A Dialectometric View of Linguistic "Gravity"

John Nerbonne, Ilse van Gemert, and Wilbert Heeringa

Alfa-informatica, University of Groningen

{nerbonne,heeringa}@let.rug.nl

P.O.Box 716, NL9700 AS Groningen, Netherlands

Tel. +31 50 363 58 15, FAX +31 50 363 68 55

1st March 2005

Abstract

Trudgill's (1974) suggestion that the diffusion of dialect features might obey a "gravity"-like law (also known as a hierarchical model or a cascade model) has been tested using individual features undergoing change in several different places, with differing results. The present paper replaces the examination of individual features with a dialectometric measure of aggregate differences, and eschews the focus on individual features undergoing change for an examination of the residue of pronunciation differences resulting from lengthy interaction. We show that the aggregate analysis indicates that geography indeed plays an overwhelming role, that there is no dominant gravity-like (inverse-square) force evident in the residue of linguistic differences, and that the role of population, while weak, is actually the opposite of that postulated by the gravity model.
Keywords: dialectology, dialectometry, linguistic diffusion, gravity theory, edit distance

1 Introduction

The oldest branch of dialectology is the study of what is today often referred to as "dialect geography", i.e. the study of the geographical distribution of language varieties, as opposed to the study of many other relations between language varieties and external conditioning factors, such as social class, age, and sex. While it is clear that geography has a massive influence on the distribution of language varieties, and that closer varieties are normally more linguistically alike than more distant ones, still there have been surprisingly few attempts to examine these relationships with an eye toward more precise formulations.

Trudgill (1974) is an honorable exception to this last generalization. Trudgill proceeds from the very plausible assumptions that closer dialects must influence each other most strongly, and that intensity of social contact is likely to determine the degree of influence. He shows how to tie these ideas together in a model which hypothesizes a gravity-like attraction between dialect varieties, where population is the analog to physical mass, and geographic distance plays its customary role. He adduces evidence in support of this view, relying on selected dialect features.

Dialectometry provides more general tools with which such relationships may be studied (Goebl 1982, Goebl 1984, Nerbonne & Kretzschmar 2003), and the present paper is an attempt to apply dialectometry to evaluate Trudgill's ideas more systematically. In fact it has been common to examine the dependence of dialect distance on geography from the earliest work on in dialectometry (Séguy 1971, Heeringa & Nerbonne 2002, Gooskens 2004). There has been no systematic examination of Trudgill's gravity hypothesis from a dialectometric perspective, however.

In the current paper we expose Trudgill's fundamental ideas to dialectometric examination. The following section presents Trudgill's ideas, their motivation and an overview of previous work. Section 3 describes the experiment, including the data sources, and Section 4 presents the results, which certainly do not provide confirmation for the importance of the role of gravity, or centripetal forces due to social interaction. The final section discusses these results, and suggests an interpretation which does not dismiss gravity, which after all, is theoretically well-founded, but which emphasizes that centrifugal forces, especially dialect differentiation, are more important.

2 Background

In this section we present, in turn, Trudgill's "gravity" theory of dialect dynamics, which might be seen as a reaction to the "wave" models of linguistic diffusion (Schmidt 1872), our own ideas on measuring the pronunciation distance between dialects, and the basic idea of testing the one via the other.

2.1 Trudgill's Gravity Model

Schmidt (1872) introduced the idea that a given linguistic change might spread in waves from a center of innovation, an idea that is at the base of many models of the diffusion of linguistic change (Wolfram & Schilling-Estes 2003, p.721). Peter Trudgill introduced an important refinement, suggesting the application of a GRAVITY MODEL, which had been used earlier in social geography, to questions of linguistic borrowing (Trudgill 1974). In Trudgill's view linguistic innovations spread as if a driving force were population size. In a typical case, an innovation spreads from a large population center directly to another intermediately sized one, often by-passing smaller, geographically intermediate sites. It then in turn spreads from the slightly smaller sites to yet smaller ones, and so on. It is as if each population center had its sphere of influence and that behavior within it is best studied with respect to the locally influential center.

The gravity model thus postulates that linguistic innovations do not simply radiate from a center, as they might in a pure version of the wave theory, but rather that they affect larger centers first, and from there spread to smaller ones, and so on. For this reason it is also referred to as a CASCADE model (Labov 2001, p.285): linguistic innovations proceed as water falling from larger pools to smaller ones. In particular, it should be possible for changes to "hurdle" immediate neighbors, instead of working only very locally.

The connection to physical gravity is suggested in Figure 1. In understanding the movement of heavenly bodies, it is best to concentrate on the nearest very massive body. Thus, even though the moon is affected by the sun's mass, its rotation is determined almost entirely by the much closer Earth. The physical theory of gravity accounts for this by postulating a force due to gravity which is inversely proportional to the square of the distance between bodies. In this way very distant bodies are predicted to have much less influence than nearby ones.

Social science uses of "gravity models" emphasize the importance of social contact and its role in suggesting and promoting the adoption of social and cultural innovations. Some of the phenomena studied by social geographers propagate spatially in a way that reflects their dependence on social contact. People generally adopt new styles of dress, new styles of housing and simple new technologies only after seeing others use them. Social contact is not merely a necessary condition for the spread of linguistic variants; contact frequency also determines the chance of adoption, and ultimately, spread. We should expect many, and perhaps very nearly all dialect variants

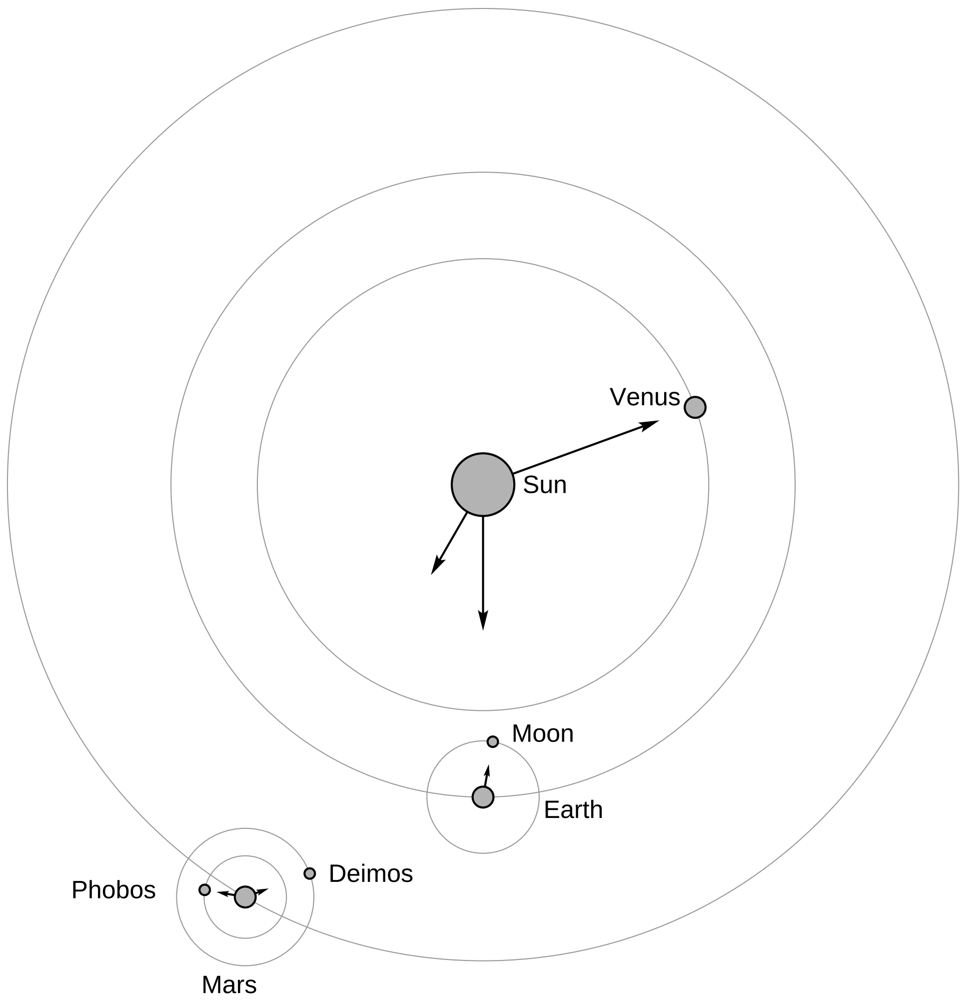

Figure 1: According to the "gravity" model of linguistic dynamics, large population centers exert a force on smaller ones in proportion to the production of their populations, just as the presence of large heavenly bodies exerts a force on smaller ones in proportion to the product of their masses. Because distance likewise plays a role which diminishes quadratically, the most important influences are local ones. Thus, just as the Earth largely determines the behavior of the Moon, so will a local population center dominate within its own vicinity.

similarly to require experience before they could be adopted, and it is also likely that the frequency of experience plays an important role.

Motivation

It is more than plausible that interaction with novel varieties disrupts customary speech habits and promotes the spread of novel linguistic variants. One indication of this plausibility is the readiness with which one interlocutor will adapt his speech to another's. We see adjustments within the time span of individual conversations, and there is evidence that lasting effects also obtain. We review these issues presently.

It has been noted in various subfields of linguistics that conversation partners regularly adjust their speech habits to "accommodate" each other's use of language. Lewis (1979) introduced a principle of accommodation in discourse analysis to account for the willingness of interlocutors to interpret each other charitably, even in the face of apparent infelicities. Language acquisition experts have noted that adults spontaneously simplify their speech in conversation with young children (Clark & Clark 1977), a phenomenon they refer to as "motherese" or "caretaker speech". Closer to home, students of dialect contact regularly note that speakers in multi-dialectal exchanges may temporarily adopt (some of) their conversation partner's dialect features (Giles 1994).

Of course, it is one thing to demonstrate a temporary accommodation to a conversation partner, and quite another to show that there are permanent effects of such accommodations either at the level of the idiolect (the speech habits of an individual) or at the level of the dialect (the speech habits common to a social or geographic group). Trudgill (1986, Ch.1) presents an overview of what is known on the first topic, along with his own studies of Englishmen in the U.S., and it is quite clear that individuals do adjust their speech habits when they live in another dialect area for a considerable length of time. Trudgill (1986, Ch.2-4) is then an extensive essay which establishes quite convincingly that dialects do borrow from one another following patterns which suggest a dependence on social contact, which in turn makes accommodation as a mechanism quite plausible.

2.2 Formulation

If social contact promotes the transfer of features, then we should be able to quantify its overall effect on entire settlements—villages or towns. The overall effect on the varieties associated with settlements should depend on the numbers of individual contacts, which in turn depends on how far apart the settlements are. Distance impedes the chance of contact, so that the further apart the settlements, the less chance of contact. Trudgill takes this idea a step further and suggests that the contact should decline, not as a linear function of distance, but rather quadratically. This seems reasonable if we consider that the area within a given distance of a settlement also increases as a quadratic function of the distance. If we imagine a dialect speakers traveling randomly from a given place of residence, then the chance of traveling to a given point should also fall quadratically with the distance from the place of residence.

The size of the settlements clearly promotes the chance of contact, however. In fact, for two settlements of size P_1 and P_2, the number of chances at social contact will rise with the product $P_1 \cdot P_2$, the number of pairs of people where the first person is from the first town, and the second from the second. Let us note that there is room for the incorporation of further factors here, including perhaps whether a town lay on a frequent trade or pilgrimage route, or whether it was a market center or seat of (local) government.

Trudgill (1983, p.75) pulls these two factors together in a formula suggesting a linguistic counterpart to the law of gravitation:

$$I_{ij} = s \cdot \frac{P_i P_j}{(d_{ij})^2}$$

where I_{ij} represents the mutual influence of centers i and j, P_i is the population of center i, etc., and d_{ij} is the distance between i and j. s is a constant needed to allow for simple transformations, but it may be viewed as "variable expressing linguistic similarity".[1] We note that Trudgill's discussion makes it clear that he would allow that s differ, depending on the similarity of the varieties he was measuring. We shall not exploit this feature of his ideas—which would indeed resist incorporation into the experiment below, but we shall take care to limit our study to fairly similar varieties. We return to this issue in Section 5 below.

The formula thus encapsulates a view of how population size and geographical distance may influence dialect differences. As our discussion has tried to show, the view accords with the notions of accommodation discussed in Section 2.1. If the result of the formula is large, it means that center i has a high level of interaction with center j, meaning that we expect their dialects to influence one another a great deal.

It will be convenient to refer below to the two consequences of the gravity theory which we have emphasized thus far, viz., that interaction should cor-

[1]It is indeed a perfect analog of the formula specifying the force due to gravity:

$$F_{ij} = G \cdot \frac{m_i m_j}{r^2}$$

in which the masses (m) of the objects play a promoting role, the distance between them (r) a suppressing one, mediated by a gravitational constant, G.

relate positively with the product of population settlement, and negatively with squared distance:

$$\begin{aligned} I_{ij} &= s \cdot \tfrac{P_i P_j}{(d_{ij})^2} \\ I_{ij} &\propto P_i P_j \\ I_{ij} &\propto 1/{d_{ij}}^2 \\ I_{ij} &\propto -{d_{ij}}^2 \end{aligned} \quad (1)$$

2.3 Work To-Date

Trudgill (1974, 225ff) examines different pronunciations of the phoneme /æ/ in southern Norway, showing that pronunciations in sites closest to Larvik, a local population center, also most closely resembled it. He chose this pronunciation because it was changing at the time the data was collected. In this way he obtained a view of a change in progress, which, indeed accorded with the predictions of the gravity model.

Callary (1975) noted a strong correlation between the height of /æ/ in Illinois speakers with the size of the city or town those speakers came from. The more urban the speaker's background, the higher the vowel pronunciation. He noted that this is an exception to the predictions of the wave theory and specifically suggested Trudgill's gravity model as a potential explanation (p. 168).

Trudgill (1986) establishes intimate borrowing in a number of ways, including especially an extensive survey of the relevant literature and also several quantitative studies of individual dialect features (e.g. pp. 42,64,111), but there seems to have been no attempt to generalize over a number of features to examine whether geographically proximate varieties in general become more similar over time. This work was not specifically presented as an investigation of the gravity model, but it reaffirms the plausibility of the underlying assumption that social contact is an important factor leading to the acceptance of change.

Bailey, Wikle, Tillery & Sand (1993) and Wikle & Bailey (1997) investigate several ongoing changes in Oklahoman varieties, concluding that while several, indeed most, follow the direction of spread from larger to smaller settlements, important exceptions actually *reversed* the trend. They show that inchoative *fixin' to* has spread from rural to urban areas, demonstrating that this direction is also possible. They attribute the reversal to the prestige ascribed to the use of this form.

Boberg (2000) examines the degree to which the gravity model can account for diffusion across the U.S.-Canada border and concludes that it has relatively little predictive power. In particular, he shows that Windsor, Ontario, which is immediately adjacent to the U.S. border, and to the large population center of Detroit, is no more "American" in its pronunciation than Toronto. He suggests that the border might need to be included in the spatial model, but does not attempt to present a more refined model, and agrees with Bailey et al. (1993) that subjective elements of prestige require attention.

Horvath & Horvath (2001) examine /l/ vocalization in Australian and New Zealand English, which they demonstrate to be a change in progress by showing that it is universally more frequent in younger speakers as compared to older ones. They conclude, however, that "a gravity model [...] does not account for the diffusion of /l/ vocalization." They suggest that this reflects an oversimplification of the model, which attributes diffusion to spatial effects without allowing that specific *places* may differ in their spatial properties.

Wolfram & Schilling-Estes (2003, p.732) report on a resisted change, an island off the coast of the American South which is not acquiescing in the widespread Southern U.S. change of /aɪ/ to /a/, which they attribute to the islanders' valuing it "as a marker of in-group identity".

In summary, research has not overwhelmingly vindicated Trudgill's postulation of a gravity-like effect in linguistic diffusion. There have been voices of affirmation, but even these have noted several counterexamples. We see two points at which a dialectometric analysis might contribute to this discussion. First, Trudgill's and others' studies might rely on fortuitously chosen features which corroborate or contradict the lasting influence of accommodation, but which might be atypical. Since the *prima facie* case for contact-induced dialect change is strong, however, further investigation about its generality is warranted. By examining the gravity hypothesis from a dialectometric perspective, the present study attempts to aggregate a larger number of linguistic variables, and thus to examine Trudgill's ideas from a more general perspective. The present study thus seeks to test Trudgill's ideas over a large range of linguistic features. Second, given the tools of dialectometry, we believe we will be in a position to quantify the strength of gravity or other other forces at play in the dynamics of dialect changes. This option is not available to those working on isolated linguistic features.

2.4 Dialect Distances

In our own work we have developed measures of the aggregate linguistic distance between varieties. We describe the method in this section.

There are several ways in which phoneticians have tried to measure the distance between two basic sounds, most of which are based on the description of sounds via a small (≤ 25) number of features (see Heeringa (2004*b*) for details). There is also a standard technique for the computational comparison of sequences, viz., Levenshtein distance, also known as (string) edit distance, and we set out to combine these techniques.

2.4.1 Segment Distances

The simplest way to compare phonetic segments is to regard identical segments as contributing nothing to phonetic distance and to regard all non-identical segments as contributing the same, unit amount of distance. The simplest distance measure is thus binary: non-identical phones contribute to phonetic distance, identical ones do not. Thus the pair [ɪ,ɒ] counts as different to the same degree as [ɪ,e]. In more sensitive versions phones are compared on the basis of their phonetic features, so the pair [ɪ,ɒ] counts as more different than [ɪ,e]. We have experimented extensively with a range of feature systems designed for this purpose, but we have not been successful in determining the relative contributions which different features should contribute to overall phonetic distance. This means that the various feature systems we have experimented with—both those of others' and those we have devised ourselves—have not proven more reliably sensitive in detecting aggregate linguistic distance that is implicit in the data of the dialect atlases.

The phonetic segment distance measure we use in this paper is based on the comparison of spectrograms of the sounds. A spectrogram is a mapping from time and frequency to intensity and captures most of the information available to the human ear. We are attracted to using spectrograms as a basis for segment distance in order to avoid the problem of determining the appropriate relative contribution of the different phonetic features. The spectrograms we used were made on the basis of recordings of the sounds of the International Phonetic Alphabet as pronounced by John Wells and Jill House on the cassette *The Sounds of the International Phonetic Alphabet* from 1995.[2] The different sounds were isolated from the recordings and monotonized at the mean pitch of each of the two speakers with the program PRAAT.[3] Next, we deployed PRAAT to obtain a spectrogram for each sound using the so-called Barkfilter which is a perceptually oriented model. On the basis of the Barkfilter representation, segment distances were calculated as curve distances between the two spectrographic mappings. The precise way in which this was done is described extensively in Heeringa (2004*b*, pp.79–119) and briefly in Gooskens & Heeringa (2004).

Because small differences in pronunciation may contribution inordinately to the perception of phonetic distance, we emphasize small differences by applying a logarithmic transformation to the curve distance obtained in the way described above. To avoid taking a logarithm of zero, we calculate a slightly modified quantity:

$$\frac{ln(distance + 1)}{ln(maximum\ distance + 1)} \times 100$$

We turn now to the Levenshtein distance, which may be regarded as a means of lifting the segment distances obtained thus far to the level of sequence distances. The basic idea behind Levenshtein distance is to imagine that one is rewriting one string into another. The rewriting is effected by basic operations, each of which is associated with a cost, as illustrated in Fig. 2.

The operations used were (i) the deletion of a single sound, (ii) the

[2]See `http://www.phon.ucl.ac.uk/home/wells/cassette.htm`.

[3]The program PRAAT is a free public-domain program developed by Paul Boersma and David Weenink at the Institute of Phonetic Sciences of the University of Amsterdam and available at `http://www.fon.hum.evua.nl/praat`.

	Operation	Cost
æ ə f t ə n ʉ n		
æ f t ə n ʉ n	delete ə	d(ə,[])=0.99
æ f t ə r n ʉ n	insert r	d([],r)=0.95
æ f t ə r n u n	replace [ʉ] with u	d([ʉ],[u])=0.76
Total		2.70

Figure 2: Levenshtein distance between two sequences is the least costly sum of costs needed to transform one string into another. The transformations shown here are associated with costs derived from spectrograms, i.e. the distance between the three-dimensional curves representing individual phonetic sounds. The pronunciations are from the *Linguistic Atlas of the Middle and South Atlantic States* (Kretzschmar et al. 1994).

insertion of a single sound, and (iii) the substitution of one sound for another. We have experimented with other operations, but we have made no use of them for this work. The operation costs used in the procedures were those derived from the distance between spectrograms in a reference database (again, we refer to Heeringa (2004*b*) for details). They consist of the measure of the distance between the sounds (in the case of substitution), and the measure of the distance between a given sound and silence (in the case of insertions and deletions).

We insist on proper alignments in the application of edit distance by requiring in general that only vowels may match with vowels, only consonants with consonants, but allowing the exceptions that [j] and [w] may align with vowels, [i] and [u] with consonants, and central vowels (here effectively only the schwa) with sonorants. Thus the [i], [u], [j] and [w] align with anything, but otherwise vowel/consonant status is respected so that unlikely matches (e.g. a [p] with a [a]) are prevented.

Comparing pronunciations in this way, the distance between longer pronunciations will generally be greater than the distance between shorter pronunciations. The longer the pronunciation, the greater the chance for differences with respect to the corresponding pronunciation in another variety. Because we would prefer not to exaggerate the effects of sounds in longer words, we normalize the raw distances obtained by dividing the raw distance by the length of the longest alignment which gives the minimum cost. The longest alignment has the greatest number of matches. We illustrate this with an example:

æ	ə	f	t	ə	∅	n	ʉ	n
æ	∅	f	t	ə	r	n	u	n
	0.99				0.95		0.76	

The total cost of $2.7(= 0.99 + 0.95 + 0.76)$ is now divided by the length of 9. One important advantage of this procedure is that word distances are now expressed as percentages of a potential maximum. In the case above we obtain a word distance of 0.3 or 30%.

Our varietal comparisons are made on the basis of 125 words, yielding 125 word distances per pair of varieties. We assay the distance between the varieties to be the mean distance in our 125-element sample. Since word distances are expressed as percentages, mean varietal distances are also percentages. All the distances between the 52 Low Saxon varieties are then arranged in a 52×52 matrix.

If we apply a Levenshtein procedure to about 100 words from several hundred field work sites, the result may be shown to verify the idea of dialect areas as used in traditional dialectology (Nerbonne, Heeringa & Kleiweg 1999). These may be reconstructed via clustering techniques, but also via the statistically more stable multi-dimensional scaling. Levenshtein distance has been shown to be consistent and valid with respect to the judgments of lay dialect speakers (Gooskens & Heeringa 2004, Heeringa 2004*b*).

Kessler (1995) first applied Levenshtein distance to phonetic transcriptions to measure the linguistic distances between (Irish) varieties. Nerbonne, Heeringa, van den Hout, van der Kooi, Otten & van de Vis (1996) and Heeringa (2004*b*, pp.213–278) have applied the techniques to Dutch (see also references there), Bolognesi & Heeringa (2002) to Sardinian, Heeringa & Gooskens (2003) to Norwegian, and Nerbonne & Siedle (2005) to German. Heeringa (2004*b*, pp.121–135) is the most complete description, and we used exactly the scheme presented there to obtain the measurements in this paper.

Although the Levenshtein technique was developed to measure the distance between sequences of phonetic segments, it measures all differences which are reflected in the phonetic transcriptions of dialect atlases, which typically consist of realizations in context, and which therefore include lexical, phonetic and morphological differences as well.

In this paper we shall use the Levenshtein distances to test the idea of linguistic gravity, to which we turn in the next section.

2.5 Dialect Distances and Gravity

The fundamental idea behind the current experiment is to test the gravity theory, which is a claim about the dynamics of dialect change, using synchronic dialect distances. Since this is methodologically innovative, let us dwell on it briefly. Examining a range of dialect sites from a fairly stable region, we reason that, if they are most strongly affected by the forces of linguistic gravity, then the patterns we find in the synchronic data should reflect the accumulated effects of linguistic gravity. Synchronic differences should reflect historical dynamics. In particular the varieties closest to one another and those involving larger populations should be linguistically most similar as well. In this way we propose to test the gravity idea, examining synchronic (linguistic distance) only. We further take care to make explicit here the assumption that the adoption of features from one dialect into another should make them more similar (than they originally were). We note one advantage which immediately accrues to this sort of probe: it does not require that we isolate ongoing changes and try to wring from them a direction. This was required in earlier examinations of the gravity hypothesis (Section 2.3).

This only makes sense if the linguistic data we are examining betrays the effect of incomplete diffusions, changes which, for whatever reason, have not propagated (yet) throughout the area we are examining or which were partially overturned by later ones. Any completely successful change will not introduce a linguistic difference which our measurements can be sensitive to.

Since we are employing an aggregate dialectometric technique, we shall be in a position to evaluate the overall tendencies shown in diffusion. We shall not restrict our attention to a small number of linguistic features, and are therefore in a position to resolve the difference of opinion with respect to the gravity model noted in Section 2.3. Our measurements will note indicative and counter-indicative phenomena alike, and also quantify which are dominant.

If we are to use synchronic linguistic distances to test claims about the diachronic development of dialects, then it is sensible to use data on the independent variables geography and population from a substantially earlier time, assuming that relative population size has been fairly stable. We imagine dialects undergoing small changes over a long period of time and continuously changing, and we want to allow enough time to lapse to give the processes of social contact a chance to accumulate effects. Finally, we settled on gathering data on population and distance (see below) from the time before the introduction of the railroads, more exactly in 1815, well before the times of modern mobility, and roughly $100 - 150$ years before the linguistic data was collected. We assume that the dialects we examine continued to influence one another from then on, and so it is preferable to examine settlements that have been fairly stable in relative size and accessibility.

We note that dialect surveys (including the one we used, introduced below) prefer older, non-mobile respondents, which means that the linguistic time lag is undoubtedly shorter than the $100 - 150$ years between the time for which population sizes are available and the time of publication of the dialect atlas, perhaps by as much as 75 years. Although we believe that it would be legitimate to apply this sort of analysis in a fairly stable dialectal situation even with no time lag, we wished to err on the side of caution and sought data that would certainly reflect the accumulation of changes over decades. Perhaps it is not superfluous to add that we concede that it is difficult to determine a most time appropriate lag in a non-arbitrary way.

Since we shall only observe the effects at a single time, we likewise assume that the situation long ago does not confound the effects we seek. This might conceivably have been the case if we had chosen a sample in which similarity unfortunately did not (originally) correlate with the chance of social contact, e.g., the Dutch of the *polders* reclaimed from the sea and uninhabited until the 1950's, or the language of areas with large percentages of migrant labor such as the older peat bogs in the north of the Netherlands.

We noted in Section 1 above that interaction was predicted to correlate positively with the product of settlement size and inversely with the square of distance.

$$I_{ij} \propto P_i P_j$$
$$I_{ij} \propto -{d_{ij}}^2$$

If it is correct, as we have just argued, that interaction should result in increased similarity, then similarity should correlate in the same way with population sizes and (inverse) distance. We shall finally be measuring linguistic distance, however, so that we shall test the following two hypotheses:

$$\begin{aligned} LD_{ij} &\propto {d_{ij}}^2 \\ LD_{ij} &\propto -P_i P_j \end{aligned} \tag{2}$$

A second assumption is likewise crucial. We shall essentially test the predictions in Section 2 by examining the correlations between linguistic distance as measured by Levenshtein distance on the one hand and geographic distance and population size on the other. We have no way of controlling for other effects in the data which are also plausible, e.g. the influence of foreign languages, the social homogeneity of the situations, or function of dialect differentiation as a mark of social differentiation. All of this is effectively "noise" in the current scheme.

Finally, let us note that while Trudgill distinguishes the attractive force of a larger settlement on a smaller one from that of a smaller settlement on a larger one (effectively using an analog of the asymmetric acceleration due to gravity), noting that one expects the smaller settlement to accommodate more to the larger one than vice versa,[4] we are restricted to observing only the long-term results of the attraction so that we do not distinguish the two cases. Viewed from another perspective, we are using a true distance measure, which is therefore symmetric. We cannot distinguish the effects

[4]The degree to which i accommodates to j is proportional to $P_i/(P_i + P_j)$.

of i on j from the inverse effects of j on i using this measure. We might be able to get some leverage on the asymmetric effects if we had data from different time points. In the present study we are attempting to evaluate an historical hypothesis on the basis of the accumulated effects it predicts.

We turn now to the details of the experiment.

3 Experiment

In this section we review our selection of data and the conduct of the experiment.

3.1 Linguistic Data

The dialect data from the *Reeks Nederlandse Dialectatlassen* (RND) were used, compiled by E. Blancquaert and W. Pée in the period 1925-1982. In these records we find the pronunciation transcripts of local speakers of each dialects in nearly 2,000 locations from which we can then choose a suitable sample (see below). On the basis of this data, linguistic distances between settlements were calculated using Levenshtein distance (see above). 125 words formed the basis of the calculations (Heeringa 2004*b*, App.B).

Table 1. Words aggregate pronunciation is based on

	Dutch	**English**	**RND**		**Dutch**	**English**	**RND**
1	mijn	my	2	64	koning	king	76
2	vriend	friend	2	65	ook	also	76
3	werk	work	4	66	geweest	been	76
4	op	on	5	67	lange	long	78

5	schip	ship	5	68	woord	word	79
6	kregen	got	5	69	kindje	baby	80
7	brood	bread	5	70	was	was	80
8	vinger	finger	6	71	dochtertje	daughter	82
9	vier	four	10	72	bos	forest	82
10	bier	beer	10	73	ladder	ladder	83
11	twee	two	11	74	mond	mouth	86
12	drie	three	12	75	droog	dry	86
13	hij	he	13	76	dorst	thirst	86
14	knuppel	cudgel	13	77	weg	way	87
15	ik	I	14	78	krom	curved	87
16	knie	knee	14	79	liedje	ditty	90
17	gezien	seen	14	80	goed	good	92
18	kerel	fellow	21	81	kelder	cellar	95
19	stenen	stones	25	82	voor	for	95
20	breder	broader	25	83	moest	must	96
21	duivel	devil	28	84	drinken	drink	96
22	gebleven	stayed	28	85	broer	brother	98
23	meester	master	29	86	moe	tired	98
24	zee	sea	29	87	dun	thin	100
25	graag	gladly	31	88	zuur	sour	100
26	steel	handle	33	89	put	well	101
27	bezem	broom	33	90	uur	hour	101
28	geroepen	called	35	91	vuur	fire	104
29	peer	pear	36	92	duwen	push	105
30	rijp	ripe	36	93	hebben	have	106
31	geld	money	38	94	stuk	piece	106

32	ver	far	39	95	brug	bridge	106
33	brengen	bring	39	96	veulen	foal	107
34	zwemmen	swim	42	97	komen	come	107
35	bed	bed	45	98	deur	door	109
36	springen	spring	47	99	gras	grass	111
37	vader	father	53	100	bakken	bake	113
38	zes	six	53	101	je	you	116
39	jaar	year	53	102	eieren	eggs	116
40	school	school	53	103	krijgen	get	116
41	laten	let	53	104	waren	were	119
42	gaan	go	53	105	vijf	five	119
43	potten	jars	56	106	hooi	hay	122
44	zijn	are	56	107	is	is	122
45	veel	much	56	108	groen	green	122
46	maart	March	58	109	boompje	little tree	124
47	nog	yet	58	110	wijn	wine	125
48	koud	cold	58	111	huis	house	126
49	kaars	candle	59	112	melk	milk	127
50	geeft	gives	59	113	spuit	spouts	127
51	licht	light	59	114	koe	cow	127
52	paard	horse	60	115	koster	sexton	128
53	tegen	against	63	116	buigen	bend	129
54	kaas	cheese	66	117	blauw	blue	131
55	dag	day	68	118	geslagen	struck	131
56	avond	evening	68	119	saus	sauce	132
57	barst	crack	70	120	flauw	flat	132
58	brief	letter	71	121	sneeuw	snow	133

59	hart	hart	72	122	doen	do	136
60	spannen	put	74	123	dopen	baptize	137
61	nieuwe	new	74	124	dorsen	thresh	138
62	kar	cart	74	125	binden	bind	139
63	zoon	son	76				

For each pair of settlements in the sample (see below), we obtain a measure of the pronunciation distance between the settlements. It is best to imagine the results as a distance chart of the sort created and distributed by automobile clubs. But the linguistic distance chart is a table in which the cells are not travel distances or travel times, as in the auto club charts, but rather linguistic distances. Every cell in the table represents the linguistic distance between the two settlements. Naturally, the diagonal contains only zeroes (the linguistic distance from a settlement to itself), and the table halves above and below the diagonal are symmetric, just as all distances are.

We shall then try to predict this distance using geographical distance on the one hand and the inverse of the populations' product on the other.

3.2 Choice of Settlements

The particular area chosen for measurements may be crucial. Naturally we wish to use data that has been collected and recorded consistently and accurately. Further, since we can only test Trudgill's formulation of the gravity idea using settlements for which it is plausible to assume that they do not differ (at least not systematically) with respect to Trudgill's linguistic similarity constant s, we should not choose an area straddling a major dialect boundary.

In fact, our first attempt at testing the ideas was flawed in a way we can attribute to our disregard for the influence of similarity. We first analyzed the predictions of the gravity theory using data from the province of Friesland (*Friesland*). Friesland is the only area in the Netherlands that is distinguished by its own, nationally recognized independent language, viz., Frisian.[5] However, dialect islands are found within the Frisian language area in which a variety known as TOWN FRISIAN is spoken, a variety intermediate between Dutch and Frisian (Heeringa 2004*a*). The towns show relatively little similarity with the villages, even with the neighboring villages. Our model could therefore explain less of the data than we had hoped (as one can see in Table 2). In hindsight, we attribute this to the violation of the assumption that the settlements involved be similar to one another to roughly the same degree.

Model	r	r^2	Δr^2
distance^2	0.212	0.045	0.045
dist.^{2} & pop.	0.221	0.049	0.005

Table 2: Correlation coefficients for the correlation of linguistic distance (Levenshtein distance, explained in § 2.4 and predictions based on distance (first line) and distance together with population (second line)) in Frisia. The correlation of geographic and linguistic distance is significant ($p \leq 0.001$), but the addition of population adds nothing further of significance (Mantel test, see below).

The distance model in Table 2 (first line) shows the degree to which the linguistic distances can be explained by squared geographic distance, while the second model includes distance and the product of the population of the settlements. As we see, the results show that the simple geographic model explains little of the variance in the data. Distance explains 4.5% of the

[5]Frisian is officially the "tweede rijkstaal" of the Netherlands, the second national language.

variance in the Frisian data, while incorporating population allows one to explain only 0.4% more. Perhaps worse, the correlation of linguistic distance with population product was actually positive, rather than negative. Settlements with large numbers of people were thus *less* likely to be similar, not more. As we shall see (below), some of the disappointments in this original sample did not arise out of accident.

The Groningen-Hengelo Sample

Figure 3: The 52 settlements in our study all lie within the Lower Saxon dialect area of the northern Netherlands.

We then sought a more suitable set of locations for the study, in particular an area with a few larger settlements and a larger number of smaller settlements varying in population size, which, moreover, does not encompass known dialect islands or significant dialect boundaries. The 52 locations we chose for the model presented in this paper are roughly between Groningen and Hengelo, with no dialect islands and few settlements with large populations. The towns and villages used in our study can be found in Figure 3, and their relative population sizes have been quite stable (see below). The selection of a given settlement was determined by its simultaneous presence in Heeringa's digitized RND data and in the historical atlas (see below).

We are aware that Dutch "city dialects" do not now enjoy great prestige, but this was not true during most of the time at which the diffusions studied here were underway, so this more recent negative prestige should not affect the chance of seeing the larger Dutch town influence smaller ones. This might have been important in view of the remarks in Section 2.3 above, postulating that prestige plays a confounding role in earlier studies. In any case it would be difficult to control for this factor in a study such as the present one, probably meaning that one would need to seek another dialect area.

We also note that the sample used here would be less than optimal if it were necessary to mirror the varieties *now* spoken in the larger towns. This has to do with the field workers' preference for older, less mobile speakers, who are not representative of the current range of speakers. For example, the variety of Groningen dialect recorded in the dialect atlas is not widely used in the city now. If we were trying to observe the effects of gravity today, we should prefer a more representative selection of today's speakers so that the population which is having the attractive effect were better represented.

But less us recall that the entire motivation in choosing older, less mobile speakers is the wish to understand what the language was like at the time of the speaker's childhood. If the RND compilers and fieldworkers choose their respondents well, then they are giving us a picture of the language approximately 50 years before their interviews (on average). So we do not believe our use of atlas material to be a liability in this respect.

An attractive aspect of this choice of settlements is that the area involves no substantial barriers such as mountain ranges, national boundaries, or large bodies of water. Since we are ultimately interested in distance as a predictor of the chance for social contact, it is clear that we wish to use a sample in which distance is likely to reflect the (inverse) chance of social contact. See Section 3.4 for more on this, however.

3.3 Population Size Data

We decided to use population sizes from the time before great mobility (see above), reasoning that these sizes would reflect the interaction of that older period, which should translate into adopted, measurable changes in the period from which our linguistic data is taken. The populations of the different settlements in our model were taken from the *Geschiedkundige atlas van Nederland; Het koninkrijk der Nederlanden 1815-1931* (Ramaer 1931) and date from around 1815.

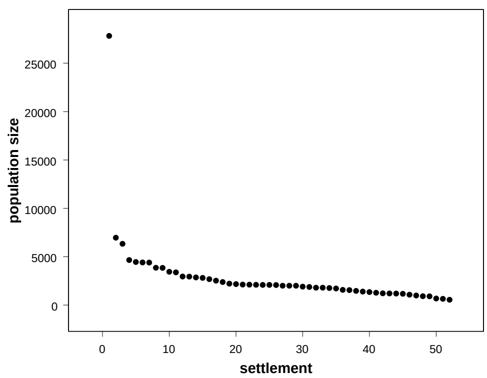

Figure 4: The populations of the 52 settlements in 1815.

Groningen had the largest number of inhabitants, with a population of 27,824. The other populations range from 553 to 6,962 with a regular distribution, as shown in Figure 4. Ramaer (1931) also provides populations for 1930, and we calculated that the populations for 1930 correlated highly with those from 1815 ($r = 0.86$) so that it is safe to say that we are dealing with a set of settlements which is quite stable in relative population size. For each pair of towns in the selection, the product of the population of those two towns was calculated, resulting in a symmetric table, like the table of pronunciations (see above).

3.4 Geographical Distances

We might have measured the distance between settlements using only an approximation such as the Euclidean distance between the longitude-latitude coordinates, i.e., the root of the sum of the squared differences in longitude and latitude. This would distort the true distances since the settlements are distributed over the surface of a sphere (the Earth), not a rectangular plane, and distance should be calculated differently in such situations. In fact we would have accepted this distortion, which would be minor in the case of the small area we are interested in.

Travel Costs

But finally, we are interested in distance as a predictor of the chance of social contact, and for this purpose it seemed best to examine the issue more exactly. Following Gooskens (2004) and Gooskens (N.d.), who in turn was inspired by van Gemert (2002), we decided that travel time should be the more reliable indicator of the chance of social contact. Even though the area we examined has no mountain ranges or large bodies of water, it did have roads at the beginning of the nineteenth century, the period when we assume that dialects we are studying were adopting one another's features. The roads surely would have promoted the chance of social contact, and we assumed it would be possible to incorporate information about them.

We decided therefore to calculate expected travel times to obtain a more reliable indicator of the chance of social contact. Using a Geographical Information System (GIS), we incorporated information on putative travel routes between the settlements in our study. In particular we used ArcInfo and ArcView in order to obtain an estimate of travel time which would reflect the relative road conditions of the time (Razavi 1997).

Van Gemert (2002) reports in more detail on the procedures needed, which we summarize here. We began with a map of the modern Netherlands because this is readily accessible for use in the GIS. We then imposed a grid on this map and assigned a (relative) "travel cost" to each of the cells created this way. For this purpose we consulted the older map. The costs indicate how difficult it was to pass through the (area of the) cell, ranging from '1' for the cost of traveling via road to '5' for crossing a cell consisting entirely of water (in the case at hand, a lake). The height and width of a cell in the grid was 500 meters. We had information about lakes, rivers and the main roads of the area (but not wetlands, see below). We discuss rivers and canals separately below. See Fig. 5, in which the darkest (black) lines represent the roads, and the dark grey lines the rivers. White areas are water surfaces such as lakes and the rest (the light grey color) is ground without special travel properties.

As noted above, we ignored the effect of the railroads because we wished to estimate travel time (costs) around 1815, when there were no railroads. Because a cell is 500 meters wide, several kinds of surface could coincide in one cell (see Fig. 5). If a road crossed a cell, we assigned the entire cell the value of a road. We did not attend to the direction the road took in a cell, which undoubtedly resulted in errors wherever the road in question did not cross a cell in a way beneficial to the route being planned. These errors tended not to accumulate because the fortuitous crossing of a road is a single 500 m. step in a long trip. The errors could only have accumulated if there had been a series of roads perpendicular to the direction of the path. For the purpose of the calculations of difficulty, we regarded this indiscriminate interpretation of the role of roads as noise.

Figure 5: A single cell in the grid with various surfaces. White represents relatively untraversable water, black lines are roads, dark grey lines are canals or rivers, and light grey other areas.

We assumed that roads are the surface most often used for traveling. That is why the pieces of the grid which contain part of a road will have a low value, indicating that they are very easy to travel on. Other surfaces, such as lakes, have a high value. It was, however, difficult to decide how to treat rivers and canals in this system. On the one hand, it is clear that waterways were very important for trade and thus promoted social contact. This argues for assigning these waterways a low travel cost. On the other hand, recalling that we are examining the time before mechanical assistance, when travel in at least one direction on a river or canal would have required that animals or humans pulled boats from the shore, the waterways may not have played an important role for non-trade purposes. For the purposes of the study reported here we experimented with various values for the rivers and canals, which mattered little in the final analysis, perhaps because it turned out that roads often ran parallel to rivers and canals.

It is clear that the role of rivers and canals might fruitfully be analyzed further. In particular, if rivers and canals were more important than roads, then we should explore assignments of values to cells in which costs for travel via waterway were lower than those for travel via road. This has not been done in the work reported on here.

After every cell in the grid has been given a value, the least costly, most easily traversable route from one place to each of the other places can be calculated. The routes drawn between Groningen and each of the other 52 locations are in shown Fig. 6, where the background shading reflects the overall cost of the trip, in particular where darker grey means a more expensive trip. The calculation of the paths was done by means of a simple, dynamic algorithm that selects a route by traveling from cell to cell, meanwhile adding up the least expensive travel costs to each cell it traverses. The algorithm assumes that the least costly path to a cell c will involve travel from a less distant cell c'. This assumption will not be strictly true. For example in modern travel it may clearly be more convenient to overshoot the destination using a relatively inexpensive means of travel, and then to "backtrack". Our algorithm would miss such cases, which we would argue were less likely before the era of cheap mobility.

In the worst case, this algorithm requires that we calculate the least costly path to each of the cells in the smallest square containing the departure and destination settlements, roughly $\ulcorner d/500 \urcorner^2$ cells, where d is the distance from source to destination (in meters), which we divide by 500 to estimate the number of steps (in cells) from the one to the other, and square to obtain the size of the entire sub-grid examined. The final cost indicates the costs through all of the cells which are traversed in going from the one location to the other. Since the maximal distance involved is less than 150 km, in the worst case this algorithm required the examination of 100,000 cells for one of the cities c along this longest axis. Note that the shortest distances from c to all of the other cites would be calculated in the process. We repeat this for each of the 52 cities in the sample, noting that, since distances are symmetric, we will have calculated some distances as we progress from city to city, keeping the total computing time below the worst case $52 \times 10^5 \approx 5 \times 10^6$ cell computations.

Figure 6: The routes calculated to and from Groningen. The background shading reflects the overall cost of the trip, where darker grey means a more expensive trip.

3.5 Analysis

Our data preparation yields three different half-matrices, one showing the linguistic distance between each pair of settlements, one showing their populations' product, and a third containing their geographic distance, measured as described above. In examining correlations involving aggregate linguistic distances we might be seen guilty of the "ecological fallacy" (Freedman, Pisani & Purves 1998, pp.148-50), i.e. of overstating correlations by examining aggregate values. We would counter this putative objection by referring to the need to characterize entire linguistic varieties with respect to some aggregate of their properties. The ecological fallacy arises, e.g. when one studies the relation between income and education not on the basis of individual incomes and educational levels, but rather on the basis of average values over several groups. We maintain that it is simply necessary to examine aggregate properties if we are to characterize entire varieties as opposed to single linguistic variables (such as the pronunciation of final /t/). Linguistic varieties have a status unlike "average individuals" which justify this step.

We therefore submitted the data described above to multiple regression analysis, exploring various approximation techniques (stepwise and simultaneous) without noticing effects in the results, to which we turn presently. We note further that, since we are dealing with distances of various sorts, the assumption of independence of observations is violated, meaning that the statistical significance of the correlation coefficients may not simply be read from a standard table or an SPSS output. Mantel (1967) suggested a permutation technique for evaluating significance in such cases. Since we are dealing with large numbers (of distances, normally $\binom{52}{2} = 1,326$), statistical significance is generally not an issue, but we report only correlation coefficients which are significant according to the Mantel test.

4 Results

In this section we examine whether the predictions of the gravity model are fulfilled.

4.1 Geographic Effects

It is a fundamental postulate of dialectology that language variety is structured geographically (Nerbonne & Kleiweg submitted, 12/2003) , so it comes as no surprise that geography is an excellent predictor of aggregate dialect distance. Our regression analyses show that we are able to account for 59% of the variance in aggregate linguistic distance in our sample (see Table 3). The correlation is positive, just as the gravity hypothesis predicts.

As it turned out, the "travel costs" as measured above and the "as-the-crow-flies" distances, approximated by Euclidean distances, noted at the beginning of this section, correlated so highly ($r = 0.92$) that the exercise in defining travel costs did not yield wildly different results in the analysis attempting to explain dialect distances. See Table 3. We note for the future that we expect detailed geographic models to be more important as we analyze mountainous areas, where the barriers to social contact are substantial (Gooskens 2004), and perhaps also when we attempt to incorporate the role of waterways more effectively.

4.2 Population

When we turn to the effect of population size, we examine the addition of the independent variable for populations product to the purely geographic model. The addition of this variable improves the purely geographic model,

	Linguistic	Distance	Travel
Linguistic	1.0	0.768	0.756
(Direct) Distance	–	1.0	0.92
Travel	–	–	1.0

Table 3: Geographic explanation of dialectic differences. Geographic distances correlate highly with aggregate dialect differences ($r \approx 0.77$). Note that the correlation is positive, just as gravity predicts. Travel costs and distances "as the crow flies" correlate so highly that distinguishing them is not interesting in this sample.

but only negligibly (Table 4), allowing a mere 0.4% more variance to be explained with respect to the exclusively geographic model, an improvement that was not even statistically significant. More importantly, and surprisingly, the fundamental relation is not inverse, as the gravity model predicts, but rather direct. That is, the larger the population product, the greater the linguistic distance—exactly the opposite of what the gravity model predicts. We note that this confirms the effect we noted with surprise in Section 3.2 above on the Frisian pilot study.

Model	r	r^2	Δr^2
Geography (d^2)	0.715	0.511	0.511
"Gravity"	0.715	0.511	0.004

Table 4: Adding population effects to obtain a gravity model adds no explanatory power to the model of the Lower Saxon data. The contribution of the population product independently is moreover positive, contradicting the predictions of the gravity model.

This is not what we expected from the model. We thought a more impressive increase of this percentage would take place after the population products were added to our model, and we are very surprised by the direction of the influence. Other things being equal, larger settlements tend to be less similar to one another than smaller settlements.

4.3 The Need for Dialectometry

To underscore the need for an aggregate view, let us note that there are counter-indicators in the data as well. There are individual features which show the negative correlation with product populations predicted by the gravity hypothesis. The pronunciation distances of the word *knuppel* inversely correlate significantly ($r = -0.14$) with the population product. Figure 7 shows a map of the distribution of this word. The inverse correlation arises because the large settlements tended to have similar pronunciations. But if our analysis were to focus on this one item, we would misread the global trend.

4.4 Quadratic?

Of course the gravity model not only predicts a positive correlation between linguistic distance and geographic distance, a prediction which is nearly synonymous with the entire enterprise of dialect geography, but it more exactly predicts that linguistic distance should be a quadratic function of geographic distance. Figure 8 provides a scatterplot of the data, together with the optimally fitting quadratic regression line. As the reader may verify, the cloud of data in the scatterplot does not suggest a quadratic relation. This visual suggestion is also born out by the attempt to model the linguistic distances not as a quadratic function of geographic distance, but rather as a linear or even logarithmic function of geography. These not only result in better apparent fits of the regression curves, but also in statistically significant rises in the correlation coefficient from $r = 0.715$ (for the quadratic curve) to $r = 0.768$ for the linear fit, and $r = 0.751$ for the logarithmic curve (which do not differ significantly). Logarithmic-like curves typically fit this data well: Séguy (1971) presented his dialect distances as

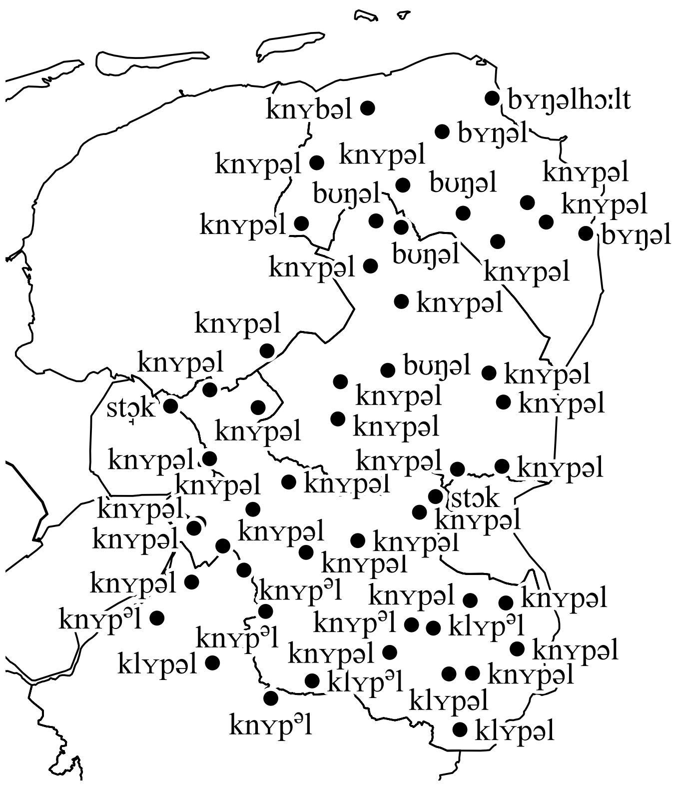

Figure 7: The pronunciation of *knuppel* ‘club’ shows the distribution with respect to population size predicted by the gravity hypothesis (an inverse correlation of pronunciation distance and population product, $r = -0.14$), but it is entirely atypical. In the absence of a dialectometric methodology, we might be tempted to choose such atypical material.

function of the square root of geographic distance, and Heeringa & Nerbonne (2002) as a function of its logarithm.

In this case the linear model is slightly better than the logarithmic model, but there is no significant difference between the two. Since linguistic distance tends to rise to a ceiling when large enough areas are examined, the logarithmic model functions in general better.

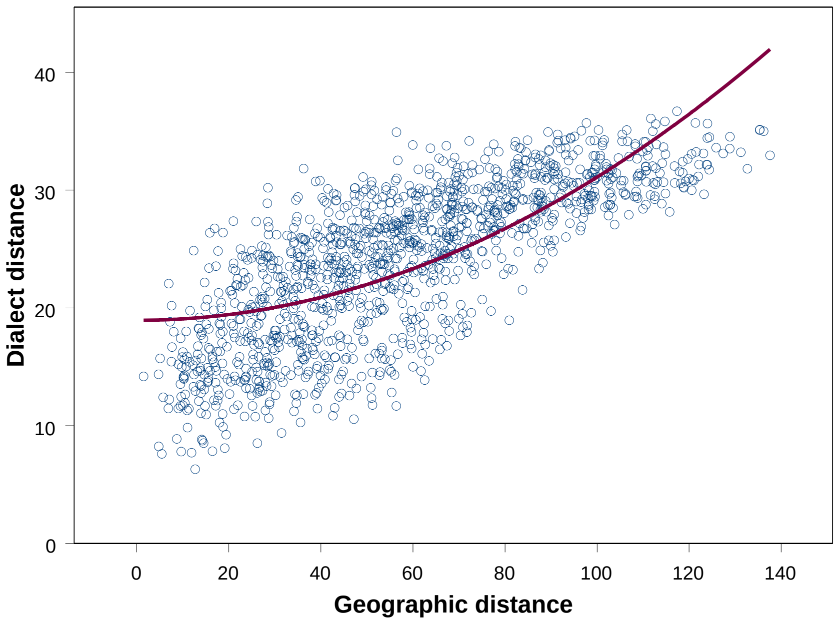

Figure 8: The linguistic distances of Lower Saxon data presented as a function of their geographic distance together with the optimal quadratic regression line. In fact there is little hint of a quadratic form in the scatter cloud.

5 Discussion

This paper has suggested a novel way of examining linguistic diffusion, viz., through synchronic measurements of linguistic similarity. We reason that the forces facilitating and impeding linguistic innovations should leave a residue of linguistic differences behind. The distribution of these differences betrays the dynamics which created them in a novel way, allowing us to examine the effects of diffusion without needing to probe ongoing changes.

The "gravity" model is not perfect in explaining the differences among dialects spoken in a certain area. There is indeed a positive correlation between dialect distance and geographic distance, but the curve does not have the predicted quadratic shape. Even more surprisingly, there is a slight positive correlation between dialect distance and combined population size (see Table 4). Together, these results suggests that the dominant effect in dialect geography is not one of attraction, but rather differentiation. The closer dialects are to one another, and the more people that are involved, the more strongly they generate and retain differentiating elements.

There are several qualifying remarks that need to be added to this conclusion. First, this emphatically does not mean that there are no gravity-like forces at work in dialect dynamics, only that they are not the strongest. The theoretical arguments establishing the plausibility of a gravity-like force derive from the need to accommodate to one's interlocutor, and this need is profoundly present in all human communication. But perhaps its effects are not lasting, and in any case they are not the strongest effects in the data we examined.

Second, we noted in Section 2.2 above that we would ignore Trudgill's "similarity" factor. It should have become clear that our experimental design, in which we crucially measured linguistic similarity as a putative result of gravity, could not also include similarity as an independent factor, at least not without complicating the analysis a great deal. It would be wrong to suspect that including similarity in the way Trudgill suggests

(Trudgill 1974, p.234) could alter the direction of the conclusion, however. In his model, similarity is postulated to promote diffusion. But since similarity correlates positively with geographical distance, its postulated effect should only strengthen the geographic one, which we have shown to be much weaker than postulated.

Third, and more generally, Wolfram & Schilling-Estes (2003, 726) criticizes the gravity model for abstracting away from too many influences which have been demonstrated to influence linguistic diffusion and retention. Wolfram and Schilling-Estes devote a good deal of discussion to the role which social networks may play, discussing in particular the work of Lesley Milroy and James Milroy (Milroy & Milroy 1985). We, too, have abstracted away from many of the forces well known in variationist linguistics, such as the effect of dialect prestige, social class, sex, and age on language varieties, relying on the one hand on the compilers of the RND to have controlled for those effectively, and focusing on a higher level of aggregation on the other. This seems reasonable, given the difficulty of obtaining data of this sort, but it also worth recalling it explicitly.

Fourth, we have not attempted to model the effects of trade in our geographic model, which we would expect to be important, and which depended largely on waterways. Incorporating the effect of trade would mean exploring the relative importance of routes over land versus over water. This would be a promising topic for future work. Our geographical modeling might also benefit from the incorporation of more detailed information concerning wetlands, which would have impeded transportation, but the fact that the location of roads (which avoid wetlands) was incorporated into the model means that information about wetlands was used indirectly.

Fifth, it is impossible to examine the mutual effect of two settlements in isolation from all others. Anytime we examine the effect of one settlement on another, we inevitably detect many effects whose causes remain obscure, but which undoubtedly involve the many other settlements. Perhaps a finer theoretical analysis can make sense of the "many body" version of linguistic attraction and differentiation, but we do not attack that problem here.

Sixth and finally, it is striking that Trudgill established the plausibility of a gravity model on the basis of individual features, while we have cast doubt upon it as a predictor of aggregate distance. It would be worth exploring how aggregate effects are related to effects among individual features. As we noted in above in this section, our negative conclusion about the predictive value of gravity models in aggregate dialect distance is compatible with the existence of gravity-like forces, but we have shown that such forces are not dominant.

We find the gravity model convincing in its intuitive justification, but unconvincing in the concrete test we put it to in the research report in this paper. We do not attempt to examine alternative models in this paper, but our basic methodology clearly supports extensive experimentation in mathematical modeling.

Acknowledgments

This research was made possible by a grant from the Dutch Organization for Scientific Research to the project "Determinants of Dialect Variation" (project 360-70-120, P.I. J.Nerbonne). We are indebted to Frits Steenhuisen and Elwin Koster for help with GIS and maps and to Pim Kooi for advice on obtaining historical demographic data. Renée van Bezooijem, Charlotte Gooskens, Bill Kretzschmar, Hermann Niebaum, Marco Spruit and Peter Trudgill commented generously on an earlier draft.

References

Bailey, Guy, Tom Wikle, Jan Tillery & Lori Sand. 1993. "Some Patterns of Linguistic Diffusion." *Language Variation and Change* 3(3):241–264.

Boberg, Charles. 2000. "Geolinguistic Diffusion and the U.S.-Canada Border." *Language Variation and Change* 12(1):1–24.

Bolognesi, Roberto & Wilbert Heeringa. 2002. "De invloed van dominante talen op het lexicon en de fonologie van Sardische dialecten." *Gramma/TTT: Tijdschrift voor Taalwetenschap* 9(1):45–84.

Callary, Robert E. 1975. "Phonological Change and the Development of an Urban Dialect in Illinois." *Language in Society* 4:155–169.

Clark, Herbert & Eve Clark. 1977. *Psychology and Language: An Introduction to Psycholinguistics.* New York: Harcourt Brace Jovanovich.

Freedman, David, Robert Pisani & Roger Purves. 1998. *Statistics.* New York: Norton. 3rd ed.

Giles, Howard. 1994. Accommodation in Communication. In *Encyclopedia of Language and Linguistics*, ed. Ronald E. Asher. Vol. I Oxford: Pergamon Press pp. 12–15.

Goebl, Hans. 1982. *Dialektometrie: Prinzipien und Methoden des Einsatzes der Numerischen Taxonomie im Bereich der Dialektgeographie.* Wien: Österreichische Akademie der Wissenschaften.

Goebl, Hans. 1984. *Dialektometrische Studien: Anhand italoromanischer, rätoromanischer und galloromanischer Sprachmaterialien aus AIS und ALF. 3 Vol.* Tübingen: Max Niemeyer.

Gooskens, Charlotte. 2004. Norwegian Dialect Distances Geographically Explained. In *Language Variation in Europe. Papers from the Second International Conference on Language Variation in Europe ICLAVE 2, June 12-14, 2003*, ed. Britt-Louise Gunnarson, Lena Bergström, Gerd Eklund, Staffan Fridella, Lise H. Hansen, Angela Karstadt, Bengt Nordberg, Eva Sundgren & Mats Thelander. Uppsala, Sweden: Uppsala University pp. 195–206.

Gooskens, Charlotte. N.d. "Traveling Time as a Predictor of Linguistic Distance." *Dialectologia et Geolinguistica.* Forthcoming. submitted, 3/2004.

Gooskens, Charlotte & Wilbert Heeringa. 2004. "Perceptual Evaluation of Levenshtein Dialect Distance Measurements using Norwegian Dialect Data." *Language Variation and Change* 16(3):189–207.

Heeringa, Wilbert. 2004*a*. "Dialect Variation in and around Frisia: Classification and Relationships." *Us Wurk; Tydskrift foar Frisistyk* 53(4).

Heeringa, Wilbert. 2004*b*. Measuring Dialect Pronunciation Differences using Levenshtein Distance PhD thesis Rijksuniversiteit Groningen.

Heeringa, Wilbert & Charlotte Gooskens. 2003. "Norwegian Dialect Examined Perceptually and Acoustically." *Computers and the Humanities* 37(3):293–315.

Heeringa, Wilbert & John Nerbonne. 2002. "Dialect Areas and Dialect Continua." *Language Variation and Change* 13:375–400.

Horvath, Barbara M. & Ronald J. Horvath. 2001. "A Multilocality Study of a Sound Change in Progress: The Case of /l/ Vocalization in

New Zealand and Australian English." *Language Variation and Change* 13(1):37–57.

Kessler, Brett. 1995. Computational dialectology in Irish Gaelic. In *Proc. of the European ACL*. Dublin: pp. 60–67.

Kretzschmar, William A., ed. 1994. *Handbook of the Linguistic Atlas of the Middle and South Atlantic States*. Chicago: The University of Chicago Press.

Labov, William. 2001. *Principles of Linguistic Change. Vol.2: Social Factors*. Malden, Mass.: Blackwell.

Lewis, David K. 1979. "Scorekeeping in a Language Game." *Journal of Philosophical Logic* 8(3):339–359.

Mantel, N. 1967. "The Detection of Disease Clustering and a Generalized Regression Approach." *Cancer Research* 27:209–220.

Milroy, James & Lesley Milroy. 1985. "Linguistic Change, Social Network and Speaker Innovation." *Journal of Linguistics* 21:339–84.

Nerbonne, John & Christine Siedle. 2005. "Dialektklassifikation auf der Grundlage Aggregierter Ausspracheunterschiede." *Zeitschrift für Dialektologie und Linguistik* . accepted for publication.

Nerbonne, John & Peter Kleiweg. submitted, 12/2003. "Toward a Dialectological Yardstick." *Journal of Quantitative Linguistics* . Paper presented at the *Conference on Quantitative Linguistics*.

Nerbonne, John, Wilbert Heeringa, Erik van den Hout, Peter van der Kooi, Simone Otten & Willem van de Vis. 1996. Phonetic Distance between Dutch Dialects. In *CLIN VI: Proc. from the Sixth*

CLIN Meeting, ed. Gert Durieux, Walter Daelemans & Steven Gillis. Antwerpen: Center for Dutch Language and Speech, University of Antwerpen (UIA) pp. 185–202. Also avail. as `http://www.let.rug.nl/~nerbonne/papers/clin96.pdf`.

Nerbonne, John, Wilbert Heeringa & Peter Kleiweg. 1999. Edit Distance and Dialect Proximity. In *Time Warps, String Edits and Macromolecules: The Theory and Practice of Sequence Comparison, 2nd ed.*, ed. David Sankoff & Joseph Kruskal. Stanford, CA: CSLI pp. v–xv.

Nerbonne, John & William Kretzschmar. 2003. "Introducing Computational Methods in Dialectometry." *Computers and the Humanities* 37(3):245–255. Special Iss. on Computational Methods in Dialectometry ed. by John Nerbonne and William Kretzschmar, Jr.

Ramaer, Johan Christoffel. 1931. *Geschiedkundige Atlas van Nederland; Het Koninkrijk der Nederlanden 1815-1931.* The Hague: Nijhoff.

Razavi, Amir H. 1997. *ArcView GIS / Avenue Developer's Guide.* Clifton Park, New York: OnWord Press.

Schmidt, Johannes. 1872. *Die Verwandtschaftsverhältnisse der indogermanischen Sprachen.* Weimar: Böhlau.

Séguy, Jean. 1971. "La relation entre la distance spatiale et la distance lexicale." *Revue de Linguistique Romane* 35:335–357.

Trudgill, Peter. 1974. "Linguistic Change and Diffusion: Description and Explanation in Sociolinguistic Dialect Geography." *Language in Society* 2:215–246.

Trudgill, Peter. 1983. *On Dialect. Social and Geographical Perspectives.* Oxford: Blackwell.

Trudgill, Peter. 1986. *Dialects in Contact*. Oxford: Blackwell.

van Gemert, Ilse. 2002. Het geografisch verklaren van dialectafstanden met een geografisch informatiesysteem (GIS). Master's thesis Rijksuniversiteit Groningen. avail. `www.let.rug.nl/alfa/scripties.html`.

Wikle, Thomas & Guy Bailey. 1997. "The Spatial Diffusion of Linguistic Features in Oklahoma." *Proceedings of the Oklahoma Academy of Science* 77:1–15. avail. at `digital.library.okstate.edu/OAS/`.

Wolfram, Walt & Natalie Schilling-Estes. 2003. Dialectology and Linguistic Diffusion. In *The Handbook of Historical Linguistics*, ed. Brian D. Joseph & Richard D. Janda. Malden, Massachusetts: Blackwell pp. 713–735.