Chapter 24

Gravity, radiation, and dialectometry

Robert Malouf
San Diego State University

Gravity laws have a long history in models of human mobility. Trudgill (1974) proposed a version of the gravity model to account for the diffusion of dialect variation, but other researchers found gravity to be a poor predictor of linguistic change. In a large-scale evaluation of the gravity law across a whole dialect region, Nerbonne & Heeringa (2007) find support for some but not all of its predictions. This contribution compares the gravity model to an alternative proposed by Simini et al. (2012) which is inspired not by gravity but by the physics of radiation and absorption. In computational simulations, the radiation model appears to be a better match than the gravity model for the general patterns found in dialect diffusion. This suggests as a future direction for dialectometric work the comparison of quantitative models of human mobility and their consequences for understanding linguistic variation.

1 Introduction

Gravity models have a long history in geography and economics as a model of human mobility (e.g., Ravenstein 1885; Zipf 1946). Taking a form familiar from Newton's law of universal gravitation, gravity models in human geography predict that the degree of interaction between two locations i and j follows the rule:

$$I_{ij} = C\,\frac{m_i\,m_j}{d_{ij}^2} \tag{1}$$

where m_i and m_j are the populations of i and j, d is the distance between them, and C is a constant. This makes intuitive sense in a general way: the opportunities for connections between locations increase as either of their populations increase, but decrease rapidly with distance.

Gravity laws have the appeal of being both conceptually and mathematically simple (though in its general form, the gravity law has a number of parameters that can be difficult to estimate). It also makes the correct empirical predictions (at least approximately) for a broad range of phenomena, including communting patterns, migrations, airline traffic, and commodity flows. More recently, gravity laws have also been applied to modeling mobile phone calls (Krings et al. 2009) and social media use (Noulas et al. 2012).

If social contact via communication, migration, and trade is the engine that drives the diffusion of dialect variation, then it stands to reason that dialect differences might reflect patterns of interaction predicted by the gravity law. Bloomfield (1933) offered the insight that dialect differences reflect differences in the 'density of communication'. Trudgill (1974) expanded on this by applying a gravity law as a measure of mutual influence between cities in modeling the diffusion of London dialect features into East Anglian English.

Since that introduction into dialect studies, the gravity model has received a number of challenges and modifications. Boberg (2000), e.g., looks at diffusion of two linguistic variants across the US/Canada border and finds that (1) is a poor fit. In this and a number of other case studies, the gravity model finds mixed support and compares unfavorably with models that take social prestige and identity into account.

2 Gravity in dialectometry

In their study of Dutch dialects, Nerbonne and colleagues developed and extended a set of dialectometric methods for large-scale measurement of linguistic variation across space and time (Nerbonne, Heeringa & Kleiweg 1999; Heeringa 2004; Heeringa et al. 2006; Nerbonne 2009; Nerbonne & Heeringa 2010). By computing edit distance between comparable forms collected at a large number of locations, they are able to quantify the degree and nature of dialect variation across a region.

Nerbonne et al.'s metrics aggregate many dimensions of linguistic variation into a single score, which makes them ideal for evaluating the predictions of the gravity model. This is in contrast to studies like Trudgull's and Boberg's, which focus on one or two hand-selected linguistic variables.

Nerbonne et al. tested how well the gravity law fits against real-world (Nerbonne & Heeringa 2007; Heeringa et al. 2007) and simulated (Nerbonne 2010) dialect areas and made two findings. First, they determined that, consistent with the gravity model, geographic distance is the most important factor in predicting dialect difference. It seems that, when taken in the aggregate, linguistic variation is not as dependent on purely social factors as the post-Trudgill work might have suggested. They do find that the relationship with distance is linear rather than quadratic. However, this is consistent with the generalized form of the gravity law which is often used in studies of human mobility.

More surprisingly, though, they conclude that the influence of population on dialect difference is very small and, if anything, points weakly in the wrong direction. Sites with larger populations are slightly more likely to differ than sites with smaller populations, contrary to (1). This would seem to be straightforward falsification of the gravity model for dialect variation, but it is also puzzling. The gravity model has been successfully used to describe many different phenomena relating to social and economic connections between locations. Why should linguistic diffusion work differently?

3 Radiation model

A possible answer to that question comes from the observation that gravity models have come under challenge in other social domains as well. Responding to the numerous theoretical and empirical shortcomings of the gravity model, Simini et al. (2012) propose an alternative, inspired not by gravity but by the physics of radiation and absorption. They start with a model of people commuting to work and a simple assumption: commuters will search for a job as close to home as possible, and will choose the closest job which provides a benefit that is greater than that which is available in their own city. From this, they derive the fundamental equation of the radiation law, which gives the probability p_{ij} of a commuter traveling from city i to city j:

$$p_{ij} = \frac{m_i\, m_j}{(m_i + s_{ij})(m_i + m_j + s_{ij})} \tag{2}$$

Here, m_i and m_j are the populations of i and j, and s_{ij} is defined as follows: if r_{ij} is the distance between city i and city j, then s_{ij} is the total population (not including the populations of i and j) living within a circle of radius r_{ij} centered on i.

One notable property of the radiation law is that it p_{ij} is not directly dependent on distance. It depends instead on s_{ij}, the population within a certain distance. If the population density is uniform, then s_{ij} is proportional to the area of a circle of radius r_{ij} and so proportional to $m_i\, d_{ij}^2$. In general, though, the relationship between d and s depends entirely on the way the population is distributed. This allowing the radiation model to, e.g., model differences in commuting patterns in rural and urban areas, something that the gravity model had not been able to give a general account of.

Simini et al. (2012) go on to validate their radiation law empirically, showing that is a better predictor of human mobility than the gravity law based on a broad range of empirical tests, concluding: "We find that the radiation model offers an accurate quantitative description of mobility and transport spanning a wide range of time scales (hourly mobility, daily commuting, yearly migrations), capturing diverse processes (commuting, intra-day mobility, call patterns, trade), collected via a wide range of tools (census, mobile phones, tax documents) on different continents (America, Europe)" (p. 97).

While the radiation model is not without critics (e.g., Masucci et al. 2013), it has held up as a superior predictor of inter-city flows of people, goods, and communication than the older gravity model. Perhaps it would be a better model of the inter-city flow of linguistic variation as well.

4 Simulations

Following Nerbonne (2010), I performed simple computational simulations to compare the predictions of the gravity and radiation models for linguistic diffusion.[1] For

---

[1] The simulations were implemented in Python and are available at `http://github.com/rmalouf/gravity`.

each of these experiments, an artificial distribution of cities was created. The first step was to create cities of varying populations of agents. The population was seeded with 500 agents, each in its own city. Next, additional agents were added one at a time by selecting an existing agent and adding the new agent to the selected agent's city until the total population of agents reaches 100,000. This is essentially Albert & Barabási's (2002) preferential attachment algorithm, and it produces a collection of cities whose populations are distributed following an inverse power law.

The next step is to locate these cities in space. The city with the largest population is taken as the 'capital' (the source of linguistic innovations) and placed at the center. The remaining cities are then distributed uniformly within a unit circle centered on the capital.

Agents are represented as a vector of 100 binary values representing linguistic variables (Holman et al. 2007; Nerbonne 2010). Agents in the capital have the value 1 (representing the innovative value) for all variables while all other agents have the value 0. Successive iterations of the simulation model the interaction between a traveler from the capital and an resident of one of the other cities. On each iteration, a city other than the capital is selected as a destination and then an agent is selected in that city. On contact with a resident of the capital, each of the selected agent's linguistic variables have a 1% chance of being set to 1 (note than some may have already been set to 1 on a previous iteration). After 50,000 iterations, the aggregate linguistic difference between the dialects spoken in each city is calculated. A city's 'dialect' is the normalized sum of the vectors of the agents living in it, and the difference between two dialects is 1 minus the dot product of the cities' dialect vectors. This ranges from 0.0 for dialects which share all properties to 1.0 for dialects that share none.

In the first simulation, destination cities were selected for agents with a probability proportional to a modified gravity law:

$$I_{ij} = \frac{m_i \, m_j}{d_{ij}} \tag{3}$$

This preserves the dependency on population, but following Nerbonne & Heeringa (2007) replaces the quadratic distance term with a linear one. In the second simulation, the probability that a destination city would be selected is given by the radiation model (1).

Figures 1 and 2 give the results of the simulation. The plots show the relationship, for all of the cities except the capital, between either geographic distance from the capital or population and the difference between the local dialect and the capital's dialect.

As Figure 1 indicates, the radiation model predicts that linguistic difference should increase with geographic distance, which is consistent with Nerbonne & Heeringa's (2007) findings. The predictions for the gravity model are not quite as clear, but while the trend is not very strong there is an increase in difference as distance increases. For both models, the curves are reminiscent in their general shape of what Nerbonne (2010) calls Séguy's Law: linguistic differences increase with distance to a point beyond which further increases in distance correlate with only small increases

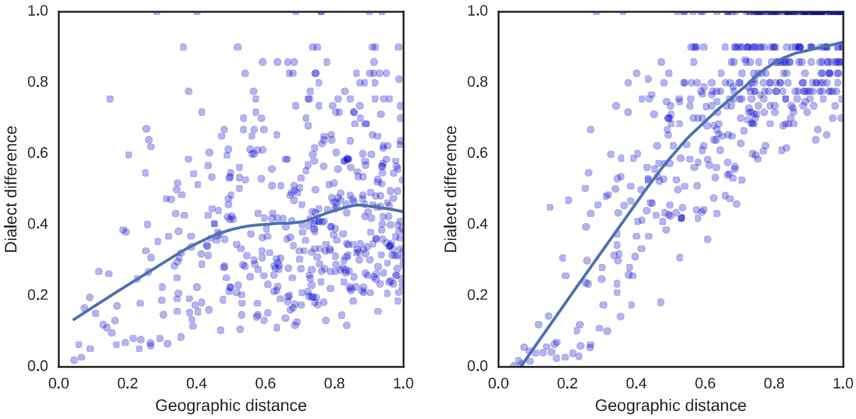

Figure 1: Geographic distance from capital vs. dialect difference, gravity (left) and radiation (right) models.

in linguistic difference. Séguy's Law leads to a sublinear increase in the dialect differences as distance increases, something both the gravity and the radiation models appear to predict.

Figure 2 shows the results for population. Based on the simulation, we can see that

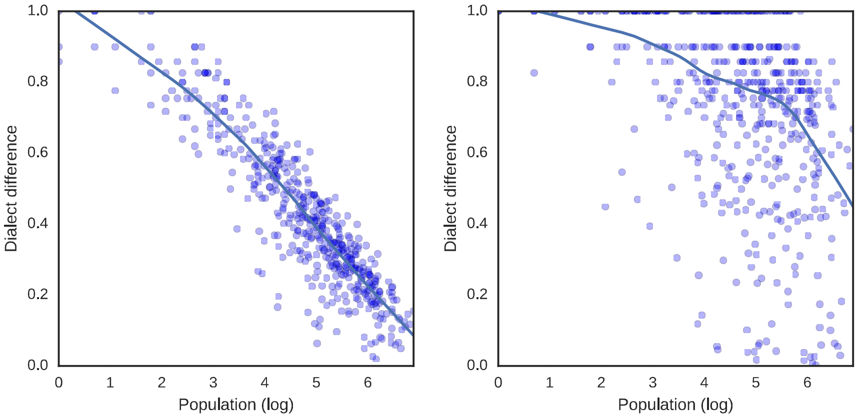

Figure 2: Population vs. dialect difference, gravity (left) and radiation (right) models.

the gravity model as specified in (1) predicts a strong inverse relationship between population size and linguistic difference. This is exactly what Nerbonne & Heeringa (2007) did **not** find in their study of Dutch dialects. For the radiation model, on the other hand, the relationship is very weak. The local regression line (fit by LOWESS) indicates a bit of a downward trend. But, many of the cities with the largest populations also have the greatest linguistic distance from the capital's dialect.

5 Conclusions

What these results show is that at least at an impressionistic level the radiation model is a better match for aggregate patterns seen in dialect variation than the gravity model. The next step would be to go beyond simulation results and to evaluate the radiation model against the kind of dialect variation data that Nerbonne & Heeringa (2007) used to test the gravity law. The radiation model has succeeded in other areas of human mobility where the gravity model has not. If it makes the correct predictions for the diffusion of linguistic variation as well, this would add confirmation to Bloomfield's (1933) hypothesis concerning the density of communication. But, if the radiation model does not fare any better than the gravity law, that would raise profound questions about the nature of the mechanisms underlying dialect diffusion and their relation to other social processes.

More generally, we can say that without dialectometric methods these questions could not even be asked. As a complement to detailed description of ongoing linguistic changes in their social context, large-scale aggregate comparison of dialects over an entire region allow us to relate language variation and change to other aspects of human dynamics.

References

Albert, Réka & Albert-László Barabási. 2002. Statistical mechanics of complex networks. *Reviews of Modern Physics* 74(1). 47.

Bloomfield, Leonard. 1933. *Language*. New York: Henry Holt & Company.

Boberg, Charles. 2000. Geolinguistic diffusion and the US-Canada border. *Language variation and change* 12(01). 1–24.

Heeringa, Wilbert. 2004. *Measuring dialect pronunciation differences using Levenshtein distance*. Rijksuniversiteit Groningen PhD thesis.

Heeringa, Wilbert, Peter Kleiweg, Charlotte Gooskens & John Nerbonne. 2006. Evaluation of string distance algorithms for dialectology. In John Nerbonne & Erhard Hinrich (eds.), *Proceedings of the Workshop on Linguistic Distances*, 51–62.

Heeringa, Wilbert, John Nerbonne, Renée van Bezooijen & Marco René Spruit. 2007. Geografie en inwoneraantallen als verklarende factoren voor variatie in het nederlandse dialectgebied. *Nederlandse Taal- en Letterkunde* 123(1). 70–82.

Holman, Eric W., Christian Schulze, Dietrich Stauffer & Søren Wichmann. 2007. On the relation between structural diversity and geographical distance among languages: observations and computer simulations. *Linguistic Typology* 11. 393–421.

Krings, Gautier, Francesco Calabrese, Carlo Ratti & Vincent D. Blondel. 2009. Urban gravity: a model for inter-city telecommunication flows. *Journal of Statistical Mechanics: Theory and Experiment* 2009(07). L07003.

Masucci, A. Paolo, Joan Serras, Anders Johansson & Michael Batty. 2013. Gravity versus radiation models: on the importance of scale and heterogeneity in commuting flows. *Physical Review E* 88(2). 022812.

Nerbonne, John. 2009. Data-driven dialectology. *Language and Linguistics Compass* 3(1). 175–198.

Nerbonne, John. 2010. Measuring the diffusion of linguistic change. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365(1559). 3821–3828.

Nerbonne, John & Wilbert Heeringa. 2007. Geographic distributions of linguistic variation reflect dynamics of differentiation. In Sam Featherston & Wolfgang Sternefeld (eds.), *Roots: linguistics in search of its evidential base*, 267–298. Berlin: Mouton de Gruyter.

Nerbonne, John & Wilbert Heeringa. 2010. Measuring dialect differences. In Jürgen Erich Schmidt & Joachim Herrgen (eds.), *Language and space: theories and methods*, 550–566. Berlin: Mouton de Gruyter.

Nerbonne, John, Wilbert Heeringa & Peter Kleiweg. 1999. Edit distance and dialect proximity. In David Sankoff & Joseph Kruskal (eds.), *Time warps, string edits and macromolecules: the theory and practice of sequence comparison.* CSLI Publications.

Noulas, Anastasios, Salvatore Scellato, Renaud Lambiotte, Massimiliano Pontil & Cecilia Mascolo. 2012. A tale of many cities: universal patterns in human urban mobility. *PloS one* 7(5). e37027.

Ravenstein, E. G. 1885. The laws of migration. *Journal of the Statistical Society of London* 48(2). 167–235.

Simini, Filippo, Marta C. González, Amos Maritan & Albert-László Barabási. 2012. A universal model for mobility and migration patterns. *Nature* 484(7392). 96–100.

Trudgill, Peter. 1974. Linguistic change and diffusion: description and explanation in sociolinguistic dialect geography. *Language in Society* 3(2). 215–246.

Zipf, George K. 1946. The P_1P_2/D hypothesis: on the intercity movement of persons. *American Sociological Review* 11(6). 677–686.