Current trends in computational geographical dialectology

Organizers
Outi Vesakoski (outves@utu.fi), University of Turku, Finland
Tatiana Kachkovskaia (tatiana.kachkovskaia@utu.fi), University of Turku, Finland
Peter Ranacher (peter.ranacher@geo.uzh.ch), PH Graubünden, Switzerland; University of Zürich, Switzerland
Karl Pajusalu (karl.pajusalu@ut.ee), University of Tartu, Estonia

We are currently in a period of the history of our disciplines where new digital datasets documenting historical language variation and dialect diversity are regularly becoming available. These range from born-digital materials to recently digitalized historical dialect atlases built upon the work of our analog forebears (linguistic microtypology of Finnic languages, EstTyp, Norvik et al. forthcoming; crowdsourced datasets such as Leemann et al., 2018 and Romansh idioms, forthcoming Grünert et al.; dialect atlases of Japanese, Huisman forthcoming; Danish, Wichman submitted; Russian, Wichmann, Jonjić & Naranjo 2026; Finnish Santaharju et al. 2025 etc.). These are arriving together with evolving computational, and especially geostatistical approaches which herald a new era in the study of the patterns and drivers of synchronic and diachronic linguistic variation. This workshop aims at reaching beyond linguistics only by promoting the usage of spatial linguistic data in interdisciplinary settings by showcasing how geostatistical methods could turn the linguistic data into new measures of human spatial variation.

Following Goebl (2010:435), the overarching aim of dialectology is to "overcome the apparently uncontrollable chaos of the data from the linguistic atlas". Since the compilation of first dialectal atlases over a hundred years ago (published in Gilliéron 1902-10, Saareste 1938, Kettunen 1940, etc), the "isogloss chaos" has begun to be tamed with quantitative and computational approaches, and toolkits have become available with easy-to-use public interfaces such as *Diatech* (Aurrekoetxea et al. 2015) and *Gabmap* (Nerbonne et al. 2011). Nevertheless, current geospatial methods allow many more possibilities to extract information from these new datasets of local linguistic variation. We wish to promote the development of the fields termed geographical dialectology, dialectometry, spatial dialectology, geolinguistics, geolinguistic dialectology etc. In this workshop we want to bring together researchers from a variety of methodological backgrounds to discuss the current trends in computational dialectology and dialectometry as well as to explore the potential of the new methods aimed at taming the "isogloss chaos".

The workshop aims to cover (at least) three different aspects of dialectometry: identifying the spatial patterns, studying the drivers of spatial dialectal patterns and cross-analysing the evolution of the dialectal landscape in interdisciplinary settings.

Today the statistical methods allow for uncovering the overall **spatial linguistic patterns** over individual isoglosses, averaging or pooling information from multiple, potentially contradictory, isoglosses (Nerbonne & Kretzschmar 2003; for overview of dialect maps, see the user interface *Dialektkarten*). Current trends in the field include further comparison of linguistic landscapes identified from the different linguistic domains: Norvik et al. (forthcoming) studied spatial variation in Finnic languages (subgroups of Uralic languages, sometimes referred as a dialect continuum) comparing phonetic and morphosyntactic linguistic features. Their result indicates different spatial patterns in these two domains.

Geostatistical methods further allow for reaching beyond the areal patterns and ask why they emerged: what are the **mechanisms** driving the observed spatial linguistic variation? These studies are needed to test a range of hypotheses and assumptions, and to reveal the factors explaining spatial linguistic variation. For example, contrary to expectations, the differences between Finnish regional dialects do not follow the isolation-by-distance hypothesis, but instead the dialects are more similar as cultural and environmental attributes of the areas become more alike (Honkola et al. 2018). Instead, Swiss German and Danish dialectal patterns follow reasonable assumptions: Swiss German dialects followed the wave model (Romano et al. 2022), and similarity of Danish dialectal variants were predicted especially with location in the same island/peninsula (Wichmann, submitted).

Holistic endeavors focusing on studies of human diversity can today use spatially rich data to **cross-analyse** linguistic, social, genetic, cultural or environmental data to reveal parallel patterns or causal mechanisms behind them. For example, Lynch et al. (2022) reported that one of the drivers of WW2 Karelian evacuees for staying in Finland was the similarity of dialect. Also Pichkar and Creanza (2023) had sociolinguistic take in their study: they found co-location of dialectal and genetic areas in England and discussed the possibility that linguistic variation would have affected population contacts through mate choice. We predict that digital dialectal datasets with fine-scale spatial variation will allow new ways in using linguistic data for studies of human diversity and history. An example of novel use of spatial linguistic variation is Nitch et al. (2025), who used dialectal areas for modeling spatial patterns in mortality to infectious diseases in 19th century Finland, exploring whether disease spread could be better modelled using dialectal areas or administrative regions as proxies of human population networks (as was done often with prediction of Covid19 spread).

The workshop aims at presenting the current trends in using fine-scale spatial linguistic data in linguistic and interdisciplinary settings. It welcomes contributions within the themes including but not restricted to: Resolving dialectal landscapes and drivers of the observed areal patterns; Showcasing cross-analyses of dialectal variation with extra-linguistic factors; Microtypological or lexical variation within a language or between sister languages; Geospatial analyses applied to solve linguistic variation; Human diversity studies in relation to dialectal landscapes.

References

Gilliéron, J. (1902-10): Atlas linguistique de la France. 13 vol. Paris: Champion.

Honkola, T., Ruokolainen, K., Syrjänen, K. J. J., Leino U, Tammi, I, Wahlberg, N. & Vesakoski, O. (2018): Evolution within a language: Environmental differences contribute to divergence of dialect groups. BMC Evolutionary Biology. 18:132.

Kettunen, L. (1940): Suomen Murteet III A. Murrekartasto. Helsinki: Suomalaisen Kirjallisuuden Seura.

Leemann, A., Kolly, M.-J., and Britain, D. (2018): The English Dialects App: The creation of a crowdsourced dialect corpus. Ampersand, 5, 1-17.

Lynch, R., Loehr, J., Lummaa, V., Honkola, T., Pettay, J. and Vesakoski, O. (2022): Socio cultural similarity with host population rather than ecological similarity predicts success and failure of human migrations. Proceedings of the Royal Society B, 289: 20212298.

Nerbonne, J. and Kretzschmar, W. (2003): Introducing Computational Techniques in Dialectometry. Computers and the Humanities 37, 245–255.

Nitsch A, Lummaa V, Ketola T, Honkola T, Vesakoski O, Briga M. (2025): The spatial distribution of pertussis, but not measles or smallpox, in pre-industrial Finland matches dialects. iScience 28(6), 112530. DOI: 10.1016/j.isci.2025.112530.

Norvik, M., Balodis, U., Metslang, H., Pajusalu, K., Saar, E., Teras, P. & Vaab, J. (2026): Continuity and change of structural features: The example of southern Finnic. In: Proceedings of SLE 2026, pp. 656-657
Pichkar, Y., & Creanza, N. (2023): Fine-scale cultural variation reinforces genetic structure in England. American journal of biological anthropology, 181(4), 626–636.
Romano, N., RAnacher, P., BAchmann, S and Joost, S. (2022): Linguistic traits as heritable units? Spatial Bayesian clustering reveals Swiss German dialect regions. J. of Linguistic Geography 10:11-22.
Saareste, A. (1938): Eesti murdeatlas I vihk. Tartu: Eesti Kirjanduse Selts.
Santaharju, J., Syrjänen, K., Honkola, T. Seppä, P., Vesakoski, O. and Leino, U. (2025): Data release: Digitized Dialect Atlas of Finnish by Lauri Kettunen. Digital Humanities in the Nordic and Baltic Countries Publications 7 (3).
Wichmann, Søren, Darja Jonjić and Matías Guzmán Naranjo (2026): Foundations of Russian dialectometry. Russkij jazyk v nauchnom osveshchenii – Russian Language and Linguistic Theory 51(1): 9-52. doi: 10.31912/rjano-2026.1.1.
Wichmann, Søren. Submitted. Elements of a Danish dialectometry

Links

Romansh idioms: https://www.atlas-rumantsch.ch/?idiom=de
Dialektkarten user interface: https://dialektkarten.ch/index.en.html
Estonian microtypology: https://esttyp.ut.ee/
Diatech user interface: http://eudia.ehu.eus/diatech/index (Aurrekoetxea, G., Santander, G., Usobiaga, I. and Iglesias, A. (2016): Diatech: Tool for making dialectometry easier. Dialectologia 17: 1-22.
Gabmap user interface (Gapmap.nl; currently no direct access): Nerbonne, J., Colen, R., Gooskens, C., Kleiweg, P. and Leinonen, T. (2011): Gapmap - A web application of dialectology. Dialectologia Special Issue II: 65-89