

# Spatial Tweetstistics with R: Geographical Distribution of English Loan Words in Spanish Tweets

Irene Checa-Garcia <sup>1\*</sup>, Virgilio Gómez-Rubio <sup>2</sup>

1. University of Wyoming

2. Universidad de Castilla-La Mancha

\*Contact author: [irene.checa@uwyo.edu](mailto:irene.checa@uwyo.edu)

**Keywords:** Loan Words, Dialectology, Anglicisms, Spatial Statistics, Twitter

In this study we look at the geographical and frequency distribution of different types of Anglicisms in tweets in Spanish. In the first part of the study we distinguish different types of Anglicisms according to several criteria: their degree of official acceptance into the language, their structural characteristics, their extension, and the availability of alternative expressions in the Spanish language for the same concept. One recurrent question in the literature about loan words, particularly in language contact situations, such as Spanish in the USA, is how frequent loan words really are. Does a language contact situation heavily influence –and if so, how heavily– the frequency of use of loan words from one language to another? A related question is which types of loan words are more frequent [1]. Previous works have contrasted loan word presence in bilingual and monolingual speakers [2] or offered a percentage of their frequency in different corpora [3], but no further statistical analysis was presented. In this work, we won't look at who is producing loan words, but rather where they are being produced.

Using a spatial data analysis [4] we will study the geographical frequency of a set of over 250 Anglicisms. We have used *R* packages **streamR** for the data collection and different packages from the *Spatial Task View* for the visualization and spatial data analysis. We have produced maps to summarize the main results and to show the spatial distribution of Anglicisms types that we have found in tweets in different Spanish speaking countries. In this way, we will show the “hottest spots” for the use of the different types of Anglicisms. And we hope to answer geographically the question of how language contact might influence loan word frequency.

## References

- [1] Medieta, Eva (1999) *El préstamo en el español de los Estados Unidos*. New York: Peter Lang.
- [2] Otheguy, Ricardo & García, Ofelia (1988) Diffusion of lexical innovations in the Spanish of Cuban Americans. In Jacob Ornstein-Galicia & George K. Green (eds.) *Research Issues and Problems in United States Spanish: Latin American and Southwestern Varieties*.
- [3] Silva-Corvalán, Carmen (1994) *Language contact and change. Spanish in Los Angeles*. Oxford: Oxford University Press.
- [4] Bivand, R., Pebesma, E. and Gómez-Rubio, V. (2013). *Applied Spatial Data Analysis with R*. Springer, New York.