The `toponym` R package: A new tool for visualizing and analyzing toponymic distributions

Lennart Chevallier & Søren Wichmann
Kiel University

Keywords: onomastics; geolinguistics; software; Xincan; Slavic

Toponomastics is a well-established branch of humanities which has a history of more than two centuries of practice based on sound, scientific principles and which is of interdisciplinary importance (Eichler 1982: 8). It is useful to distinguish between what we may call particularistic and pattern-finding approaches: The former include a focus on the etymologies of individual place names and what they can tell us about histories of ethnic groups. The latter are mainly concerned with the identification of morphemes recurring across numerous place names and what the distributions found may reveal. Since around 1990, the personal computer began to play a decisive role in some toponymic studies of the pattern-finding kind (e.g., Campbell 1991), and the availability of GIS software have made such studies more common. Nevertheless, the particularistic approaches are still vastly dominant in the literature on toponyms (cf. Tent 2015), and it is symptomatic that the recent textbook by Cacciafoco and Cavallaro (2023) is almost devoid of distributional maps. We believe, however, that the possibility to process and map large amounts of place names using dedicated software has a great potential to boost toponomastics. For this reason, we have developed the package `toponym` within the computational environment of R, an environment that enjoys a wide popularity. R allows for the integration of countless tools from statistics, linguistics, geography, etc. The package provides an interface to the data at https://www.geonames.org/, which contains around 12.5 million entries for toponyms pertaining to 253 countries and 9 feature classes (populated places, water bodies, mountains, etc.). Our software facilitates the search for specific strings (sequences of letters) in the entire database and instantly produces maps of toponyms filtered in a number of different possible ways. One among several functions provided is specialized in finding and exploring toponyms that are characteristic of a certain region as opposed to the rest of a particular country. Users may specify an individual region by clicking on a map or directly select it from a list of official administrative units. We demonstrate the functionalities of the package and illustrate its scientific utility through two case studies, one involving the small Xincan linguistic family of southeastern Guatemala, and one involving varieties of Slavic languages formerly spoken in eastern Germany.

References

Campbell, Jon C. (1991), Stream Generic Terms as Indicators of Historical Settlement Patterns, *Names* 39(4), 333–366.

Eichler, Ernst (1982), *Ergebnisse der Namenforschung im deutsch-slawischen Berührungsgebiet*, Berlin: Akademie-Verlag.

Tent, Jan (2015), Approaches to Research in Toponymy, *Names* 63(2), 65–74.

Perono Cacciafoco, Francesco and Francesco Cavallaro (2023), Place Names: Approaches and Perspectives in Toponymy and Toponomastics, Cambridge: Cambridge University Press.