On the relation between structural diversity and geographical distance among languages: Observations and computer simulations

ERIC W. HOLMAN, CHRISTIAN SCHULZE, DIETRICH STAUFFER, and SØREN WICHMANN

Abstract

Modern linguistic typology is increasingly less concerned with what is possible in human languages (universals) and increasingly more with the question "what's where why?" (Bickel 2007). Moreover, as several recent papers in this journal show, typologists increasingly turn to quantitative approaches as a means to understanding typological distributions. In order to provide the quantitative study of typological distributions with a firm methodological foundation it is preferable to gain a grasp of simple facts before starting to ask the more complicated questions. In this article the only assumptions we make about languages are that (i) they may be partly described by a set of typological characteristics, each of which may either be found or not found in any given language; that (ii) languages may be genealogically related or not; and that (iii) languages are spoken in certain places. Given these minimal assumptions we can begin to ask how to express the differences and similarities among languages as functions of the geographical distances among them, whether different functions apply to genealogically related and unrelated languages, and whether it is possible to distinguish in some quantitative way between languages that are related and languages that are not, even when the languages in question are spoken at great distances from one another. Moreover, we may investigate the effects that factors such as ecology, migration, and rates of linguistic change or diffusion have on the degree of similarities among languages in cases where they are either related or unrelated. We will approach these questions from two perspectives. The first perspective is an empirical one, where observations primarily derive from analyses of the data of Haspelmath et al. (eds.) (2005). The second perspective is a computational one, where simulations are drawn

Linguistic Typology 11 (2007), 393–421 1430–0532/2007/011-0393
DOI 10.1515/LINGTY.2007.027

upon to test the effects of different parameters on the development of structural linguistic diversity.

Keywords: *areal typology, computer simulation, isolation by distance, spatial autocorrelation*

1. Introduction

This article investigates the relationship between typological similarity and geographical distance among languages. Because of the effect of diffusion it is expected that we should observe the phenomenon known, e.g., among ecologists, as SPATIAL AUTOCORRELATION or, among population geneticists, as ISOLATION BY DISTANCE (IBD), that is, a relation where increased geographical distance correlates with greater differences – in this case among languages. This relationship is expected to obtain whether or not the languages in question are related. Nevertheless, while spatial autocorrelation is expected to occur universally, the effect might be enhanced or diminished by different factors such as initial similarity among languages vs. initial dissimilarity, features of the physical environment, rates of diffusion, rates of internal language change, speed of migration, and perhaps other factors. How do these various factors affect the relationship between structural similarity and geographical distance? For instance, should we expect languages that are related to be as different from one another as languages that are not related, even at great geographical distances?

Another set of questions which this article will be concerned with is to what extent qualitatively different typological datasets may have an effect on our ability to discern the footprints of language history. More specifically, we will investigate whether it makes a difference whether one uses binarily as opposed to ternarily or quarternarily etc. encoded features and we will also look at whether or not it has an effect on one's results whether the values of the features used stand in a graded relationship to one another or not (illustrative examples and more clarification of this issue will be given below).

In the following section the problem area is introduced by means of an overall comparison of structural diversity among related vs. unrelated languages of the world. The remainder of the article presents both empirical data and various computer simulations serving to investigate the impact of different factors on linguistic diversity. The impact of ecological factors is best studied empirically since it is very complicated to model geographical factors in a simple and adequate fashion. On the other hand, factors such as rates of diffusion and speed of migration are hard to study empirically in a systematic way because we cannot compare several different areas where such factors are known to differ, given that we have too little exact knowledge about prehistory. Moreover, it is impossible to tease different factors apart so as to study their individual impacts. Using computer simulations, however, we can hold certain parameters constant and vary others so as to study the effect of each. The effects of rates of internal language change are studied both empirically and by means of simulations.

Methodologically, the computer simulations to be presented are the centerpiece of our article. To our knowledge, computer simulations have never before been used to answer the specific questions raised here regarding the relationship between structural differences and geographical distances. Nevertheless, a small but growing number of researchers have drawn upon computer simulations and mathematical models to investigate other aspects of linguistic evolution, including the development of linguistic diversity (Abrams & Strogatz 2003, Sutherland 2003, Patriarca & Leppänen 2004, Mira & Paredes 2005, Schulze & Stauffer 2005, Wang & Minnett 2005, Kosmidis et al. 2005, Schwämmle 2005, Oliveira et al. 2006a, b, Pinasco & Romanelli 2006), the development of taxonomic dynamics (Wichmann, Stauffer, et al. 2007), language change (Nettle 1999a, Niyogi 2002, Prévost 2003, Baxter et al. 2006), and the evolution of language structure (Cangelosi & Parisi (eds.) 2002, Nowak et al. 2002, Christiansen & Kirby (eds.) 2003, Wang, Ke, & Minett 2004, de Boer 2006, Niyogi 2006). Without looking at geographical distances, Itoh & Ueda (2004) analyzed differences between many Eurasian languages and made a computer simulation similar to the Ising model well known in Statistical Physics. As regards the investigation of empirical correlations between linguistic differences and geographical distance there are precursors in the field of dialectometry, which was initiated by Séguy (1973) and developed further in many subsequent works, including Goebl (1984, 2005), Nerbonne et al. (1996), and Kretzschmar (1996). Finally, Cavalli-Sforza & Wang (1986) studied lexical similarities as a function of geographical distance for some Micronesian speech communities.

2. The relation between structural diversity and geographical distance among related vs. unrelated languages

The data drawn upon in this section are provided by *The World Atlas of Language Structures* (Haspelmath et al. (eds.) 2005, henceforth *WALS*). *WALS* contains 138 maps showing the distribution of different phonological, lexical, and grammatical features for a sample of languages that varies in size among maps from roughly 100 to 1,200. The present study draws on 134 of the 138 features, excluding features that involve redundant data. Each feature has anywhere from two to nine discrete values. The total number of languages from which data are drawn in *WALS* is 2,560. The present study excludes pidgins, creoles, and sign languages, leaving 2,488 languages. The genealogical classification used in *WALS*, which we also adopt here, represents an attempt to follow the views of the majority of specialists and results in 205 families and isolates. Reflecting this consensus, languages in the same family are here called related, and languages in different families are called unrelated.

Many of the *WALS* features are correlated in the sense that the presence of a certain value of one feature to some degree implies the presence of a certain value of another feature – a well-known phenomenon in typology. Some readers may protest that it is problematical to use all 134 non-redundant features of *WALS* without taking into account the fact that some are correlated. But upon closer reflection this ceases to be a problem because we are averaging over many comparisons of language pairs in this study. We can imagine that if we compared one pair of languages for which only five correlated word-order features were available with another pair of languages for which five uncorrelated features were available, the first pair of languages could potentially stand out as falsely being more similar than the second. But the present study is not concerned with comparisons of individual pairs of languages, only with similarities averaged over hundreds of pairs of languages, so we should not expect to see any effect on our results due to statistical implicational relations. Indeed, drawing upon Holman (no date) we have redone our different investigations using a reduced sample of 47 mutually uncorrelated features with no noticeable changes in results, but we prefer to present investigations using all 134 features in order not to create additional problems of scarcity of data.

The typological difference between languages as a function of the geographical distance between them was measured in the following way. For each pair of languages, their distance was calculated from the latitudes and longitudes in the *WALS* database, where the location of each language is defined as a spot somewhere near the center of the region where the language is spoken (see Comrie et al. 2005: 7 for more detail). Pairs of languages were then grouped according to distance in ranges such as 0–500 km, 500–1,000 km, 1,000–2,000 km, etc. For each of the 134 features, the average difference between the paired languages in a group was defined as the number of pairs with different values of the feature, divided by the number of pairs for which feature was attested in both languages. These proportions were averaged across the 134 features and expressed as a percentage to represent the overall difference of the language pairs in a group. Figure 1 plots difference as a function of mean distance for respectively related and unrelated languages in the world. The curve for related languages does not extend as far as the curve for unrelated languages because few language families contain languages separated by very large distances.

We realize that the effect of treating languages as points in space is to inflate the calculated distances from large languages to others. To determine whether this effect distorts the curves in Figure 1, we repeated the analysis, excluding languages with more than 5,000,000 speakers; the resulting figure is practically indistinguishable from Figure 1. The effect of large languages may be more prominent in Eurasia, however, where most of the 50 or so largest languages in

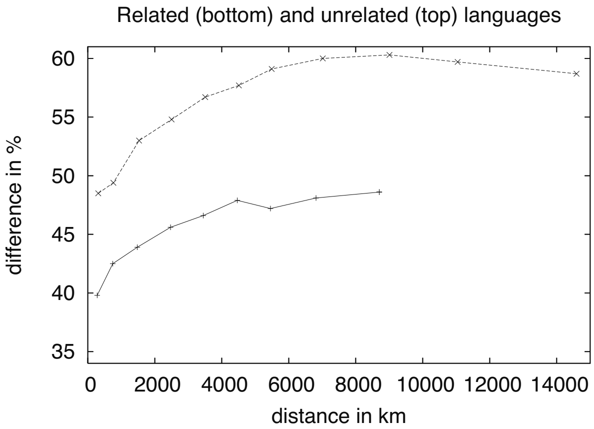

Figure 1. *The relationship between geographical distance in km and the amount of structural differences among related and unrelated languages in the world*

the world are spoken. Even in this area the effect would produce marked distortion only in the 0–2,000 km range, where languages which are in fact neighbors are treated as being far apart. Thus, in Eurasia languages are expected to falsely stand out as more similar at given short distances than languages in other parts of the world. This is the only problem caused by our approach. It should be kept in mind when comparing regions (as we, in fact, do towards the end of this article), but it can be ignored when averaging over the world's languages, as we do throughout most of this article.

Both curves in Figure 1 provide clear evidence of spatial autocorrelation. Difference is least between languages less than 1,000 km apart, and then increases with increasing distance. For languages more than about 5,000–6,000 km apart, the curves appear to approach asymptotes. Bootstrap tests, reported on in Holman et al. (no date), indicate a significant effect of distance as well as a significant difference between the curves. Spatial autocorrelation is a phenomenon that has been studied within different disciplines under different labels. For instance, within population genetics the equivalent term "isolation by distance" (IBD) was defined by Wright (1943) and refers to situations where large genetic correlations are found among spatially proximal populations and drop off smoothly as the distances among the populations decrease. Discussion and examples are found in Epperson (2003: 14–25). One illustrative example is the study by Sokal & Menozzi (1982) of different allele frequencies for HLA blood group loci in European and Middle East populations. Large autocorrelations were found within a range of approximately 700 km, lesser positive values within an approximate 700–1,400 km range and negative values beyond this range. A notion from geography which is similar to IBD in biology is "Tobler's First Law of Geography", which states that "everything is related to everything else, but near things are more related than distant things" (Tobler 1970). For linguistics, Nerbonne & Kleiweg (2007) define what they call the "fundamental dialectological postulate" as follows: "Geographically proximate varieties tend to be more similar than distant ones". Their Figure 1, which is very similar to our Figure 1, illustrates how differences among dialects depend on distance (similar curves have been published by Séguy 1971 and Goebl 2001). This principle or postulate, then, is equivalent to IBD. For our purposes the term "spatial autocorrelation" seems to be the most appropriate since it is the most general one, being used across several disciplines.

Up to this point we have demonstrated that typological similarities among languages, whether the languages are related or not, tend to increase with decreased distance, and we have related this phenomenon to the cross-disciplinarily well-known notion of spatial autocorrelation. Our interpretation is that the autocorrelations found are due to diffusion. We see that at around 8,000 km the curves have flattened out, and interpret this as the point at which effects of diffusion can no longer be seen. Another result obtained is that related languages on average are more similar to one another than unrelated languages at any degree of geographical distance (for individual pairs of languages this does not necessarily hold true, cf. Wichmann & Holman 2007). In the remainder of this paper we will investigate different additional factors that may affect spatial autocorrelations among languages.

3. Differences in spatial autocorrelations and their possible causes: Intuitive/common-sensical explanations

The example in Figure 1 shows a difference in spatial autocorrelation. In this case the responsible factor must be relatedness vs. unrelatedness of the languages compared. We might, however, imagine that other factors could affect the curves such that languages, whether related or not, become either more or less similar at given geographical distances, pushing the curves down or up.

Ecological differences. Possibly certain features of geography are conducive to the enhancement of similarities among languages, whether those similarities are due to diffusion or inheritance. Nettle (1999c) has pointed out that there is a correlation between the amount of annual rainfall and linguistic diversity in terms of the amount of different genealogical lineages. He suggests that an inverse correlation between genealogical diversity and ecological risk holds because in areas with greater ecological risk people need to establish larger networks of exchange in order to mitigate this risk, a situation which disfavors linguistic diversity. Earlier, Johanna Nichols had observed that "high genetic diversity is evidently favored by coastline, tropical to subtropical climate, and (at least in some cases, such as the Caucasus and Himalayas) mountains" (Nichols 1992: 233–234). While these correlations may, indeed, hold between GENEALOGICAL (linguistic genetic) diversity and geography, it is not given that geography and STRUCTURAL diversity are similarly correlated, such that we should expect to see differences in curves for spatial autocorrelations depending on ecological factors. This is worth investigating. The two factors relating to coastal vs. inland and tropical vs. non-tropical environments are relatively easy to investigate empirically in a systematic fashion, but it is more difficult to examine the impact of highlands vs. lowlands, the main problems being the following: information about whether given languages are spoken in highlands or lowlands is often not available; it is difficult to know where to draw a boundary between what counts as highland and lowland; and one and the same language – particularly a language having a large number of speakers – may be spoken both in highlands and lowlands. Thus, below we examine the two first factors only: coast vs. inland and tropical vs. non-tropical.

As mentioned in Section 1, some factors are better investigated using computer simulations than drawing upon empirical data, simply because the relevant empirical data cannot be isolated and accessed in any systematic way. They include the following.

Differences in genealogical diversity. If a large area was populated in a single wave of migration, as has been argued from genetic evidence to be the case for the Americas (most recently Stone & Stoneking 1998, Silva Jr. et al. 2002, Tarazona-Santos & Santos 2002, Zegura et al. 2004), then all the languages of that area might share a single ancestor or descend from relatively similar languages. It seems likely that the languages in such an area could be more similar than would generally be the case in other parts of the world. We have already seen (Figure 1) that related languages are more similar, even at large geographical distances, than unrelated languages. But we cannot know whether this effect is expected to hold at great time depths such as would correspond to a putative "Amerind" since no such families have been established with certainty. Therefore this situation has to be simulated.

Rate of diffusion. Possibly languages in some parts of the world show less diffusion (i.e., less transfer of features from one language to the other) than languages of other parts due to differences in patterns of subsistence, preferred marriage patterns, ecological barriers, or other.

Rate of migration. A fast migration of people speaking the same language into a territory which was previously not occupied might produce a situation with small differences even among languages spoken very far apart. On the other hand, speakers would quickly cease to be in contact, and the absence of diffusion would produce an effect in the opposite direction of more differences. Thus, it is difficult to predict the effect of speed of migration.

Rate of language change. Little is known about the rate of change contributed by sheer internal restructuring. The heavy focus on contact-induced change in the literature on diachronic typology may distract one from considering changes that are due to sheer internal changes. If changes in the typological profile of a language were mostly due to contact we would, over a span of many tens of thousands of years of language evolution, see all languages become similar to one another and would not expect the present diversity. So we should consider the possible impact of the rate of internal language change on linguistic spatial autocorrelation.

These various thought experiments lead us to doubt that there will be simple and monolithic explanations for differences among linguistic spatial autocorrelations. Instead, they suggest that a variety of factors could affect the distributions. Factors that pull in the same direction may be difficult to tease apart, and it may be difficult to judge what the relative impacts of factors that pull in opposite directions are. In the following we therefore investigate each factor separately.

4. Investigating ecological factors using empirical data

In this section we look at the two sets of contrasting factors that involve ecology and are amenable to reasonably systematic investigations: coast vs. inland and tropical vs. non-tropical languages. As a guide for locating each language we used the geographical coordinates provided in the *WALS* database. These coordinates roughly correspond to the geographical center of the extension of each language and were produced by the editors for the purpose of placing dots on the *WALS* maps. Obviously, for larger languages it is problematical to identify their location with a single dot. Thus, for classifying languages in coastal vs. inland we have added some extra criteria (see below). For classifying languages as tropical or non-tropical, however, we used the given coordinate as the sole criterion; this means that for languages spoken on both sides of the Tropic of Cancer or the Tropic of Capricorn their categorization is somewhat arbitrary.

4.1. *Coast vs. inland*

The definition of which languages are spoken in coastal ranges requires a certain degree of judgment. We therefore limited the investigation to all languages in *WALS* for which 45 or more features are attested. We have defined a coastal language as one which is spoken within 100 km of the ocean (the Black Sea is excluded, the Baltic included). For national languages we required that a long strip of coast pertains to the country. Thus, for instance, German is classified as inland. Such large languages are a minority in the dataset. When we had serious doubts about how to categorize a language we simply excluded it. The resulting sample consists of 167 coastal languages and 154 inland ones.

Figure 2 shows spatial autocorrelations for related vs. unrelated languages, as in Figure 1, but now we further subdivide each into coastal vs. non-coastal. The criss-crossing curves suggest that there is no appreciable difference in the behavior of coastal vs. non-coastal languages.

4.2. *Tropical vs. non-tropical*

The tropics are standardly defined as the region in latitude by the Tropic of Cancer in the northern hemisphere, at approximately 23.5° N latitude, and the Tropic of Capricorn in the southern hemisphere at 23.5° S latitude. Languages whose *WALS* coordinates lie within the tropics are classified as tropical, the rest as non-tropical.

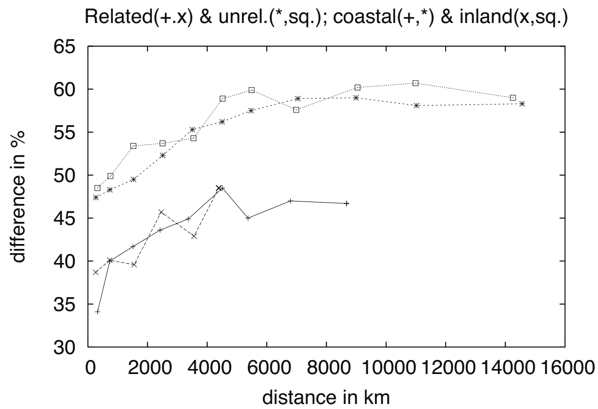

Figure 2. *The relationship between geographical distance in km and linguistic differences among related vs. unrelated coastal vs. inland languages in the world*

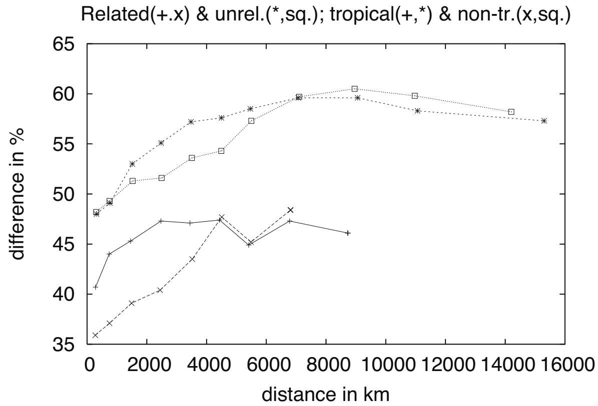

Figure 3. *The relationship between geographical distance in km and linguistic differences among related vs. unrelated tropical vs. non-tropical languages in the world*

Figure 3 shows the behavior of related vs. unrelated languages subdivided into tropical vs. non-tropical. There may be some differences here, tropical languages being somewhat more different than non-tropical ones, but only at relatively short distances. The 4,000–6,000 range where the differences break down would correspond to the maximal east-west extension of continents. It is not clear whether the differences are really significant, but it is tempting to relate the possible differences to Nettle (1999c). Nettle discusses the greater amount of linguistic diversity in tropical areas. His criterion for diversity is different from ours since Nettle simply looks at diversity in the sense of how many different language families are represented in different areas, whereas we look at how different the languages are typologically. Nevertheless, his and our observations concur inasmuch as what leads to genealogical diversity must be the tendency for languages to drift apart structurally. We may be observing a tendency for tropical languages to drift apart more rapidly than non-tropical ones, but given the absence of statistical tests to support this hypothesis we will not take this any further. It would be imprudent to try to explain an observation which may not even be correct. For the moment we will restrict ourselves to state as a hypothesis that there may be a slight tendency for tropical languages confined to a given continent to be more different than non-tropical ones within the same continent. Future work may reveal whether this observed tendency is significant.

5. Investigating other factors using computer simulations

In this section we look at further factors that might possibly affect patterns of linguistic spatial autocorrelation. First we briefly present a computational model suitable for such an investigation and subsequently we present the results of the implementation of this model.

5.1. *Brief description of the computational model*

We model population expansion on a square lattice. Initially only the top line is occupied, but the lattice is gradually populated as new speakers are born. The people in the top line either all speak the same language or each one speaks a randomly selected language. Language dynamics is simulated by allowing for the following four processes, each of which happens with a certain prespecified probability. People may SHIFT to a neighboring language with a probability which decreases with increasing size of the original language. Languages may CHANGE with a certain prespecified probability. Languages may suffer the effects of DIFFUSION and people may MIGRATE. The language model used is one where a language is characterized by a certain number of FEATURES, which can be imagined to represent distinctive typological features. These features can have two or more VALUES. All in all there are six variable parameters in this model. The settings used are described in the following sections. A more detailed description of the computational model is provided in Appendix 1.

5.2. *Results of the implementation of the model*

5.2.1. *Introduction.* A simulation using standard parameter settings is shown in Figure 4. Here and in the following figures we show the result after 70 iterations, when the shapes of the curves no longer change. Only the topmost 10 lines of the $10,001 \times 10,001$ lattice are analyzed, comparing pairs of sites with the first one in the top line and the second one at distance d exactly below the first. (The values remain constant also for longer distances up to 30.) We start from many (+) or one ancestor (x). (For many ancestors, the results cease to change already after 5 iterations.) For larger diffusion $q \simeq 1$ instead of 0.9, the initially fragmented population would have changed later into one dominated by only one language, cf. Schulze & Stauffer (2006) and Stauffer, de Oliveira, et al. (2006). We assume that features are ordered and measure differences between languages accordingly (see Section 6 below for more detail).

The standard parameter settings, then, are:

internal language change: $p = 0.5$
diffusion: $q = 0.9$
language shift: $r = 0.9$

migration: $s = 0.5$
number of features: $F = 8$
number of feature values (states, choices): $Q = 5$

Since at present we do not know how to translate empirical data into expected absolute probabilities for language change, diffusion, language shift, and migration, all the values that we operate with should be looked upon as highly abstract. The same goes for the results of measuring structural differences and geographical distances. Nevertheless, a comparison of the empirical data with the results of the computer simulations suggests that a simulated distance of 2 roughly corresponds to a real-world distance of 5,000 km. As suggested by Figure 1, this is roughly the distance at which the inverse correlation between distance and structural similarity among languages of different families begins to cease being discernable. With one exception (see Figure 6) all the simulations for the corresponding distance of ≈ 2 similarly show a weakening of the correlation.

An important result of the simulations shown in Figure 4 is that a difference between having many ancestors and having one continues to be preserved over long distances, even if this difference is diminished somewhat. This would mean that given two situations where all else is equal, we may be able to distinguish between languages sharing a common ancestor and unrelated languages by means of typological data. As we will see, this "preservation of history" does not result from all parameter settings, but it is the rule rather than the exception.

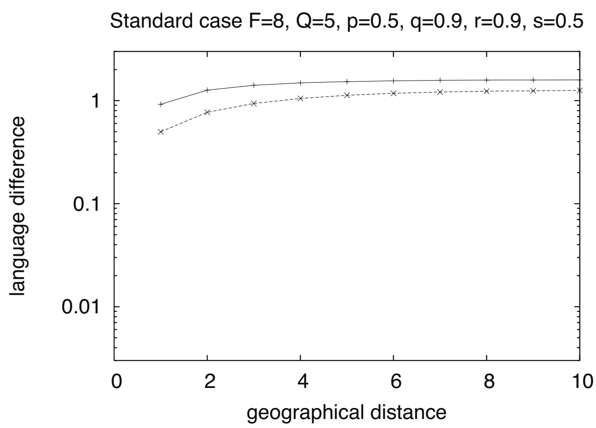

Figure 4. *A simulation using standard parameter settings. Legend: + = many ancestors, x = one ancestor.*

In the following we vary the settings to study the effects of each individual parameter.

5.2.2. *Rate of diffusion.* Decreasing the rate of diffusion makes initially related languages more different, but the effect of such a decrease diminishes with the distance. There is hardly any effect to be seen for the situation with many ancestors, cf. Figure 5. That is, diffusion has a greater effect on the degree of similarities among related than among unrelated languages. This is a highly interesting result which is initially somewhat counter-intuitive, but it can nevertheless be brought to accord with the real-world situation. We know that diffusion is a highly potent force in language change. Nevertheless, we also see an immense structural diversity among the world's languages. What could explain this apparent discrepancy is that whereas diffusion may make languages that are in contact with one another more similar, these regional similarities contribute to inter-regional diversity. It is possible to change from a situation where the population is distributed equally over many different languages to one where more than half of the population speaks a single language, but that requires the diffusion rate to be 0.999, as witnessed by simulations not shown here.

5.2.3. *Rate of migration.* We have tried to vary the migration probability to 0.1 and let the simulation run for 200 iterations to get stationary results. This has virtually no effect. The curves are so similar that no difference can be made out visually although there are minor differences in the data.

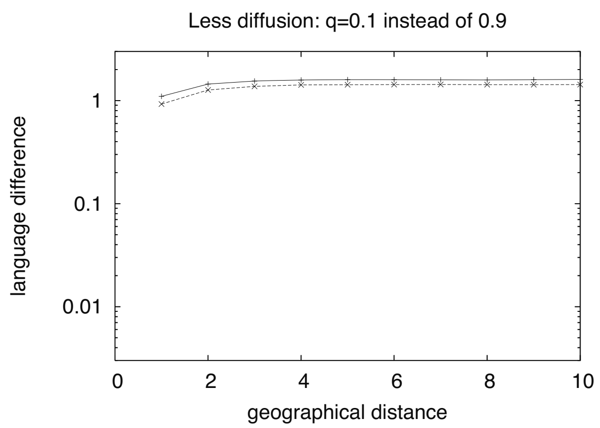

Figure 5. *The effect of varying diffusion rate from 0.9 (Figure 4) to 0.1 (this figure). Legend: + = many ancestors, x = one ancestor.*

5.2.4. *Rate of language shift.* A simulation where the rate of language shift was varied from 0.9 to 0.5 showed that this has no effect in the case of many ancestors. Although more language shift should reduce the number of languages, it does not affect the overall structural diversity as measured in the total number of differences among language pairs. It stands to reason that in the extreme case where the world's population was divided up into, say, speakers of Chinese and English, the differences among these two languages might still correspond to the average differences among the current languages of the world. The rate of language shift does seem to have a small effect in the case of a single ancestor. With more shift, the offspring of an initially uniform language become more similar, but this effect evaporates at large distances. Again, this is intuitively obvious. As in the case of the standard parameter settings the curves continue to be distinct at large distances. Given the high degree of similarity with Figure 4 we do not show the graph here.

5.2.5. *Rate of language change.* Comparing Figures 4 and 6 shows that reducing the rate of language change has no effect when we start with many ancestral languages, but has a drastic effect when there is only one. In a situation of many ancestors the effects of language changes will tend to cancel one another out. In a situation of just one ancestor, internal language change is a major contributor to diversity.

We have tested whether a similar result is obtained when empirical data are used. It ought to be the case that two curves for spatial autocorrelation among

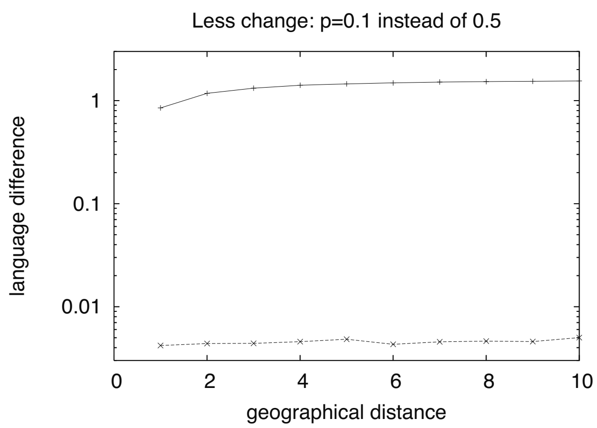

Figure 6. *The effect of varying the rate of language change from 0.5 (Figure 4) to 0.1 (this figure). Legend: + = many ancestors, x = one ancestor.*

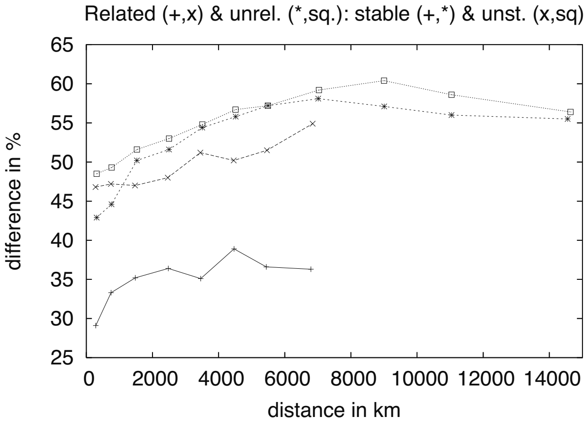

Figure 7. *Spatial autocorrelation among related vs. unrelated languages in a world-wide sample with respect to stable vs. unstable features*

related languages are far apart when one curve describes the number of differences among languages with respect to features that are highly stable and another describes the number of differences with respect to highly unstable features. On the other hand, for unrelated languages there should be no obvious differences in the corresponding curves for stable and unstable features. In Wichmann & Holman (no date) we test the performances of four different stability metrics for typological features. The most reliable of these is briefly described in Appendix 2 below. After having applied this metric to the *WALS* data we selected the 33 most stable features and the 33 most unstable ones (see Appendix 3 for the lists of features) and then, for each set of features, calculated the differences among respectively related and unrelated languages. Figure 7, which should be compared to Figures 4 and 6, shows that the empirical data conform to the general predictions of the simulations. The average differences among related languages with respect to highly stable features is appreciably larger – 10 % or more at any geographical distance – than the differences with respect to highly unstable features. For unrelated languages, however, the curves are largely similar. If the curve for stable features tends to show slightly fewer differences, this might be attributed to the relatedness among some languages assumed not to be related in the *WALS* classification, but this effect is so small as to barely be noticeable.

6. Fingerprints of language history seen through qualitatively different datasets

6.1. *Introduction*

The data from *WALS* that were used to produce the examples in Figure 1 involve a mixture of qualitatively different encodings. Features can have from 2 to 9 values and for some there is an internal relationship among these values while for others there is no such relationship.

An example of a binary feature is the presence vs. absence of future as an inflectional category (Dahl & Velupillai 2005). The linguistic typological database developed at the Max Planck Institute for Psycholinguistics in Nijmegen, which was drawn upon by Dunn et al. (2005), consists exclusively of binary features. An example of a *WALS* feature having nine values is the way that the plural category of nouns is expressed (Dryer 2005). Thus, the plural can be expressed by means of (1) a prefix, (2) a suffix, (3) stem change, (4) tone, (5) a mixture of the preceding, (6) reduplication, (7) a separate word, or (8) a clitic; finally, (9) some languages do not have a nominal plural category.

When there is an internal relationship among the values we use the term "ordered feature". An example of an ordered feature is the inventory of vowel qualities (Maddieson 2005). In Maddieson's formulation of this feature, 2–4 vowels count as "small", 5–6 as "average", and 7–14 as "large". Thus there are three values. Presumably a language normally does not change its inventory of vowels directly from small to large but has to pass through a stage where the inventory is average. The same would hold for a change in the opposite direction. This feature, then, is (probably) ordered. The expression of the plural, in contrast, is largely unordered in the sense that a language can mostly change from having any of the possible values to any other. For instance, a language can just as easily go from not having a plural to having either a prefix or a tone expressing the plural. We say "largely unordered" because a language may be unlikely, for instance, to go from having no plural to having a mixture of different types of plural (value 5). It is typical for the *WALS* features that it is not always easy to decide whether they are ordered, semi-ordered, or unordered. The feature involving velar nasal consonants (Anderson 2005) may serve as an example of this indeterminacy. Some languages do not have a velar nasal, others may have such a sound but not in the beginning of words (as in English), and yet others may allow a velar nasal in the beginning of the word (this is common in, for instance, languages of Africa, South-East Asia, or aboriginal Australia). For languages that allow final consonants it is rare to find cases where an initial velar nasal is allowed but not a final one. Thus, if a language goes from not having a velar nasal to allowing velar nasals word-initially we might expect that in most cases it would first pass through a stage where it only had word-final velar nasals (that is, if the language allows final consonants). But this is far from certain and would require more investigation. The example, then, serves to illustrate that classifying *WALS* features into ordered vs. unordered is an idealization. In reality, order is a matter of degree that would have to be determined by studying how often any pair of values of a given feature co-occur in genealogical language groups with a short history of differentiation. By this method, which has been suggested by Maslova & Nikitina (no date), we could develop an idea about the changes among feature values that are more likely to take place and thus determine to what degree a feature is ordered. To produce the various plots of empirical data in this article the *WALS* data were treated as unordered.

In the present context we are interested in studying the effects of different numbers of feature values on the outcome of historical linguistic investigations and we also want to know more about the effects of ordered vs. unordered features. For the investigation of ordered vs. unordered features we must necessarily assume that a clean distinction between the two can be made, that is, we assume that the features used are of an ideal type.

For an unordered feature any difference in values will count as 1, whereas for an ordered feature a difference is counted as the absolute difference between the two values. For instance, for the feature of vowel inventory sizes mentioned above, a difference between value 1 (small inventory) and value 3 (large inventory) counts as 2 differences, whereas a difference between value 1 and 2 or value 2 and 3 both count as 1 difference.

6.2. *Ordered vs. unordered features*

Given that ordered features encode more information it is to be expected that such features will preserve the evidence for initial relatedness better than unordered ones. Figure 8 is nevertheless an important demonstration of the validity of these expected results. It should be compared to Figures 4 to 6, but in particular to Figure 4, which has the same parameter settings as Figure 8 except that Figure 4 involves ordered features. We see that the curves end up meeting one another, which is an important difference compared to Figures 4 to 6.

6.3. *Features having two vs. more than two values*

The different behavior of binary as opposed to many-valued features is similar to that of unordered vs. ordered features, in fact so similar that there is no need to provide an additional illustration. Just as in Figure 8, the curves for many ancestors and one ancestor narrow in on one another for $Q = 2$ and end up meeting each other. This is an important difference against Figure 4 and the results of other simulations in this paper: in the other simulations the curves never meet. Thus, it will get harder to discern a difference between related and

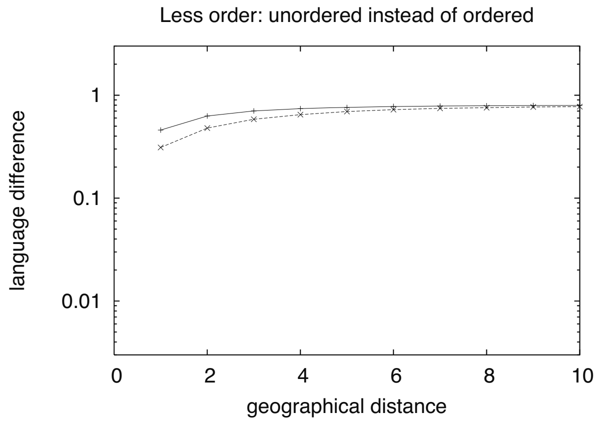

Figure 8. *The effect of using unordered (this figure) as opposed to ordered features (Figures 4 to 6). Legend: + = many ancestors, x = one ancestor.*

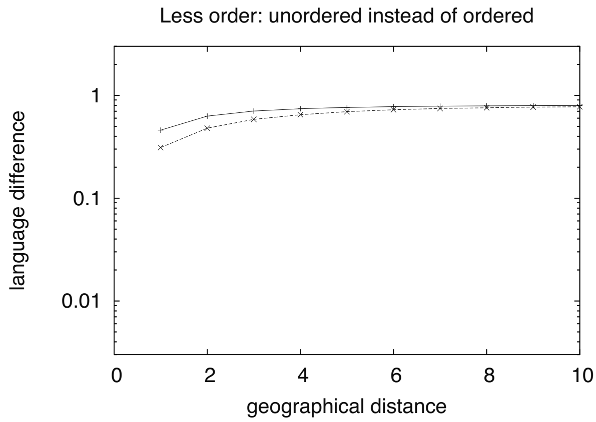

Figure 9. *The effect of using binary features (here) as opposed to many-valued ones (Figure 4). Legend: + = many ancestors, x = one ancestor.*

unrelated languages in binary descriptions. For $3 \leq Q \leq 9$ we merely see a parallel shift since the absolute differences may increase for increasing Q as opposed to binary ones. Thus, the curves for $Q = 3$ and $Q = 9$ are similar to those for $Q = 5$ shown in Figure 4, but the curves for $Q = 3$ lie lower than the curves for $Q = 5$ and those for $Q = 9$ lie higher (not shown).

6.4. *Conclusions regarding different results for different encodings of data*

Sections 6.2 and 6.3 have shown that diffusion, language shift, and migration cannot cause languages that share a common ancestor to look typologically as dissimilar as unrelated languages at large geographical distances when we average over many languages. (As shown by Wichmann & Holman 2007, a single pair of related languages, for instance, within Indo-European, Irish and Marathi, may be typologically as dissimilar as the majority of unrelated language pairs, but in the present article we are looking at the behavior of whole groups of languages.) Conceivably a fast rate of internal language change and a lack of diffusion among some related languages could make these languages become mutually more different than a group of unrelated languages having a slow rate of change and much diffusion. But even with such a radical tweaking of parameters we have not been able to produce simulation results where the curves for a single vs. many ancestral languages trade their usual places. More importantly, there is reason to believe that the average rate of internal language change (p) across typological features does not differ much between languages or geographical regions. We saw in Section 4 that different ecological factors have no appreciable effects on differences among languages, and Wichmann & Holman (no date) have shown that rates of change in individual typological features correlate across different geographical areas, which indicates that p cannot be random.

The choice of typological features and the way that they are encoded may cause related and unrelated languages to be indistinguishable in their behavior with respect to spatial autocorrelations. These two factors, are to a large extent, although not entirely, under the linguist's control. Different typological databases in existence may serve to illustrate the range of choices. The database collected under the auspices of the Pioneers of Island Melanesia project (PIM) of the Max Planck Institute for Psycholinguistics in Nijmegen, i.e., the dataset drawn upon for Dunn et al. (2005), consists entirely of binary features (which are also, by definition, unordered). The database represented by *WALS* is a mixture of everything from binary to 9-ary features, a few of which may be interpreted as ordered, some as unordered, and others as semi-ordered. Finally, the Autotyp database constructed by Balthasar Bickel and Johanna Nichols (http://www.uni-leipzig.de/∼autotyp/) has as many values as the researcher feels necessary for capturing all important differences among functionally related categories. This set of values can be merged to fewer values if such a reduction is opportune for a given purpose. It is also possible to modify the number of feature values in *WALS* and the PIM data. The *WALS* dataset can be recast as binary features, each representing the presence or absence of a given feature value. By the opposite approach, some of the binary PIM features that pertain to related linguistic categories and have mutually exclusive distributions could be recast as many-valued features. Thus, much is up to the researcher. A linguist cannot decide to make a feature which does not exhibit any ordered behavior ordered. But unordered features can be excluded to strengthen the utility of the dataset for the purpose of making historical inferences.

7. General discussion and conclusions

In the beginning of this article we showed that typological similarities among languages strongly depend on the geographical distance among them. This dependency is not surprising, but the correlation between typological differences and geographical distance has nevertheless not previously been studied systematically. Known among biologists as isolation by distance, the dependency has been investigated intensively within genetics for well over half a century. While geneticists began employing computer simulations to study IBD-effects more than a quarter of a century ago, this is the first linguistic study to use such a strategy.

We then went on to investigate how differences in the curves might come about. Our computer simulations indicate that diffusion may cause languages to become more similar. This is hardly surprising, but a less trivial finding was that the effect is stronger among languages which share a common ancestor than among unrelated languages. Migration apparently contributes little to differences in diversity. Language shift has no effect for unrelated languages and only a small effect for related ones. Similarly the effect of language change is only noticeable for related languages. Here, however, the effect is quite drastic. Thus, for related languages diffusion, language shift, and the rate of internal change may all affect the degree of diversity, whereas unrelated languages should show similar curves for spatial autocorrelation in different parts of the world given that nothing seems to affect these distributions.

As a preliminary test of this prediction Figure 9 plots differences among languages that are (thought to be) unrelated in Africa, the New World, and Eurasia (the last corresponding to the combination of the areas Eurasia and SE Asia + Oceania, as defined in *WALS*). Indeed, there do not seem to be any major differences, at least not within the 0–6,000 km range. This is suggested by the crisscrossing nature of the curves. For Eurasia we see fewer differences at around the 1,000 km distance than in the other areas. This is probably an artificial effect of the way that distances are calculated, identifying the location of languages with single points in space. As discussed in Section 2, this approach is expected to inflate the degree of similarity in areas where many large languages are found. Beyond the 6,000 km point languages of the Americas appear to show fewer differences than Eurasian languages at corresponding ranges. If statistically significant, this behavior could be due to a deep genealogical relationship among several of the New World languages. However, as we report

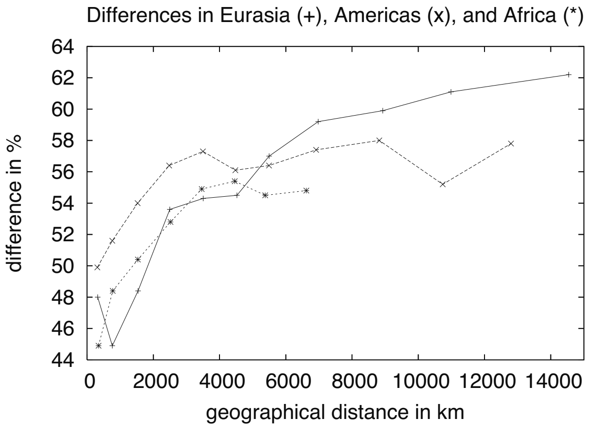

Figure 10. *The relationship between geographical distance in km and typological similarity in Africa, the New World, and Eurasia*

in Wichmann, Holman, et al. (no date) it has not been possible to show the somewhat smaller number of differences among New World languages to be statistically significant. At present the important insight to be derived from Figure 9 is that curves for spatial autocorrelations have a consistent behavior across continents.

We found that it is generally the case that related languages are more similar than unrelated ones at large geographical distances, i.e., that evidence for common descent is preserved. But we also found that this preservation of history is to a great extent dependent on the nature of the data. At large geographical distances, binary or unordered features will tend to obscure the differences between related and unrelated languages, making them look equally dissimilar. This is an important methodological lesson for linguists wishing to make inferences about language history using typological data – an approach which is becoming increasingly more popular in historical linguistic research.

From a historical linguistic perspective the bottom line of the present investigation is that relatedness among groups of languages should be evident from their behavior with respect to spatial autocorrelation given an optimal choice and encoding of data and given the assumption that internal language change is roughly and on average similar across languages. We have not found conditions other than the rate of internal change and the choice and encoding of data that could blur the differences between groups of related and groups of unrelated languages even at large geographical distances. Computer simulations are a useful tool for making one's assumptions explicit, testing hypotheses, and making predictions about real-world behavior. They do not simply substitute for and illustrate a hypothesis, but represent datasets in their own right. Simulated data need to be interpreted just like "real" data. Thus, the final word on the relation between structural diversity and geographical distance among languages is not said in this article. We expect that other researchers will challenge our interpretations and hope that our findings may be tested by means of alternative simulation models and by attempts to verify or falsify the predictions by means of empirical investigations.

Received: 10 July 2006
Revised: 6 July 2007

University of California – Los Angeles
Universität zu Köln
Max-Planck-Institut für evolutionäre Anthropologie

Correspondence addresses: (Wichmann; corresponding author) Department of Linguistics, Max-Planck-Institut für evolutionäre Anthropologie, Deutscher Platz 6, 04103 Leipzig,Germany; e-mail: wichmann@eva.mpg.de; (Holman) Department of Psychology, University of California, Los Angeles, California 90095-1563, USA; e-mail: Holman@psych.ucla.edu; (Schulze) Institut für Theoretische Physik, Universität zu Köln, 50923 Köln, Germany; e-mail: cs@thp.uni-koeln.de; (Stauffer) Institut für Theoretische Physik, Universität zu Köln, 50923 Köln, Germany; e-mail: stauffer@thp.uni-koeln.de

Acknowledgements: Author names are given in alphabetical order. Holman was responsible for sparking off the inquiry and plotting empirical data. Wichmann suggested the research strategy and did most of the write-up. Stauffer provided the computer simulations, using a model proposed by Schulze. We thank Cecil H. Brown for extensive discussions of many of the issues touched upon here, Brigitte Pakendorf for making us aware of the concept of isolation by distance in biology, Hans-Jörg Bibiko for providing us with the *WALS* data in a suitable format, and Bernard Comrie, Charles Nunn, Frans Plank, and three anonymous referees for valuable comments on the manuscript.

Appendix 1: The computational model

In previous studies applying computer simulations to problems of language evolution different models have been developed. The "Viviane model" (Oliveira et al. 2006a, b) simulates the first occupation of a large continent by human beings who initially all speak one language, and the growing diversity of languages during this colonization. However, languages in that model were simply numbered consecutively, preventing a simulation of structural differences observed empirically and shown in Figure 1. Simulations of spatial autocorrelations were only possible in a modified version, presented in Stauffer, Schulze, & Wichmann (2007, in press). The language learning model of Nowak et al. (2002) has a similar disadvantage. The computer models of Abrams & Strogatz (2003) (followed by Patriarca & Leppännen 2004 and Pinasco & Romanelli

2006), of Kosmidis et al. (2005), and of Schwämmle (2005) deal with relatively few languages, not with the thousands reflected in Figure 1. Thus we modify the "Schulze model" (Schulze & Stauffer 2005, 2006) (also used by Teşileanu & Meyer-Ortmanns 2006). Since readers of the present article may not know this model or may not have easy access to the physics literature we describe it in some detail in the following.

A large square lattice is occupied by people speaking one language each. Each language (or grammar) is characterized by F features, each of which can take Q different integer values from 1 to Q. Mostly we use $F = 8$. For the simplest binary case $Q = 2$ efficient bit-string algorithms have been used in the past, allowing larger F, but for the present purposes we vary Q up to 9 and thus use only simpler programs (written in Fortran and available as langpotts26.f from stauffer@thp.uni-koeln.de).

Initially, only the top line of the lattice is occupied, and all others are empty. The people in the top line either all speak the same language or each one independently selects randomly one of the Q^F possible languages. Mainly, what we are interested in is to study the differences between these two different initializations. (For a single ancestor, the one initial language has the integer part of $(1+Q)/2$ for all its features, i.e., the central value for odd Q.)

For each time step (human generation), each occupied lattice site i can change its features and that of its neighbors following four probabilities p, q, r, s for four different processes (i) to (iv):

(i) Shift (r): If a fraction x of people in the whole population speaks the language of site i, then site i shifts with probability $(1-x)^2 r$ to the language of one of its four lattice neighbors, randomly selected. (If this site is still empty, the new language is that used for the initialization, i.e., either the single one or a randomly selected one.) This shift takes into account the tendency for humans to give up speaking minority languages.

(ii) Change (p): Each of the F features is randomly changed with probability p. For zero diffusion probability (see next process) this change is to a randomly selected value for unordered features and to the old value ± 1 for ordered features; see Section 5 for this distinction. (If in the ordered case the new value would be 0 or $Q+1$, the old value is kept for this feature.) This change describes the language changes from one generation to the next.

(iii) Diffusion (q): In the case of diffusion, with probability q the new value is taken during the change of process ii from one of the four lattice neighbors, selected randomly and independently for each of the F features. In this way it is simulated how a language may take over traits from other languages. (If this neighboring site is still empty no diffusion takes place.)

(iv) Migration (s): Each of the four nearest neighbors (North, East, South, and West) is checked, and nothing happens if it is already occupied. If it is empty, then with probability s it becomes occupied, with the same language features as on the original site i. This original site i remains occupied. Migration, then, simulates the peaceful colonization of uninhabited territory by an expansion of the population.

In earlier papers published in physics journals, the first three processes were respectively denoted by the terms "flight", "mutation", and "transfer", while process (iv), which was not introduced in earlier papers, presumably would be called "diffusion" there.

The differences between languages on the top line and those on the lattice line separated by d lattice spacings are calculated in two different ways for the ordered and the unordered features, discussed in Section 6: as the average number of features which are different (unordered case), and as the average sum of the absolute differences in the features (ordered case). For binary features, $Q = 2$, this distinction vanishes. (In both cases we average only over occupied sites.)

Rates of changes for different features are assumed to be equiprobable, whereas in "real life" different features have different rates of change. We tested what would happen when differences in rates of change were introduced into the simulations and saw no qualitative effects. For this reason, and because it is preferable not to introduce too many free parameters, we assume that rates of changes for different features are equal. In Wichmann & Holman (no date) we vary rates of changes in simulations in order to test different metrics for measuring rates of change in *WALS* features. For the present purposes, however, this type of exercise is not relevant. Similarly, we have checked the effect of gaps in the data in order to approximate a "real life" situation and again see no changes in the particular simulations presented here.

The Schulze model and its variants were simulated without migration in several publications (mostly reviewed in Schulze & Stauffer 2006), as a function of the three probabilities p, q, r and the total population N. For $r \simeq 1$ and large but finite N a sharp phase transition was found between heterogeneity and homogeneity. Either the system ends up being heterogeneous, with each possible language spoken by about the same number of people (if the population is not large enough, then a roughly random selection of all possible languages is spoken). Otherwise the system ends up dominated by one language spoken by the majority while the others mostly speak minor variants of this dominating language; this is what we label "homogeneity" here, somewhat imprecisely since the situation is not completely homogeneous (in our other works involving simulations, heterogeneity and homogeneity, as just defined, were termed "fragmentation" and "dominance", respectively). Both final states, heterogeneity and homogeneity, can be reached either from several random ancestors or from a single ancestor. If we start with one person whose offspring lets the population grow up to a final stationary value we necessarily start with homogeneity, and if we also end with homogeneity, we may have a maximum of the number of spoken languages at some intermediate time (Schulze & Stauffer 2005) as in Nettle (1999b). In the $p-q$-plane, one finds a transition line separating heterogeneity (large p, small q) from homogeneity (small p, large q). If instead we select $r \ll 1$, final homogeneity may become impossible. In the present simulation of $L \times L$ square lattices, the population is $N = L^2$ and should be compared with the possible number Q^F of languages. Perhaps for $N \to \infty$ also the time which homogeneity needs to emerge from heterogeneity goes to infinity; thus mathematical limits should be considered with caution.

Appendix 2: A brief description of the metric for typological feature stabilities

The philosophy behind our metric, discussed in more detail in Wichmann & Holman (no date), is that if one given feature more often tends to have the same value for languages that are related than does another given feature, then the first of the two may be considered to be more stable. But when calculating figures representing stability an additional factor has to be taken into account, namely the tendency for traits to be similar among languages that are NOT related. Thus, stabilities are calculated for each feature as follows. Within each group of related languages, we look at all the pairs of languages for which the feature is attested in both languages, and find the proportion of such pairs for which the feature has the same value. This proportion is then averaged across all groups of related languages, with each group weighted by the square root of the number of language pairs for which the feature is attested; the square-root weighting of pairs produces a nearly equal weighting of languages. The weighted average proportion is called R, for similarly behaving related languages. As a baseline, we also look at all the pairs of unrelated languages for which the feature is attested in both languages, and find the proportion of these pairs for which the feature has the same value. This proportion is called U, for similarly behaving unrelated languages. Stability is called S, and is defined as follows:

$$S = \frac{R-U}{1-U},$$

where the numerator is the degree to which similarity is enhanced in related languages compared to the baseline of unrelated languages, and the denominator is the maximum possible enhancement from the baseline. The resulting percentage has a maximum value of 100 % if related languages are identical with respect to the given feature, and it has an expected value of 0 % if unrelated languages are just as similar as related languages. A metric related in spirit to ours is described in Nichols (1995: 347).

Appendix 3: The 33 most and 33 least stable features in *WALS*

33 most stable features (in descending order of stability): 31: Sex-based and Non-sex-based Gender Systems; 118: Predicative Adjectives; 30: Number of Genders; 119: Nominal and Locational Predication; 29: Syncretism in Verbal Person/Number Marking; 85: Order of Adposition and Noun Phrase; 28: Case Syncretism; 32: Systems of Gender Assignment; 83: Order of Object and Verb; 86: Order of Genitive and Noun; 40: Inclusive/Exclusive Distinction in Verbal Inflection; 39: Inclusive/Exclusive Distinction in Independent Pronouns; 10: Vowel Nasalization; 47: Intensifiers and Reflexive Pronouns; 73: The Optative; 121: Comparative Constructions; 18: Absence of Common Consonants; 21: Exponence of Selected Inflectional Formatives; 84: Order of Object, Oblique, and Verb; 57: Position of Pronominal Possessive Affixes; 89: Order of Numeral and Noun; 9: The Velar Nasal; 90: Order of Relative Clause and Noun; 63: Noun Phrase Conjunction; 61: Adjectives without Nouns; 137: N-M Pronouns; 81: Order of Subject, Object, and Verb; 66: The Past Tense; 79: Suppletion According to Tense and Aspect; 42: Pronominal and Adnominal Demonstratives; 99: Alignment of Case Marking of Pronouns; 44: Gender Distinctions in Independent Personal Pronouns; 87: Order of Adjective and Noun.

33 least stable features (in descending order of stability): 68: The Perfect; 4: Voicing in Plosives and Fricatives; 108: Antipassive Constructions; 106: Reciprocal Constructions; 78: Coding of Evidentiality; 64: Nominal and Verbal Conjunction; 136: M-T Pronouns; 105: Ditransitive Constructions: The Verb 'Give'; 102: Verbal Person Marking; 3: Consonant-Vowel Ratio; 52: Comitatives and Instrumentals; 72: Imperative-Hortative Systems; 19: Presence of Uncommon Consonants; 38: Indefinite Articles; 132: Number of Non-Derived Basic Colour Categories; 37: Definite Articles; 15: Weight-Sensitive Stress; 41: Distance Contrasts in Demonstratives; 45: Politeness Distinctions in Pronouns; 110: Periphrastic Causative Constructions; 1: Consonant Inventories; 22: Inflectional Synthesis of the Verb; 11: Front Rounded Vowels; 16: Weight Factors in Weight-Sensitive Stress Systems; 60: Genitives, Adjectives, and Relative Clauses; 34: Occurrence of Nominal Plurality; 133: Number of Basic Colour Categories; 76: Overlap between Situational and Epistemic Modal Marking; 128: Utterance Complement Clauses; 115: Negative Indefinite Pronouns and Predicate Negation; 59: Possessive Classification; 135: Red and Yellow; 58: Obligatory Possessive Inflection.

References

Abrams, Daniel & Steven H. Strogatz (2003). Modelling the dynamics of language death. *Nature* 424: 900.

Anderson, Gregory D. S. (2005). The velar nasal (ŋ). In Haspelmath et al. (eds.) 2005, 42–45.

Baxter, Gareth J., Richard A. Blythe, William Croft, & Alan J. McKane (2006). Utterance selection model of language change. *Physical Review E* 73: article no. 046118.
Bickel, Balthasar (2007). Typology in the 21st century: Major current developments. *Linguistic Typology* 11: 239–251.
Cangelosi, Angelo & Domenico Parisi (eds.) (2002). *Simulating the Evolution of Language*. Berlin: Springer.
Cavalli-Sforza, Luigi Luca & William S.-Y. Wang (1986). Spatial distance of lexical replacement. *Language* 62: 38–55.
Christiansen, Morten & Simon Kirby (eds.) (2003). *Language Evolution*. Oxford: Oxford University Press.
Comrie, Bernard, Matthew S. Dryer, David Gil, & Martin Haspelmath (2005). Introduction. In Haspelmath et al. (eds.) 2005, 1–8.
Dahl, Östen & Viveka Velupillai (2005). Tense and aspect. In Haspelmath et al. (eds.) 2005, 266–281.
de Boer, Bart (2006). Computer modelling as a tool for understanding language evolution. In Nathalie Gontier, Jean Paul van Bendegem, & Diederik Aerts (eds.), *Evolutionary Epistemology, Language and Culture: A Non-Adaptationist Systems Theoretical Approach*, 381–406. Dordrecht: Springer.
Dryer, Matthew S. (2005). Coding of nominal plurality. In Haspelmath et al. (eds.) 2005, 138–141.
Dunn, Michael J., Angela Terrill, Ger P. Reesink, Robert A. Foley, & Stephen C. Levinson (2005). Structural phylogenetics and the reconstruction of ancient language history. *Science* 309: 272–275.
Epperson, Bryan K. (2003). *Geographical Genetics*. Princeton, N.J.: Princeton University Press.
Goebl, Hans (1984). *Dialektometrische Studien anhand italoromanischer und galloromanischer Sprachmaterialien aus AIS und ALF*. 3 volumes. Tübingen: Niemeyer.
— (2001). Arealtypologie und Dialektologie. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher, & Wolfgang Raible (eds.), *Language Typology and Universals: An International Handbook*, Volume 2, 1471–1491. Berlin: Walter de Gruyter.
— (2005). Dialektometrie. In Reinhard Köhler, Gabriel Altmann, & Rajmund G. Piotrowski (eds.), *Quantitative Linguistics: An International Handbook*, 498–531. Berlin: Walter de Gruyter.
Haspelmath, Martin, Matthew Dryer, David Gil, & Bernard Comrie (eds.) (2005). *The World Atlas of Language Structures*. Oxford: Oxford University Press.
Holman, Eric W. (no date). Approximately independent typological features of languages. Manuscript, submitted.
Holman, Eric W., Søren Wichmann, & Cecil H. Brown (no date). Linguistic and cultural diffusion in a comparative perspective. Manuscript, submitted. Available at http://email.eva.mpg.de/~wichmann/diffusion-submitted.pdf
Itoh, Yoshiaki & Sumie Ueda (2004). The Ising model for changes in word ordering rules in natural languages. *Physica D – Nonlinear Phenomena* 198: 333–339.
Kosmidis, Kosmas, John M. Halley, & Panos Argyrakis (2005). Language evolution and population dynamics in a system of two interacting species. *Physica A* 353: 595–612.
Kretzschmar, William A. (1996). Quantitative areal analysis of dialect features. *Language Variation and Change* 8: 13–39.
Maddieson, Ian (2005). Vowel quality inventories. In Haspelmath et al. (eds.) 2005, 14–17.
Maslova, Elena & Tatiana Nikitina (no date). Stochastic universals and dynamics of cross-linguistic distributions: The case of alignment types. Manuscript. Available at http://www.stanford.edu/~emaslova/Publications/Ergativity.pdf, accessed April 11, 2007.
Mira, Jorge & Angel Paredes (2005). Interlinguistic simulation and language death dynamics. *Europhysics Letters* 69-6: 1031–1034.
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 Gert Durieux, Walter

Daelemans, & Steven Gillis (eds.), *Proceedings of CLIN '95*, 185–202. Antwerpen: Universiteit Antwerpen, Centrum voor Nederlandse Taal en Spraak.

Nerbonne, John & Peter Kleiweg (2007). Toward a dialectological yardstick. *Journal of Quantitative Linguistics* 14: 148–167.

Nettle, Daniel (1999a). Using social impact theory to simulate language change. *Lingua* 108: 95–117.

— (1999b). Linguistic diversity of the Americas can be reconciled with a recent colonization. *Proceedings of the National Academy of Sciences of the U.S.A.* 96: 3325–3329.

— (1999c). *Linguistic Diversity*. Oxford: Oxford University Press.

Nichols, Johanna (1992). *Linguistic Diversity in Space and Time*. Chicago: University of Chicago Press.

— (1995). Diachronically stable structural features. In Henning Andersen (ed.), *Historical Linguistics 1993: Selected Papers from the 11th International Conference on Historical Linguistics, Los Angeles 16–20 August 1993*, 337–355. Amsterdam: Benjamins.

Niyogi, Partha (2002). The computational study of diachronic linguistics. In David Lightfoot (ed.), *Syntactic Effects of Morphological Change*, 351–365. Cambridge: Cambridge University Press.

— (2006). *The Computational Nature of Language Learning and Evolution*. Cambridge, Mass.: MIT Press.

Nowak, Martin A., Natalia Komarova, & Partha Niyogi (2002). Computational and evolutionary aspects of language. *Nature* 417: 611–617.

Oliveira, Viviane M. de, Marcelo A. F. Gomes, & Ing Ren Tsang (2006a). Theoretical model for the evolution of linguistic diversity. *Physica A* 361: 361–370.

Oliveira, Viviane M. de, Paulo R. A. Campos, Marcelo A. F. Gomes, & Ing Ren Tsang (2006b). Bounded fitness landscapes and the evolution of the linguistic diversity. *Physica A* 368: 257–261.

Patriarca, Marco & Teemu Leppänen (2004). Modeling language competition. *Physica A* 338: 296–299.

Pinasco, Juan Pablo & Liliana Romanelli (2006). Coexistence of languages is possible. *Physica A* 361: 355–360.

Prévost, Nathalie (2003). The physics of language: Towards a phase transition of language change. Doctoral dissertation, Simon Fraser University.

Schulze, Christian & Dietrich Stauffer (2005). Monte Carlo simulation of the rise and fall of languages. *International Journal of Modern Physics C* 16: 781–787.

— (2006). Computer simulation of language competition by physicists. In Bikas K. Chakrabarti, Anirban Chakraborti, & Arnab Chatterjee (eds.), *Econophysics and Sociophysics: Trends and Perspectives*. Weinheim: Wiley-VCH.

Schwämmle, Veit (2005). Simulation for competition of languages with an ageing sexual population. *International Journal of Modern Physics C* 16-10: 1519–1526.

Séguy, Jean (1971). La relation entre la distance spatiale et la distance lexicale. *Revue de linguistique romane* 35: 335–357.

— (1973). *Atlas linguistique et ethnographique de la Gascogne*, Volume 6: *Notice explicative*. Paris: Centre national de la recherche scientifique.

Silva Jr., Wilson A., Sandro L. Bonatto, Adriano J. Holanda, Andrea K. Ribeiro-dos-Santos, Beatriz M. Paixão, Gustavo H. Goldman, Kiyoko Abe-Sandes, Luis Rodriguez-Delfin, Marcela Barbosa, Maria Luiza Paçó-Larson, Maria Luiza Petzl-Erler, Valeria Valente, Sidney E. B. Santos, & Marco A. Zago (2002). Mitochondrial genome diversity of Native Americans supports a single early entry of founder populations into America. *American Journal of Human Genetics* 71: 187–192.

Sokal, Robert R. & Paolo Menozzi (1982). Spatial autocorrelations of HLA frequencies in Europe support demic diffusion of early farmers. *The American Naturalist* 119: 1–17.

Stauffer, Dietrich, Suzana Moss de Oliveira, Paulo Murilo C. de Oliveira, & Jorge S. Sá Martins (2006). *Biology, Sociology, Geology by Computational Physicists*. Amsterdam: Elsevier.

Stauffer, Dietrich, Christian Schulze, & Søren Wichmann (2007). Computer-Simulation des Wettbewerbs zwischen Sprachen. In Stefan Kolling (ed.), *Beiträge zur Experimentalphysik, Didaktik und computergestützten Physik: Festschrift zum 70. Geburtstag von Prof. Dr. rer. nat. Hans-Josef Patt*, 351–370. Berlin: Logos.

— (in press). Birth, survival and death of languages by Monte Carlo simulation. *Communications in Computational Physics*. Available at http://de.arxiv.org/abs/0704.0691

Stone, Anne C. & Mark Stoneking (1998). mtDNA analysis of a prehistoric Oneota population: Implications for the peopling of the New World. *The American Journal of Human Genetics* 62: 1152–1170.

Sutherland, William J. (2003). Parallel extinction risk and global distribution of languages and species. *Nature* 423: 276–279.

Tarazona-Santos, Eduardo & Fabrício R. Santos (2002). The peopling of the Americas: A second major migration? *The American Journal of Human Genetics* 70: 1377–1380.

Teşileanu, Tiberiu & Hildegard Meyer-Ortmanns (2006). Competition among languages and their Hamming distances. *International Journal of Modern Physcis C* 17: 259-278.

Tobler, Waldo (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography* 46: 234–240.

Wang, William S.-Y., Jinyun Ke, & James W. Minett (2004). Computational studies of language evolution. In Chu-Ren Huang & Winfried Lenders (eds.), *Computational Linguistics and Beyond*, 65–106. Taipei: Institute of Linguistics, Academia Sinica.

Wang, William S.-Y. & James W. Minett (2005). The invasion of language: Emergence, change and death. *Trends in Ecology and Evolution* 20-5: 263–296.

Wichmann, Søren & Eric W. Holman (2007). Pairwise comparisons of typological profiles. Preprint of paper submitted to the proceedings of the conference "Rara & Rarissima: Collecting and interpreting unusual characteristics of human languages", Leipzig, 29 March – 1 April 2006. Available at http://de.arxiv.org/ftp/arxiv/papers/0704/0704.0071.pdf

— (no date). Assessing temporal stability for linguistic typological features. Submitted. Available at http://email.eva.mpg.de/~wichmann/WichmannHolmanIniSubmit.pdf

Wichmann, Søren, Eric W. Holman, Dietrich Stauffer, & Cecil H. Brown (no date). Similarities among languages of the Americas: An exploration of the *W*ALS evidence. Submitted to the proceedings of the conference "Genes and Languages", University of California at Santa Barbara, September 8–10, 2006. Available at http://email.eva.mpg.de/~wichmann/SantaBarbWichmannRevSubmit.pdf

Wichmann, Søren, Dietrich Stauffer, F. Welington S. Lima, & Christian Schulze (2007). Modelling linguistic taxonomic dynamics. *Transactions of the Philological Society* 105: 126–147.

Wright, Sewall (1943). Isolation by distance. *Genetics* 28: 114–138.

Zegura, Stephen L., Tatiana M. Karafet, Lev A. Zhivotovsky, & Michael F. Hammer (2004). High-resolution SNPs and microsatellite haplotypes point to a single, recent entry of Native American Y chromosomes into the Americas. *Molecular Biology and Evolution* 21: 164–175.