the odds of that inscription being written in a given alphabet and convert these odds into probabilities.



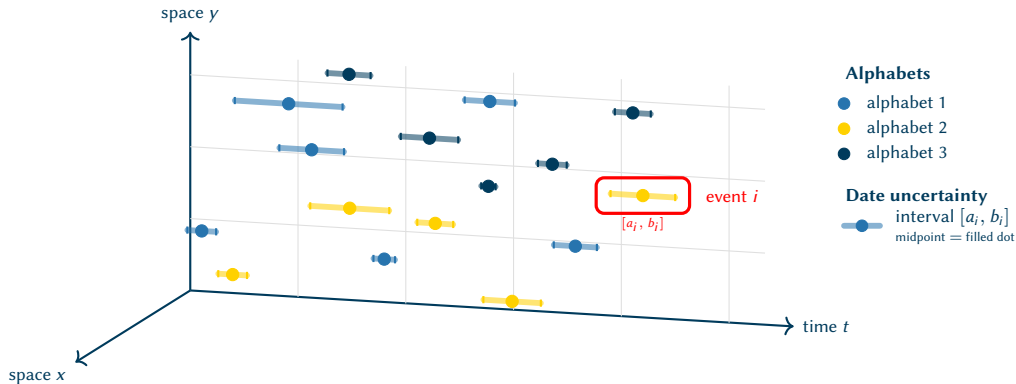
Marked Point Process modeling of alphabet dynamics

Each inscription is a point in space-time and **marked** with an alphabet. The date is almost always an *interval* $[a_i, b_i]$.



Marked Point Process modeling of alphabet dynamics

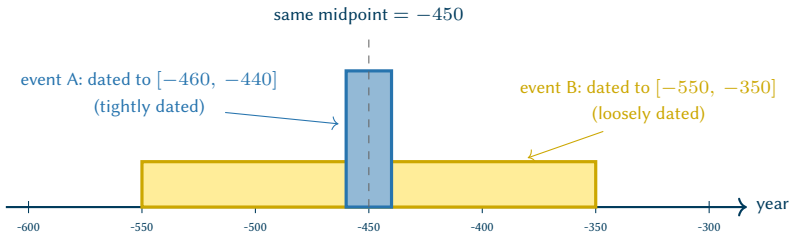
Each inscription is a point in space-time and **marked** with an alphabet. The date is almost always an *interval* $[a_i, b_i]$.



For each event i : **given location, time interval, and the alphabets of nearby earlier events, what's the probability i is in alphabet k ?**

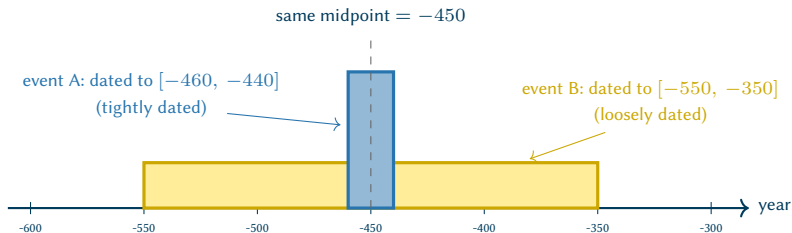
Problems with midpoint dating

- Most spatiotemporal analyses represent the dating interval as a midpoint.



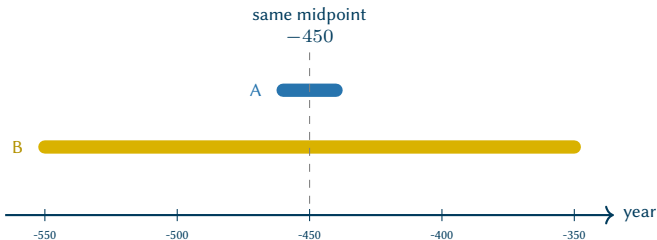
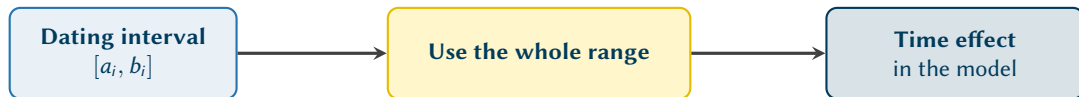
Problems with midpoint dating

- Most spatiotemporal analyses represent the dating interval as a midpoint.



- A midpoint-only model would treat A and B as *identical*.
- Both are centred on -450 , but A's date is more certain than B's.

Modeling dating uncertainty

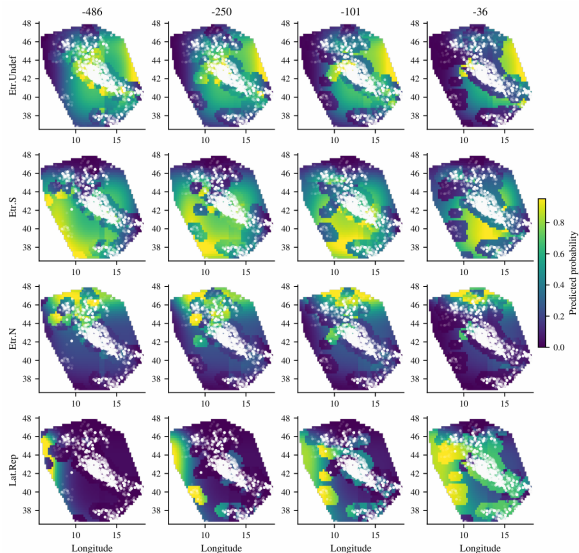


- Our model allows the **width of the dating interval** to affect the estimated time effect by using higher order moments of the date ranges.

Results and Interpretation

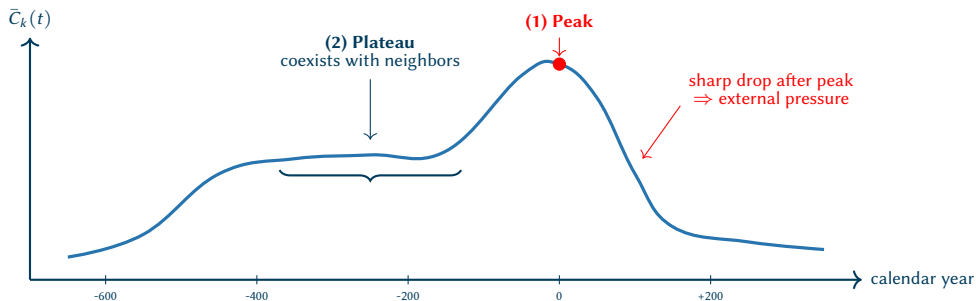


Spatiotemporal probability fields for selected alphabets



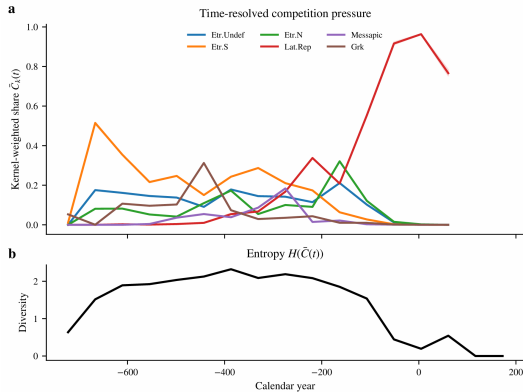
Change in competition pressures

Each curve shows $\bar{C}_k(t)$ for one alphabet k . Three things to look for:

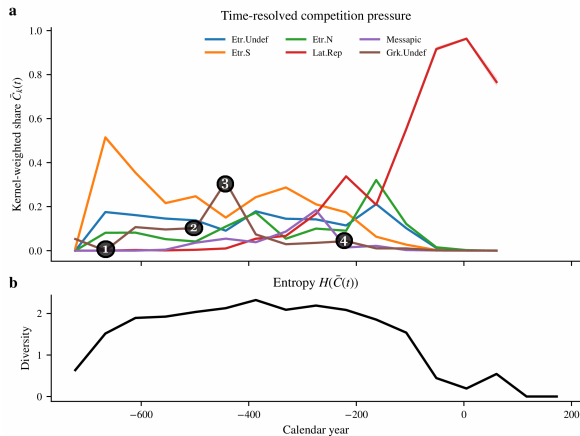


Interpretation: $\bar{C}_k(t)$ is a kernel-weighted *share* of the neighborhood, which is robust to overall inscription rates and recovery bias.

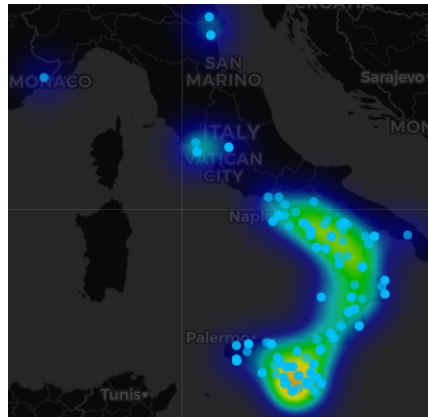
Interpreting Alphabet Competition Over Time: Historical Events



Interpreting Alphabet Competition Over Time: Greek Alphabet

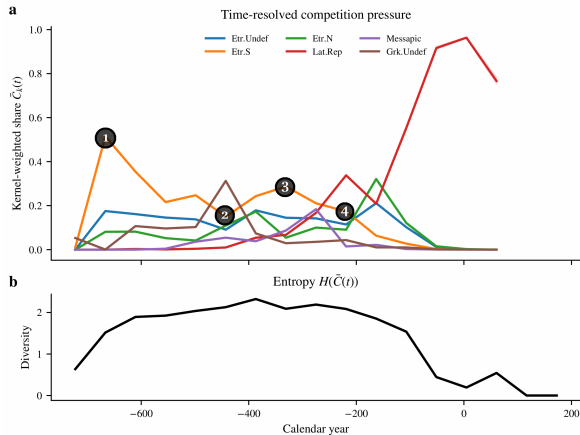


(a) Time-resolved competition pressure

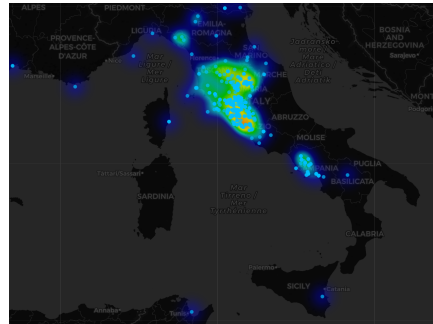


(b) Heatmap of inscriptions in the Greek alphabet

Interpreting Alphabet Competition Over Time: Southern Etruscan Alphabet

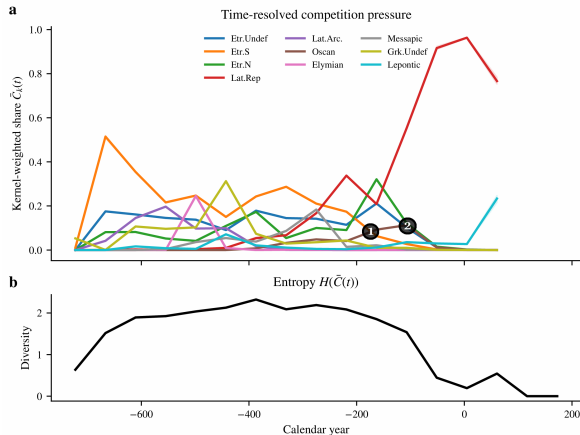


(a) Time-resolved competition pressure

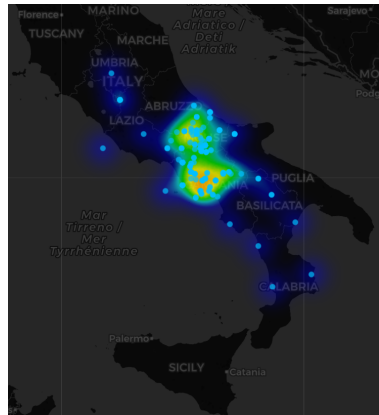


(b) Heatmap of inscriptions in the Southern Etruscan alphabet

Interpreting Alphabet Competition Over Time: Southern Etruscan Alphabet and Oscan (not just Latin!)

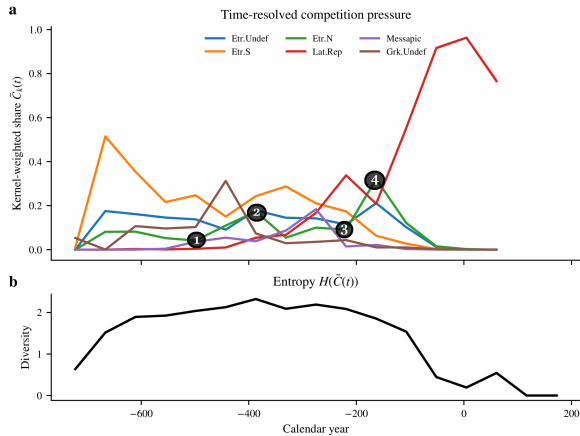


(a) Time-resolved competition pressure

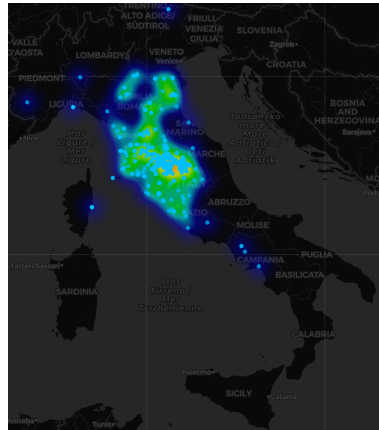


(b) Heatmap of inscriptions in the Oscan alphabet

Interpreting Alphabet Competition Over Time: Northern Etruscan Alphabet



(a) Time-resolved competition pressure



(b) Heatmap of inscriptions in the Northern Etruscan alphabet

Conclusions

- Roman conquest does explain *some* changes in alphabet competition, but colonization efforts and migration by several other actors (including Punics, Greeks, Syracusans, Oscans, Etruscans, etc.) are also essential for understanding alphabet dynamics in 1st millennium BCE.

Conclusions

- Roman conquest does explain *some* changes in alphabet competition, but colonization efforts and migration by several other actors (including Punics, Greeks, Syracusans, Oscans, Etruscans, etc.) are also essential for understanding alphabet dynamics in 1st millennium BCE.
- The privileging of literary texts perpetuates ancient Greco-Roman narratives and silences the voices of the other groups of language speakers in Italy whose voice is recorded in the vibrant multi-lingual inscriptional record.

Conclusions

- Roman conquest does explain *some* changes in alphabet competition, but colonization efforts and migration by several other actors (including Punics, Greeks, Syracusans, Oscans, Etruscans, etc.) are also essential for understanding alphabet dynamics in 1st millennium BCE.
- The privileging of literary texts perpetuates ancient Greco-Roman narratives and silences the voices of the other groups of language speakers in Italy whose voice is recorded in the vibrant multi-lingual inscriptional record.
- Spatiotemporal point processes provide fine-grained insights into the relative contribution of different factors in influencing the attestation of different alphabets in ancient Italy.

Thank you for listening!

Comments? Questions?

References I

-  Agostiniano, Luciano (1977). *Iscrizioni anelleniche di Sicilia I. Le iscrizioni elime*. Firenze: L. S. Olschki.
-  — (2021). *Iscrizioni anelleniche di Sicilia I. Le iscrizioni elime. Appendice 1978-2020*. Roma: L'Erma di Bretschneider.
-  Bakkum, Gabriël Cornelis Leonides Maria, ed. (2009). *The Latin dialect of the Ager Faliscus: 150 years of scholarship*. Amsterdam: Amsterdam University Press. 2 pp. ISBN: 978-90-5629-562-2 978-90-485-0875-4.
-  Crawford, Michael Hewson et al. (2011). *Imagines Italicae: a corpus of Italic inscriptions*. 3 vols. Bulletin of the Institute of Classical Studies supplement 110. London: Institute of Classical Studies, School of Advanced Study, University of London. ISBN: 978-1-905670-30-7 978-1-905670-35-2 978-1-905670-36-9.
-  De Simone, Carlo and Simona Marchesini (2002). *Monumenta Linguae Messapicae*. Wiesbaden: L. Reichert. ISBN: 978-3-89500-228-1.

References II

-  Dubois, Laurent (1989). *Inscriptions grecques dialectales de Sicile. 1.* Collection de l'Ecole Française de Rome 119. Num Pages: 303. Rome: École Française. ISBN: 978-2-7283-0185-0.
-  — (1995). *Inscriptions grecques dialectales de Grande Grèce. 1: Colonies eubéennes. Colonies ioniennes. Emporia.* École Pratique des Hautes Études, IVe Section, Sciences Historiques et Philologiques III, Hautes études du monde gréco-romain 21. Num Pages: 202. Genève: Droz. ISBN: 978-2-600-00085-7.
-  — (2002). *Inscriptions grecques dialectales de Grande Grèce. 2 Colonies achéennes.* Hautes études du monde gréco-romain 30. Genève: Droz. ISBN: 978-2-600-00690-3.
-  — (2008). *Inscriptions grecques dialectales de Sicile. 2.* Hautes études du monde gréco-romain 40. Genève: Droz. ISBN: 978-2-600-01340-6.
-  Lejeune, Michel (1974). *Manuel de la langue Vénète.* Indogermanische Bibliothek 1. Reihe. Heidelberg: C. Winter. ISBN: 978-3-533-02353-1.
-  Prosdocimi, Aldo Luigi (1965). “Per una edizione delle iscrizioni della Val Camonica”. In: *Studi Etruschi* 33, pp. 575–599.

References III

-  Prosdocimi, Aldo Luigi and G. B. Pellegrini (1967). *La lingua venetica I: Le iscrizioni*. Fir: Istituto di glottologia dell'Università di Padova–Circolo Linguistico Fiorentino.
-  Rix, Helmut (2002). *Sabellische Texte: die Texte des Oskischen, Umbrischen und Südpikenischen*. Handbuch der italischen Dialekte / begr. von Emil Vetter 5. Heidelberg: Winter. 227 pp. ISBN: 978-3-8253-0853-7.
-  Rix, Helmut and Gerhard Meiser (2014). *Etruskische Texte: editio minor*. Neu Bearbeit. Studien zur historisch-vergleichenden Sprachwissenschaft Bd. 4. Hamburg: Baar. ISBN: 978-3-935536-71-4.
-  Tibiletti Bruno, Maria Grazia (1990). “Nuove iscrizioni camune”. In: *Quaderni Camuni* 49-50, pp. 29–171.
-  Wallace, Rex E. (2007). *The Sabellic languages of ancient Italy*. Languages of the world / Materials 371. München: LINCOM Europa. 72 pp. ISBN: 978-3-89586-990-7.
-  — (2008). *Zikh Rasna: a manual of the Etruscan language and inscriptions*. Ann Arbor, Mich.: Beech Stave Press. 271 pp. ISBN: 978-0-9747927-4-3.