

Dialect Variation as Geography

Sheila Embleton, Dorin Uritescu† and Eric Wheeler

embleton@yorku.ca; eric.wheeler@sympatico.ca

York University, Toronto, CANADA

It is well known that dialect variation correlates with many factors, including the sociolinguistic ones of age, gender and social class (e.g. Chambers 2009). But the obvious (and one might say, traditional) explanation for dialect variation is geography: the further apart people are, the more their languages can vary (for example, cf. Nerbonne, et al. 2007). How true is this?

Theory-driven thinking would suggest that the more difficult it is for two groups to communicate, the more possibility there is for variation between the groups (cf. Wheeler 2007). That suggests that distance between communities, both direct and as a function of travel time or travel distance would correlate with dialect variation. A first test of this idea, using Romanian data (Embleton, Uritescu and Wheeler 2012, 2017), demonstrated a correlation with r-squared of about 0.80. There were some small improvements when travel distance or travel time was substituted for direct distance “as the crow flies”. Clearly, in this case, geography is an important but not exclusive correlate of dialect variation. The question arises of whether or not the scale of distance is important. The Romanian area examined is about 250 km long, whereas the distance from Beijing to Hong Kong is about 2000 km. It is to be expected that there is “dialect variation” between the two Chinese locations, but does it accord with the distance?

At the other extreme, we have data from the Mambila region of Nigeria and Cameroon, where the distances are well under 100 km, but perhaps the cultural and social distances are stronger, and they override the simple geographic factors.

Using consistent approaches to the measures of geography and dialect variation, we report the quantitative correlation in these three situations (Northwest Romania, Chinese, Mambila), and discuss some of the issues that impact the answers. While positive results in any study are always welcome, other results can illustrate the limitations of a valid hypothesis. Yes, geography does correlate with dialect variation, but not always in the same degree, and for good reasons, such as the appropriate choice of linguistic data, varying geographic factors, and over-riding social conditions.

References

- Chambers, J.K. 2009. *Sociolinguistic Theory: Linguistic Variation and its Social Significance*, 3rd ed, Oxford UK and Cambridge US: Blackwell. First ed. 1995, second ed. 2003.
- Embleton, Sheila, Dorin Uritescu and Eric S. Wheeler. 2012. Effective Comparisons of Geographic and Linguistic Distances. In Gabriel Altman, Peter Grzybek, Sven Naumann, and Relja Vulcanović (eds.). *Synergetic Linguistics. Text and Language as Dynamic Systems*. Vienna: Praesens Verlag. pp. 225-232
- Embleton, Sheila, Dorin Uritescu and Eric S. Wheeler. 2017. An Expanded Quantitative Study of Linguistic vss Geographic Distance Using Romanian Dialect Data. Proceedings of Qualico 2016. In Lu Wang, Reinhard Köhler, Arjuna Tuzzi (eds.), *Structure, Function and Process in Texts*, Lüdenscheid, Germany: RAM-Verlag, pp. 25-33.
- Nerbonne, John, Peter Kleiweg, Franz Manni and Wilbert Heeringa. 2007. Projecting Dialect Distances to Geography: Bootstrap Clustering vs. Noisy Clustering. <http://www.let.rug.nl/~nerbonne/papers/Nerbonne-Kleiweg-Manni.pdf>
- Wheeler, Eric S. 2007. Language Change in a Communication Network. pp 689-698 in Peter Grzybek, Reinhard Köhler (editors). *Exact Methods in the Study of Language and Text*. Berlin and New York: Mouton de Gruyter.