



Applications and Explanations in Dialectometry

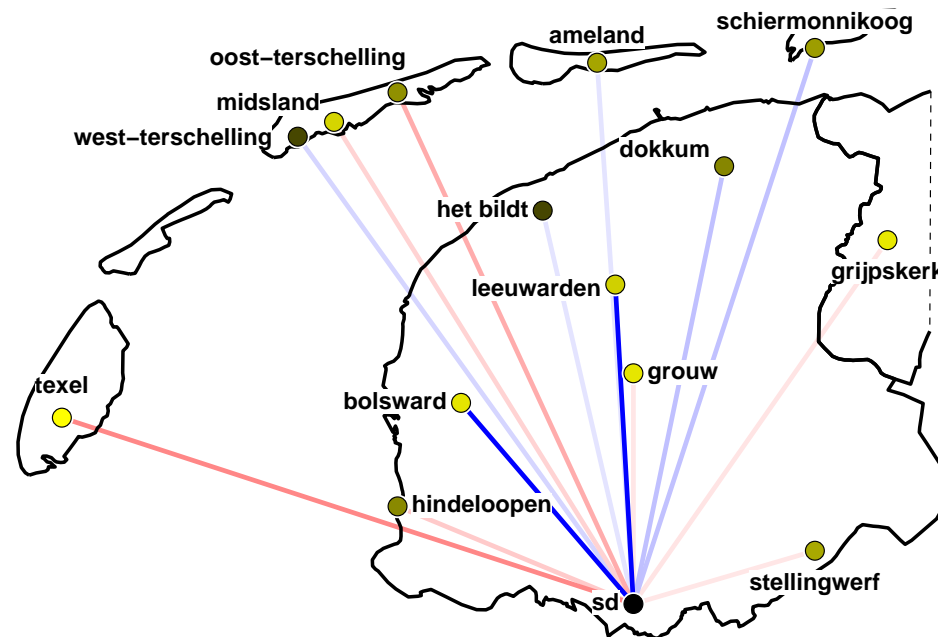
John Nerbonne
University of Groningen, Humanities Computing
j.nerbonne@rug.nl

LSA Linguistics Institute
Harvard and MIT, Summer 2005
<http://www.let.rug.nl/nerbonne/teach/dialectology/>



Applications

Measuring convergence and divergence of varieties (Heeringa and Nerbonne, 2000):
Town Frisian is converging toward standard Dutch (sd)

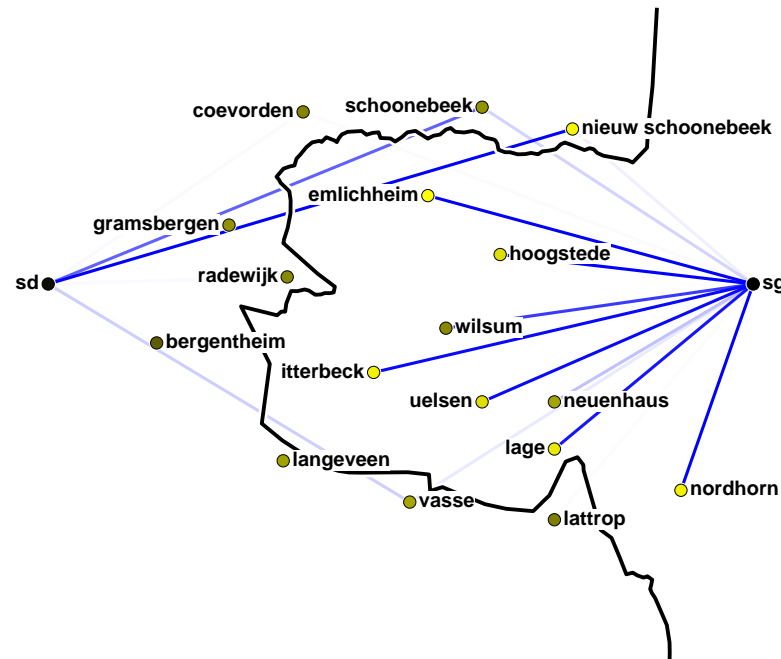


Blue indicates convergence, red divergence. Yellow increases with change.



Applications, cont.

Measuring effect of borders, standard languages (Heeringa et al. 2000) Saxon is diverging at Dutch-German border in Bentheim, 1974-2000



Blue indicates convergence toward standard Dutch (sd) vs. standard German (sg).



New Questions

Now that we can measure linguistic distances reliably, can we ask the fundamental question more satisfactorily?

How does geography influence linguistic variation?



Geography and Linguistics

- Part of larger assembly of questions on geography and culture
- How has geography influenced the spread of culture?
- What does the pattern of culture reveal about cultural dynamics?



Trudgill's Gravitational View

Peter Trudgill suggests that language varieties may be subject to a “gravity law”, being attracted to one another in a way like the way planets are attracted to the sun.

$$F = G \frac{m_1 m_2}{r^2}$$

F is the force due to gravity,

m_1, m_2 the masses of the two objects attracting each other,

r the distance between them, and

G is a “universal gravitational constant.”



Linguistic Cohesion via Gravity

$$F = G \frac{m_1 m_2}{r^2} = G \frac{p_1 p_2}{r^2}$$

F is the attractive force,

m_1, m_2 the populations (p_1, p_2) of the two settlements,

r the distance between them, and

G won't be speculated on

Idea: social contact promotes linguistic accommodation and linguistic similarity.



Motivating Linguistic Cohesion via Gravity

Chance of social contact should be

- proportional to the product of settlement size and
- (if travel is random) inversely proportional to squared distance

Notate bene: we measure linguistic dissimilarity, which we postulate stands in inverse relation to the attractive force of social contact.



Predictions of Linguistic Cohesion via Gravity

$$F = G \frac{p_1 p_2}{r^2} = 1/D$$

$$D \propto 1/G \frac{r^2}{p_1 p_2}$$

$$D \propto r^2, \quad D \propto 1/p_1 p_2$$

F is ling. attraction, which should produce similarity

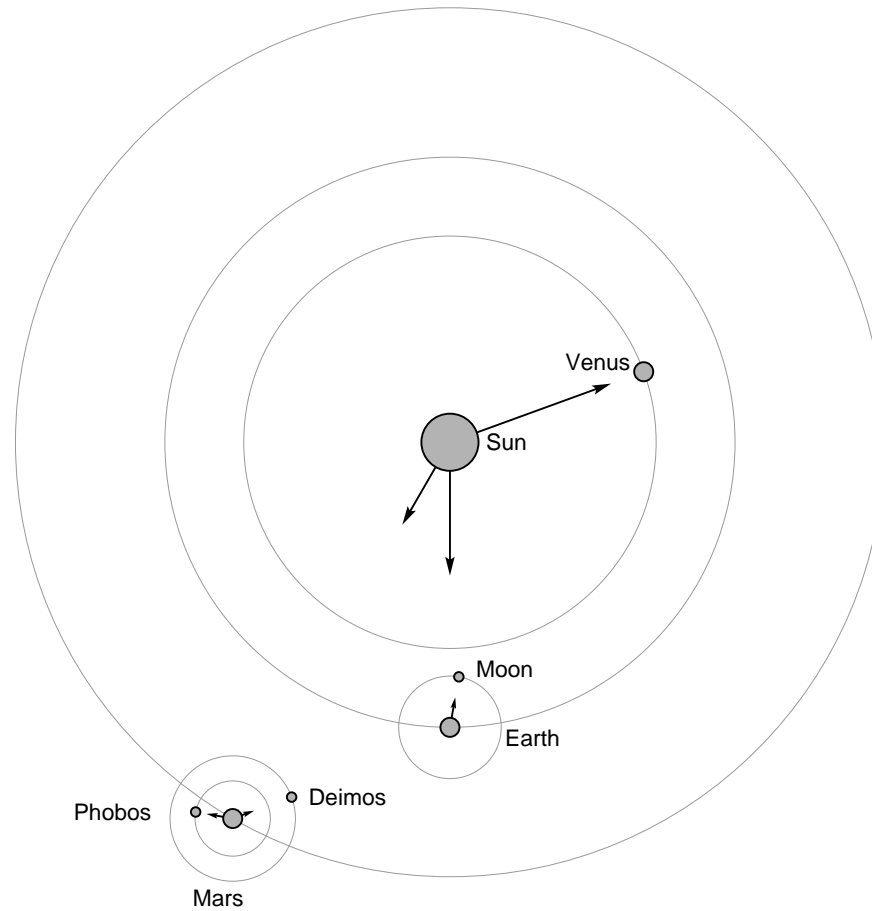
D is ling. dissimilarity

p_1, p_2 the populations of the two settlements, and

r the distance between them



Celestial Gravity





Gravity Studies

- Trudgill examined changes in progress in East Anglia, Norway
- Callary (1975) noted /æ/ changes in Am. mid-west followed degree of urbanization.
- Bailey et al. (1993) noted changes primarily in the direction predicted by gravity in Oklahoma, but also counterexamples. Inchoative *fixin' to* spread from rural to urban areas.
- Boberg (2000) criticizes gravity for ignoring political border (U.S./Canada), shows effect of border.
- Horvath (2001) see little confirmation of gravity in predicting /l/ vocalization in Australia and New Zealand



Sample





Distance vs. Travel Costs

- Social contact should depend more directly on travel costs, not distance as the crow flies
- Ilse van Gemert's thesis in *Informatiekunde*:
 - Estimate travel time (cost) using GIS
 - Investigate whether travel time improves prediction of linguistic distance.
- Tricky issue: GIS is designed to analyse variables with respect to places, not pairs of places.



GIS cells



- darkest - roads
- dark grey - waterways
- white - wetlands
- light grey - other (farmland, woods, ...)



Find Cheap Routes (Greedy)

```
find_route(loc,dest,cost,bound);
do until loc=dest:
    loc_cost <- max_int;
    for loc' in neighbor(loc) do:
        if cost(loc') < loc_cost
            then next <- loc';
                loc_cost <- cost(loc')
        endif;
    endfor;
    path <- path ^ next;
    cost <- cost + loc_cost;
    loc <- next;
enddo;
return path, cost;
end find_route.
```

Alternatives: Breadth-first, depth-first, branch-'n-bound



Least cost travel routes

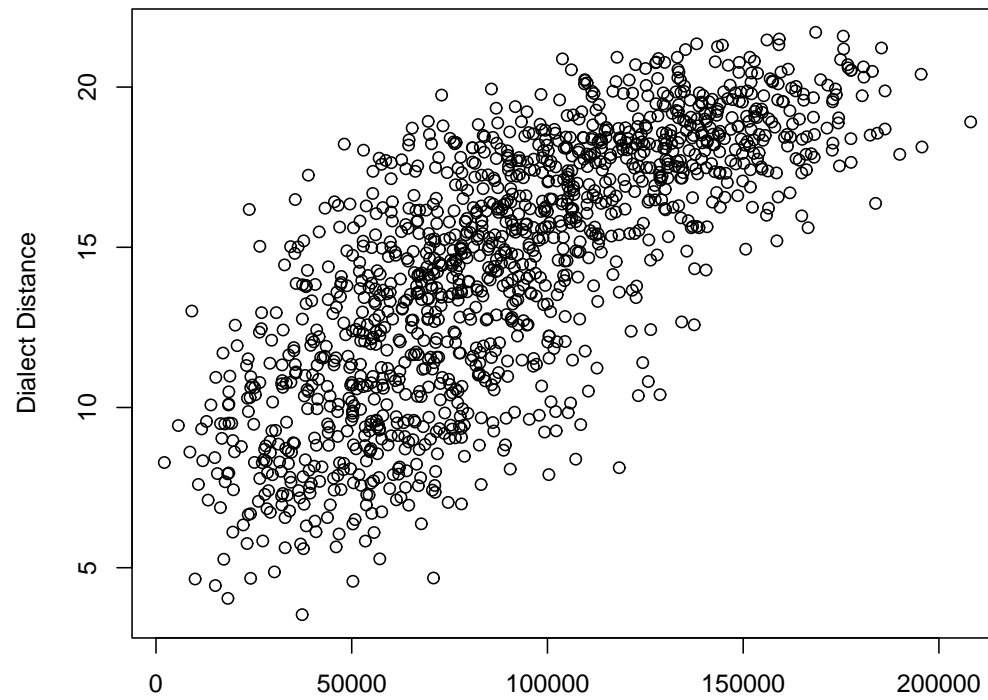


Correlation travel cost, as-the-crow-flies distance ($r = 0.92$)



Look at Data

Linguistic Distance vs. Geographic Distance

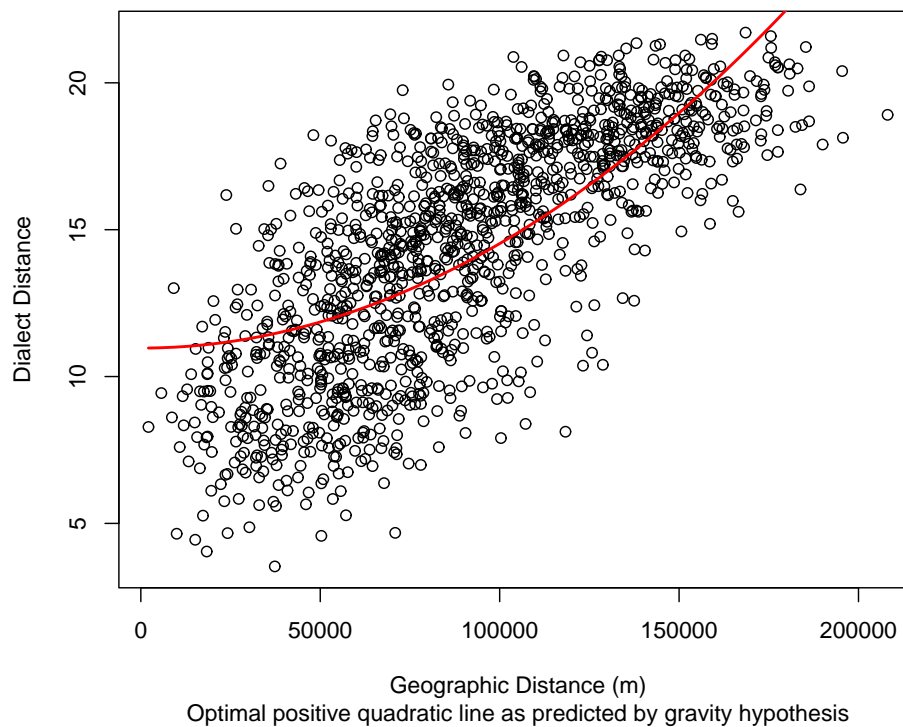


Gravity predicts a positive quadratic effect!



Quadratic?

Linguistic Distance vs. Geographic Distance

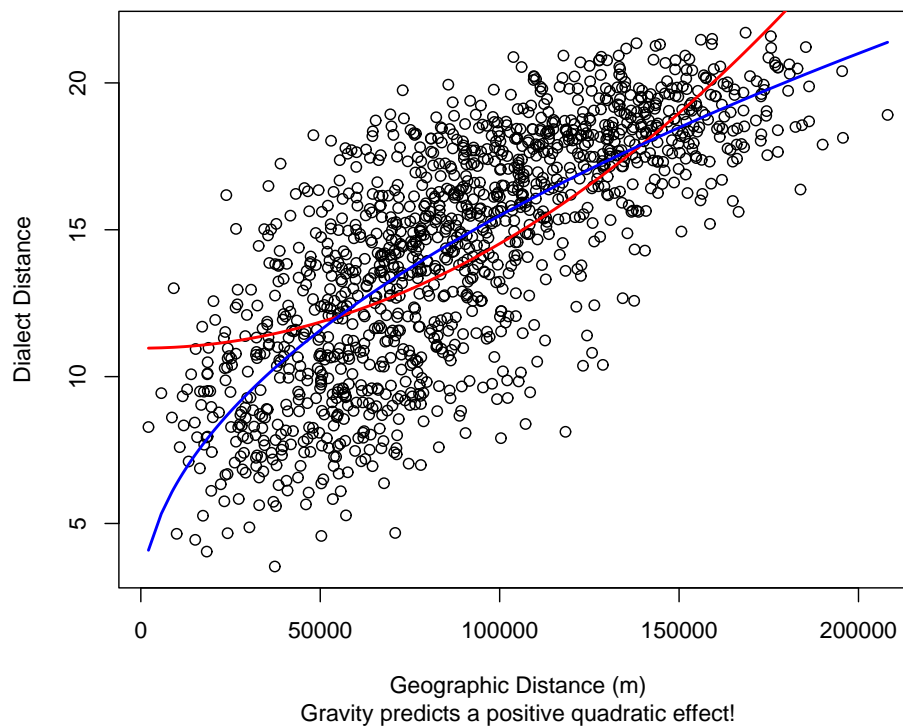


Shape? Zero? ($r^2 = 0.5$)



Function of $\sqrt{\text{dist}}$?

Linguistic Distance vs. Geographic Distance



Shape? Zero? ($r^2 = 0.58$)

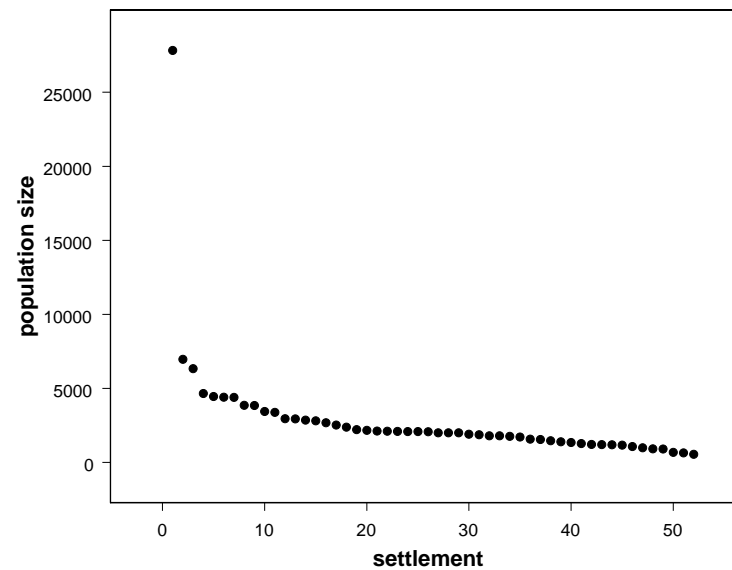


LAMSAS

- LAMSAS phonetic distances correlate $r = 0.557$ w. geography, $r = 0.588$ w. root geography (symbols), $r = 0.531$ and $r = 0.552$ (features)
- Lowman's portion correlates $r = 0.476$ w. geography, $r = 0.501$ w. root geography (symbols), $r = 0.390$ and $r = 0.413$ (features)



Range of Populations



Ramaer (1931) *Geschiedkundige atlas van Nederland; Het koninkrijk der Nederlanden 1815-1931*



Interpreting Results

Trudgill's gravity model

- attraction is relatively stronger over short distances
- therefore linguistic distances should be relatively smaller over these short geographic distances

Observations

- Linguistic distance increases positively with geographic distance, *but*
- effect is proportionately **greater** over short distances rather than proportionately smaller



Speculation on Cultural Dynamics

Not **attraction**, as Trudgill postulates, but rather **repulsion/fission/differentiation** is the fundamental cultural dynamic.

It is natural to see this grow relatively weaker over long distances.

In spite of enormous linguistic pressures toward accommodation.



Further Results

- very weak, and surprisingly, also **positive correlation** of ling. distance with population size
 - likewise suggests fundamentally repellent force
- Van Gemert (2002) and Gooskens (2004) attempt improvement using 19th cent. travel time instead of geography
 - no improvement in (flat) Netherlands (van Gemert), massive improvement in rugged Norway (Gooskens)



Further Questions

- How does varietal distance compare to other indicators of cultural affinity?
—schooling, dress, church, recreation, architecture, ...
- How does varietal distance compare to other indicators of genetic relationship?
—genetic distance, patronymic distance
- Are there better (secondary) predictors of varietal distance?
—waterways, trade connections, marriage patterns, pilgrim routes, ...