



HAL
open science

Toponymic representation of spaces within the city through real estate ads

Alicia Bianchi, Giovanni Fusco, Karine Emsellem

► **To cite this version:**

Alicia Bianchi, Giovanni Fusco, Karine Emsellem. Toponymic representation of spaces within the city through real estate ads. 23rd European Colloquium on Theoretical and Quantitative Geography (ECTQG 2023), Universidade de Coimbra, Sep 2023, Braga, Portugal. hal-04329864

HAL Id: hal-04329864

<https://hal.science/hal-04329864>

Submitted on 7 Dec 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Toponymic representation of spaces within the city through real estate ads

Alicia BLANCHI¹; Giovanni FUSCO²; Karine EMSELLEM³

¹Université Côte d'Azur, CNRS, UMR ESPACE, KCityLabs, France, alicia.blanchi@kcitylabs.fr

²Université Côte d'Azur, CNRS, UMR ESPACE, France, giovanni.fusco@univ-cotedazur.fr

³Université Côte d'Azur, CNRS, UMR ESPACE, France, karine.emsellem@univ-cotedazur.fr

Keywords: Spatial analysis, real estate ads, toponym, spatial representation, fuzzy geographical location

Toponyms mark, segment and organise the urban space since their creation. They evolve with time and space, reflecting the historical, social, and cultural evolution of cities, including in terms of urban planning (Le Squère, 2006; Jordan, 2020; Giraut & al. 2016). They are perceived and identified by populations, thus contributing to the social representation of the different subspaces that compose the city. Toponyms respond to both spatial and social dimensions, making them difficult to identify. Their location and characterization present blurred and uncertain boundaries (Bunel, 2021). Several studies have dealt with the delimitation of fuzzy socio-spatial phenomena (Hu & al. 2019, Vasardani & al. 2013).

In this direction, we propose to reconstruct and characterise the toponymic landscape of the city from geolocated real estate ads, using an appropriate spatial analysis of the pieces of information on the urban space present in their text.

A geographic database was used, containing tens of thousands of pieces of geographic information extracted from real estate ads texts of the French Riviera through a combination of natural language processing (NLP) and artificial intelligence (Cadorel & al., 2021). The geographical information extracted, cleaned, and structured, is based on the places and their socio-spatial context, made of spatial relationships and associated attributes, recognizable in expressions like "*in the residential neighbourhood of Cimiez*".

The naming and delimitation of places based on a multitude of extracted information are complex processes, leading to blurred toponyms with unclear boundaries, even if spatial barriers can sometimes segment their use. To model the significance of toponyms in space, we have developed a protocol based on spatial analysis of geographical information extracted and geolocated with decametric precision on the street network. The street network is a fundamental feature for the perception and cognition of urban space (Lynch 1960), and we can assume that it can be a pertinent basis for its segmentation through toponymy.

Geolocated ads can be seen as a spatial sampling of toponym use. We thus interpolated toponym use with a kernel density-based approach on the street network with adaptive parameters for each toponym (Gelb, 2021; Lachance-Bernard & al., 2011). The toponyms are those present in the

database of extracted geographical information associated with a relationship of spatial situation (Blanchi & al., 2022).

From there, we extracted the relevant phenomenon (use of a given toponym) by imposing a minimal threshold and some spatial coherence in toponym use. Artefacts linked to erroneous geolocations or low presence of ads (a known bias introduced by the real estate market) could thus be reduced. Nevertheless, the presence of toponym use in space is not sufficiently indicative of its representation within the urban toponymic landscape. We therefore also assessed its absence within the same space, using the same interpolation method of all geolocated real estate ads that do not mention the toponym through a notion of situation.

Finally, we assessed the strength of a toponym in space using the concepts of core-space and support-space, derived from fuzzy set theory. First, the ratio between the kernel densities of presence and absence of toponym use on each street segment is used to model graduated presence of the phenomenon. Then, using the equivalent of an alpha-cut approach, the core was defined as the space where toponym use was predominant, and the support as the space of extended but more contested use. More precisely, morphological closures and connection checks were also carried out to impose spatial coherence on core- and support-spaces. Their mapping reveals how place-name usage produces meaningful patterns in intra-urban space, independently of administrative constraints.

However, some subspaces can be shared by different place names. Within the context of real-estate valorisation, some toponyms can be used more widely beyond their core. Supports overlap cores in an asymmetric way, showing the different power exerted by toponyms in the social representation of space. A local hierarchy among neighbouring toponyms can thus be calculated and mapped, producing new insight in the social meaning of the identified spatial patterns.

It was also possible to further characterise toponyms through their associated geographic entities. Entities like “*neighbourhood*”, “*sector*”, “*street*”, “*hill*” have different meanings in social perception and cognition. Finally, a correspondence analysis was carried out using toponyms, associated attributes, and localisation factors (features of space put forward by real estate ads, including proximity to interest points, urban services and environmental amenities and view factors). The interpretation of the factorial axes allowed for the identification of main oppositions within the toponymic landscapes of the cities of the French Riviera. In Nice, for example, a first major opposition is between what are perceived as toponyms of residential places on the hills surrounding the city, and two different subspaces within the latter: the historic and tourist-oriented city and the more utilitarian city signified by the presence of urban facilities and services.

In conclusion, the spatial analysis of toponyms and their relationships in space allowed us to represent the patterns of toponymic landscape as they are perceived, appreciated, and identified by people. However, it is important to note that our data comes from real estate ads, which are punctual data influenced by the dynamics of the real estate market and real estate valorisation.

References

A Blanchi, A., Fusco, G., Emsellem, K., Cadorel L. (2022). ‘Studying urban space from textual data: Toward a methodological protocol to extract geographic knowledge from real estate ads’. In O. Gervasi; B. Murgante; S. Misra; A.M.A.C. Rocha; C. Garau. (eds.) *Computational Science and Its*

Applications – ICCSA 2022 Workshops. Proceedings Part II, 13378, Springer, pp.520-537, 2022, Lecture Notes in Computer Science.

Bunel, M. (2021). ‘Un état de l’art sur l’imprécision spatiale et sa modélisation’, *Cybergeo: European Journal of Geography*, Cartography, Images, GIS, document 966.

Cadorel, L., Bianchi, A., Tettamanzi G. B., A. (2021). ‘Geospatial Knowledge in Housing advertisements: Capturing and Extracting Spatial Information from Text’. In *Proceedings of the 11th Knowledge Capture Conference (K-CAP '21)*, December 2–3, 2021, Virtual Event, USA. ACM, New York, NY, USA, 8 pages.

Gelb, J. (2021). ‘spNetwork, a package for network kernel density estimation’. *The R Journal*.

Giraut F., Houssay-Holzschuch M., (2016). ‘Place Naming as Dispositif: Toward a Theoretical Framework’. *Geopolitics*, vol. 21, n° 1, p. 1-21.

Hu Y., Mao H., McKenzie G. (2019). ‘A natural language processing and geospatial clustering framework for harvesting local place names from geotagged housing advertisements’, *International Journal of Geographical Information Science*, 33:4, 714-738.

Jordan, P. (2020). ‘Role of Place Names in Relating People and Space’. *Handbook of the Changing World Language Map*. Springer, Cham.

Lachance-Bernard, N., Produit, T., Tominc, B., Nikšič, M., Goličnik Marušić, B. (2011). ‘Network based Kernel Density Estimation for Cycling Facilities Optimal Location Applied to Ljubljana’. In: Murgante, B., Gervasi, O., Iglesias, A., Taniar, D., Apduhan, B.O. (eds) *Computational Science and Its Applications - ICCSA 2011. ICCSA 2011*. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg.

Le Squère, R. (2006). ‘Analyse des perceptions, usages et fonctions des toponymes actuels des territoires ruraux et urbains de Bretagne’, *Cahiers de sociolinguistique*, vol. 11, no. 1, pp. 81- 99.

Lynch, K. (1960). *The image of the city*. MIT Press.

Vasardani, M., Winter, S., Richter K-F. (2013). ‘Locating place names from place descriptions’. *International Journal of Geographical Information Science*. 27.