Various Variation Aggregates in the LAMSAS South

John Nerbonne

Rijksuniversiteit Groningen

9700 AS Groningen

The Netherlands

nerbonne@let.rug.nl

Telephone: +31 (50) 363 58 15

FAX: +31 (50) 363 68 55

www.let.rug.nl/~nerbonne

submitted to Catherine Davis and Michael Picone (eds.)

Language Variety in the South III.

Tuscaloosa: University of Alabama Press. 2005.

Abstract

Techniques for analyzing variation in the aggregate have improved considerably over the past several years, but we still have no techniques which allow one to extract linguistic structure from aggregate comparison.[1] The present paper explores one means of comparing aggregate analyses and linguistically motivated restrictions, essentially the technique of aggregating over a restricted input set. Using the Southern states data which Guy Lowman collected as part of the LAMSAS, we compare aggregate analyses restricted to vowel differences to those using the complete data set.

1 Introduction

A major contemporary challenge in the analysis of linguistic variation is to relate dialectometric methods which aggregate over the entirety of available data, e.g., an entire linguistic atlas or the collected records of a field worker, on the one hand, to linguistic analyses on the other, which seek characterizations of variation in terms of a small number of parameters, e.g., adjustments in the pronunciation of segments throughout a lexicon (e.g., that /l/ is darker in one area than another ([ɫ] or [lˠ] vs [l], [l̡] or [lʲ])), or even more ambitiously, that entire sets of segments are affected by a change for which an insightful linguistic characterization exists, e.g., the affrication of the German stop series in the south or the vowel shifts that Labov has identified in American pronunciation (Labov 1994, Ch.6).

It is clear that a linguistic characterization—to the degree that it is general and accurate—is to be preferred, since it approaches the scientific ideal of a more general and economical description. It is more general and economical to note once, for an entire lexicon of pronunciations, that /l/ is always pronounced as dark ([ɫ]) than it is to note this for each word in a variety (probably hundreds of times). Researchers would agree that this is the more insightful description as well. But some dialectologists counter that the facts of language variation are often more rebarbative, making very general characterizations at best poor approximations and at worst, simply wrong (Chambers and Trudgill 1998, [1980], § 2.1). A closer look at most linguistic atlases inevitably reveals numerous exceptions to virtually all of the simpler characterizations of dialect differences. See the left side of Fig. 1 for an illustration from the American South of a frequent characterization, the monophthongal pronunciation of the vowel in <u>night</u> [nat] is mapped. The phenomenon is much less general in the data used here than many linguistic characterizations would have it.

1.1 Dialectometry

Dialectometry has arisen as a means of seeking general characterizations through aggregation of all available data (Goebl 1982, Goebl 1984, Nerbonne and Kretzschmar 2003). This work proceeds from the assumption that varieties, i.e., sums of speech dispositions in a given community ought to be characterized in their relations to one another, not merely arbitrarily selected features. Dialectometry has shown that exceptions need not disturb characterizations, e.g., of dialect areas, if they are evaluated together with the rest of the data (and analyzed statistically); it has vindicated the postulation of dialect areas (exceptions and non-coinciding isoglosses notwithstanding); and it has provided the first rigorous foundation for the intuition that dialects may be organized along a continuum (Nerbonne, Heeringa and Kleiweg 1999, Heeringa and Nerbonne 2002).

Dialectometry achieves these results at the price of abstracting to the level of a measurement of difference between varieties. In measuring differences, the dialectometrist deliberately abstracts away from the details of what has contributed the difference, in an abstraction step that is inherent to the strength of the approach, but which at the same time loses the connection to the linguistic characterization.[2] Lest we appear to be painting with too broad a brush here, let us hasten to add that it is possible to apply dialectometric techniques to material which has been linguistically prepared, e.g., to a matrix of properties, some or all of which might be linguistically abstract. The techniques are general in this sense. But even in this case the result will be a measure of difference which no longer bears any trace of its origin.

This paper suggests one way of linking dialectometric characterizations to more detailed linguistic characterizations, and this is simply to aggregate over a linguistically interesting subset of the data.

2 LAMSAS

The *Linguistic Atlas of the Middle and South Atlantic States* (LAMSAS) comprises dialect material collected on the Eastern seaboard of the United States from 1933 through 1974. Our focus here will be on the pronunciation of vowels in part of the data from the South, namely the part collected by Guy Lowman in 1933-1936. Lowman and Raven McDavid were together responsible for 95% of the data collected in LAMSAS, but Nerbonne and Kleiweg (2003) document the degree to Lowman and McDavid differed in the data they collected, and suggests that it is sensible to analyze them separately.

The LAMSAS material is admirably accessible for reanalysis (see `http://hyde.park.uga.edu/lamsas/`, (Kretzschmar 1994)) and contains the responses of 1162 informants who were interviewed in 483 communities. The responses to 151 different items is included in the web distribution, which formed the basis for the work here.

We focus here on Lowman's data from North Carolina, Virginia, West Virginia, and the District of Columbia. We likewise include data from Maryland and Delaware in order to provide context for our comparisons. The map in Fig. 1

indicates the range of sites included in this study. This subset of the data included 238 field work sites, and 57,833 phonetic transcriptions of words and brief phrases or roughly 243 per site. Since we shall focus on vowels below, let us note that there is a total of 1,132 different vowels (different combinations of basic segment plus one or more diacritics) in this data.

2.1 Southern Vowels

We chose to analyze vowels in this study in order to illustrate the aggregation at a level below that of the entire data set. Vowels are often remarked to be the more fluid bearers of varietal differences in general, and distinct markers of American Southern speech in particular (Labov 1994, 201ff.), which makes them an interesting candidates for analysis.

In order to contrast aggregate dialectometric techniques with those focusing on individual features, it is useful to inspect the distribution of some example features. Fig. 1 presents the frequencies of two well-studied features of Southern speech, the monophthongal pronunciation of vowels in night etc. [nat] (instead of [naɪt]), and the diphthongal pronunciation of a range of vowels that are pronounced as monophthongal in the speech of (most) Northerners in words such as afternoon [æᵊ f t ən ʉn] vs. [æf t ər n u n].

The LAMSAS words and phrases used to collect frequencies for [aɪ/a] were Ohio, dragonfly, dry spell, five, he died with, lightwood, miles, my wife, night and nine. To check on diphthongal vs. monophthongal pronunciation (e.g., [æ^{ə} vs. [æ] in afternoon, we examined the first vowels in Alabama, Asheville, Baltimore, France, Wednesday, afternoon, cleans up, fog, half-past seven, hundred and ten. As the reader may satisfy herself by examining Fig. 3, both of these features are far cry from being definitive in characterizing the Southern U.S. speech area.

The idea behind the aggregation in dialectometry is that the sum of speech differences in a variety should provide the most reliable basis for characterizing its relations to other varieties. If one imagined collecting maps like those in Fig. 1, and superimposing them on one another, the sum would be an aggregate dialectometric map.

3 Measuring Pronunciation Differences

Various phoneticians have proposed methods to measure the difference between pairs of phonetic segments (Vieregge, Rietveld and Jansen 1984, Almeida and Braun 1986). Our work has been to embed these systems (and others, see Heeringa (2004) for a current survey) into a larger ones in which the distance between sequences is assayed. Fig. 2 illustrates the main idea behind the procedure.

The task of choosing an appropriate feature set is not trivial. Most segments include one or more diacritics (on average, each phonetic base segment is accompanied by 0.56 diacritics, and some segments bear three and even four

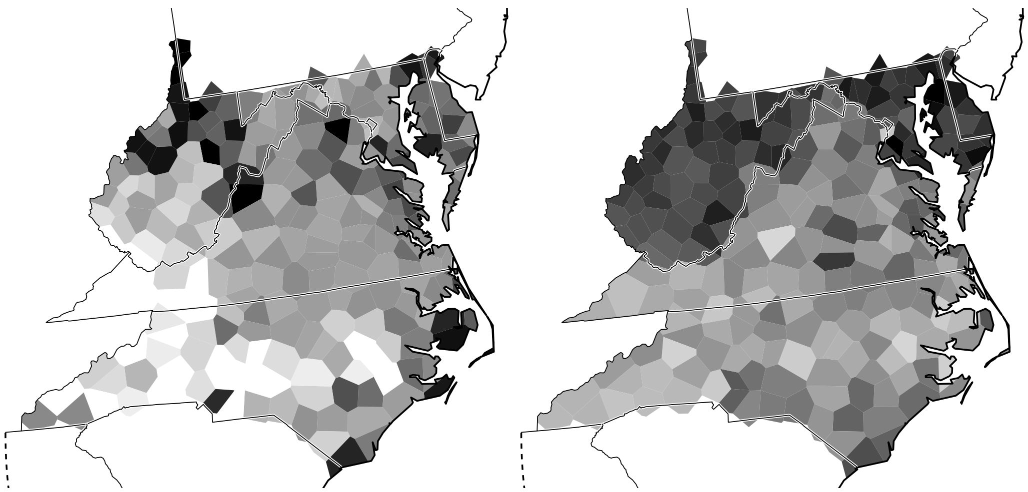

Figure 1: The darker polygons in the map on the left above show how frequently the vowel in night (and similar words) is pronounced [aɪ], and the light ones where it is pronounced [a]. The dark areas in the map on the right show monophthongal pronunciations of the first vowel in afternoon (and similar words) [æ] as opposed to diphthongal ones [$æ^{ə}$]. The idea behind dialectometric aggregation is to sum over all such differences, and this indeed gives a reliable indication of dialect differences. We note in passing that even the [a/aɪ] shibboleth does not have the clean distribution in American speech which dialect maps sometimes suggest, at least not in the LAMSAS data from the 1930's.

diacritics). As we noted above, a total of 1,132 different vowels are transcribed in the LAMSAS database, which, incidentally represents an editorial simplification with respect to the field workers notes (Kretzschmar 1994, ch.5). We should be cautious in attributing a high degree of reliability to the details of the transcriptions since the interviews were conducted informally, by a single interviewer, who was also responsible for the transcription, and who could not fall back on recordings to verify his notes. Most LAMSAS interviews were conducted without the benefit of mechanical recording apparatus, including all the interviews that we use here. We shall not focus on these data preparation issues here, however.

We used a feature set derived from Kretzschmar (1994, p.116) which we summarize in the table below. The table notes not only the feature but also the complete range of values that we used to interpret the feature. Even though the features are those suggested by Kretzschmar (1994), and the number of values is determined by the number of different distinctions we found in the database (see also Kretzschmar et al. p.118 on the number of distinctions), still we are responsible for the relative weights assigned to the different features. Heeringa (2004, Ch.7) finds that the segment measurements are robust with respect to small changes in relative weighting of features, and this is fortunate since it is difficult, if not impossible to set the relative weights in a non-arbitrary way. We comment further on this below.

v-advanced	-3, -2, -1, 0, 0.4, 1, 1.4, 2, 2.4, 3
v-high	-1.75, -1.5, -1.25, -1, -0.75, -0.5, -0.25, 0, 0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75
v-rounded	-1, -0.5, 0, 0.5, 1
v-long	-0.5, 0, 0.5, 1
v-stress	0, 0.35, 0.7
v-nasal	0, 1
v-rhotic	0, 1
v-super	0, 1
v-pharyng.	0, 1
v-voice	0, 1

Perhaps it is useful to note first what is not represented. Diphthongs were represented by two segments and differences between them will effectively be analyzed as the sum of differences between the first and second parts, respectively. This means that diphthongs are not represented via particular feature configurations.

The feature names reflect their normal phonetic interpretation. The stress which is marked on a syllable is interpreted as a property of the vowel, which is why it appears on the list above. Vowels receive either stress, secondary stress, or no stress. Vowels were interpreted as voiced except when explicitly marked as voiceless, in which case they bore the feature [- voice]. Lowman rarely added a diacritic indicating the "pharyngealization" of a vowel, and the [v-pharyng.] feature interprets that. Vowels written as superscripts (e.g., the second parts of laxing diphthongs) are not interpreted through a feature [± super]—but rather through a weighting. Comparisons involving superscripted vowels count only 50% of what they would cost if they segments compared were both non-superscripted. The idea behind this naturally is that such minor articulations should contribute less to pronunciation difference.

The range of values reflects the number of distinctions made in the data, where we have occasionally taken the liberty of simplifying. We found 15 height distinctions in vowels, all of which may be represented in the values here. But we could simplify the six degrees of rounding distinguished in the LAMSAS data to only five, as we did not find more than five in the data analyzed here.

The distance between two segments was taken to be a logarithm of the sum of the differences in the feature values, more specifically $\log(1 + \text{sum})$. We employ a logarithm to de-emphasize large differences, following Heeringa and Braun (2003, 264–265), and in accordance with the idea that we are dealing with a psychophysical regularity (Stevens 1975). Since the values of some features may differ more than those of others, this effectively weights some features as more important than others. Advancement may differ by as much as 6, while rounding can not differ by more than one. Diacritics representing stress, rhotism, pharyngealization and devoicing were each capable of adding

	Operation	Cost
æəf t ən ʉn		
æf t ən ʉn	delete ə	d(ə,[])=0.3
æf t ər n ʉn	insert r	d([],r)=0.2
æf t ər n u n	replace [ʉ] with u	d([ʉ],[u])=0.1
Total		0.6

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 phoneticians' work on the distance between individual phonetic sounds. The pronunciations are from Savannah, Georgia (top) and Lancaster, Pennsylvania (bottom) (both in LAMSAS). We do not illustrate the algorithm which guarantees that the least costly set of operations is used to determine the overall cost. See Heeringa (2004) for detailed explanations and algorithms.

maximally one unit of difference, and intermediate differences, including those indicated by diacritics, were interpolated. The differential weightings of the features are given implicitly by the difference in the extreme values which the feature can take on.

In order to "lift" the segment distances to the level of sequences, we used the operations (i) the insertion of a single sound, (ii) the deletion of a single sound, and (iii) the substitution of one sound for another. Other operations are possible. The operation costs used in the procedures were those assigned by the feature differences explained above (see Heeringa (2004) for details on a range of alternatives). They consist of the 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). A "silence" was defined using the features of the voiceless subscripted schwa [ə]. All of the measurements here were carried out using the freely available RuG/L04 package (`www.let.rug.nl/~kleiweg/L04`).

4 Results

We apply a Levenshtein procedure to all of the pairs of phonetic transcriptions from each pair of the 238 sites mentioned above. The procedure is adjusted to allow for multiple pronunciations at a single site (multiple responses) (Nerbonne and Kleiweg 2003, p.349). In this section we compare the analysis of vowels to an analysis of the entire phonetic transcription in in order to determine how important the vowels are. When we analyse the entire phonetic transcription, consonants and vowles, we use the simplest segment differentiation—segments are either identical or different. We refer to this as a "phone-based" analysis. Heeringa (2004, p.174,p.186) demonstrates that phone-based methods work nearly as well as the more refined feature-based methods.

In the experiment restricted to vowels the Levenshtein procedure serves to align the strings so that we have the best chance of comparing the proper vowels. We compare the vowel results to those obtained from the entire set of transcriptions using the crude segment difference measure.

The result of the analysis is a pronunciation distance chart, comparable to the distance charts of automobile clubs. Just as in those, there are distances in each

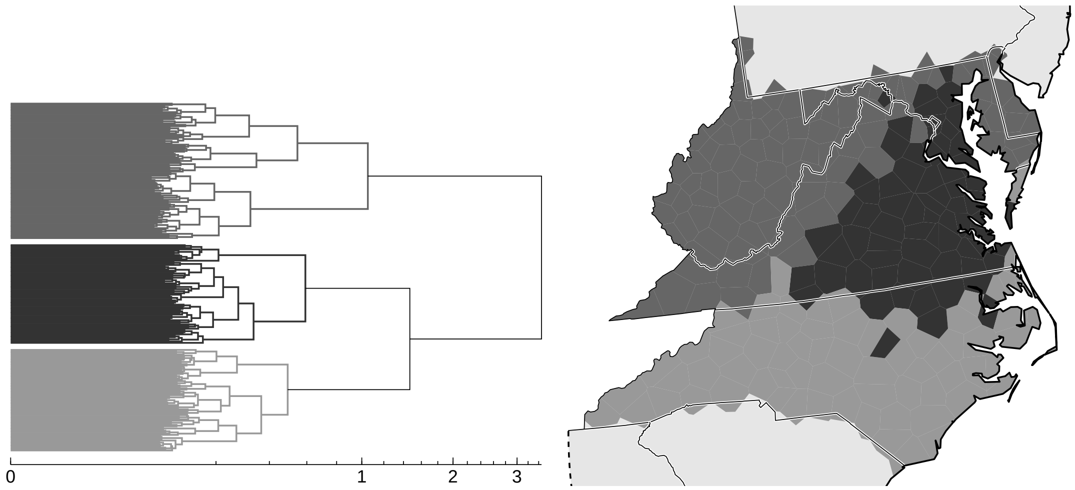

Figure 3: The result of (Ward's) clustering applied to the vowel analyses. The dendrogram on the left shows that the major break is indeed the North-South border, and the map on the right shows the areas identified by clustering. Vowels indeed distinguish the LAMSAS South and its subareas. The cophenetic correlation between the original distances and the distances represented in the dendrogram is $r = 0.62$ (Sokal and Rohlf 1962).

(place × place) cell, only here they indicate pronunciation distance rather than traveling distance. Such distance matrices may be analyzed using multi-dimensional scaling (Nerbonne et al. 1999), or via various forms of clustering. We shall focus on clustering here.

Clustering is an exploratory technique that seeks groups in data. It is most easily understood procedurally. We begin with a list of sites that we gradually connect via a tree. We work from a half-matrix of distances calculated by the procedures above. It is a half-matrix since we can ignore one half due to the symmetry of distance (the distance from a to b is always the same as the distance from b to a). We select the shortest distance, then fuse the two points, a and b, which are involved. This corresponds to adding a node to the list of sites and drawing branches to the two elements now fused by virtue of the node. The distance from the new, fused ab point to each of the others in the half-matrix must then be assigned, and there are several ways of doing this. Ward's method minimizes the squared differences from each of the original points to the newly fused one. It simultaneously minimizes $(d(a,x) - d(ab,x))^2$ and $(d(b,x) - d(ab,x))^2$ for all other x in the distance matrix.

Fig. 3 shows the result of clustering the vowel distances, including a projection to the map of the areas. We note that the procedure mostly finds geographically coherent speech areas, even though no geographic information is input to it, confirming that the approach makes dialectological sense; that the South emerges as clearly distinct in this process; and that Eastern Virginia (Piedmont) emerges within the South. We attribute the fact that the Eastern Virginia cluster is not entire coherent to instability of clustering itself. Shortly, we examine an alternative view which reflects the structure of the pronunciation differences more faithfully (Fig. 4). These results in themselves indicate that the dialectometric techniques are performing well when applied to the restricted data set.

In order to test our hypothesis that vowels are responsible for a great deal of the dialectal differences in the south we compare the analysis developed thus far with an analysis of the entire pronunciation, vowels and consonants. In order to keep the latter analysis simple, we do not include differential weights for phonetic segments. A segment is either identical to another to which is compared (distance = 0) or it is not (distance = 1). This is coarse, but the quantity of information compensates for the coarseness of the comparison.

A "composite cluster map" offers an alternative view of the distance matrix. This sort of map was developed by Kleiweg, Nerbonne and Bosveld (2004) in order to visualize the information in a (pronunciation) distance table as it projects to a map. The map reflects more of the information in the distance table than maps such as the one in Fig. 3, and it compensates to some degree for the instability in the cluster procedure. We obtain these composite cluster maps by repeatedly clustering the distance table, adding random amounts of noise to the distances. The structure which emerges under the addition of noise should be stable. We draw each border around all of the subgroups in each dendrogram obtained through the repeated clustering. The darkness of borders reflect the frequency with which it is drawn. The map reflects rather reliably and sensitively the structure in a distance table. Software for drawing these maps is available at `www.let.rug.nl/~kleiweg/L04`.

Fig. 4 compares the vowel analysis to the analysis based on the entirety of pronunciation material available. The fact that the analyses result in such comparable maps indicates that the vowels are probably responsible for a good deal of the aggregate differences. We cannot exclude at this point the possibility

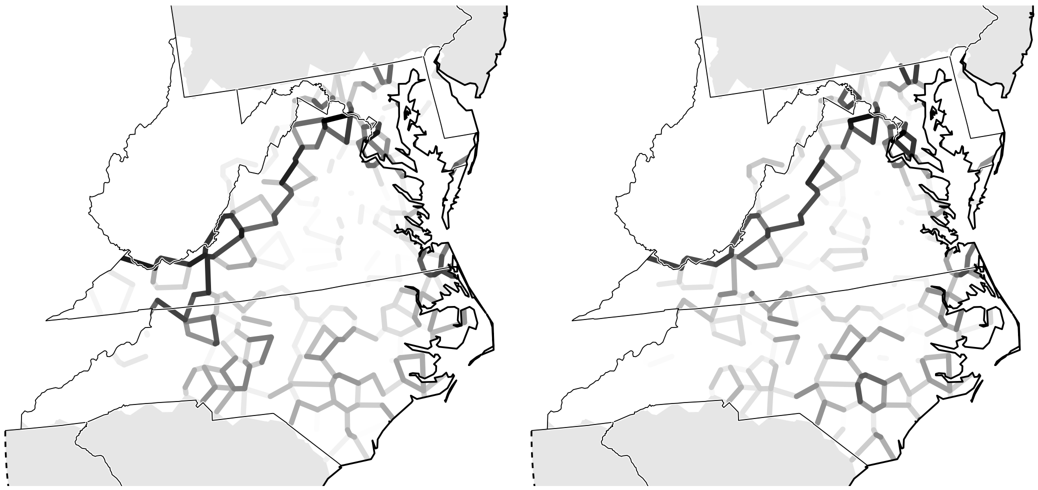

Figure 4: Two "composite cluster maps" of the sort developed by Kleiweg, Nerbonne and Bosveld (2004). We obtain the map on the left by analyzing all pronunciations while using the crudest notion of segment distance, viz., identity, and including all segments, in particular, consonants as well. On the right is the analysis restricted to vowels, but using the finer measure of segment distance (see text). Vowels indeed distinguish the LAMSAS South and its subareas as one can discern a border separating Virginia from West Virginia and Maryland, but the south itself remains quite complex. The composite cluster map is obtained by drawing every border implicit at any level of clustering, where clustering is repeated with random amounts of noise. The darkness of borders reflects the frequency with which it is drawn. The distances obtained using only vowels correlate closely with the distances obtained from entire transcriptions ($r = 0.936$).

that the signal is in fact redundant, so that any number of restrictions might be sufficient to determine the overall relations. To do that, we would need to reanalyze, using other subsets of material.

We can likewise note that the distances between sites assigned by the algorithm using vowels and feature-based segment differences correlates closely with the distances assigned by the algorithm using entire phonetic transcriptions but only the crude identity/nonidentity between segments ($r = 0.936$). We may therefore conclude that the vowels account for 87.6%$(= r^2)$ of the variance in pronunciation and that attention to vowel pronunciation is an excellent indication of dialectal identity in the Southern United States. Since signals may be redundant, we may not conclude that no other linguistic features will be as successful and we explicitly warn against the conclusion that other features could not account for more than 12.4% $(= 100 - 87.6\%)$ of the variance of pronunciation.

5 Conclusions and Prospects

We argued in the introduction that dialectometry, which has been successful in delineating global trends among dialects, needs to be enhanced in order to interpret detailed linguistic claims. Linguists' claims about dialect delineations may be overeager or even inexact about what characterizes a dialect area, but they are unquestionably superior in the degree to which they attempt generalization over the data, a property we take to make them scientifically interesting.

This paper has presented a first, crude means of teasing out the linguistic structure in large-scale dialectometric comparison. We have compared the results of dialectometric analysis performed on the entire data set of Lowman's southern pronunciations in LAMSAS to the the data set restricted to consist only of vowels and extracted from the first. The results have been encouraging—aggregate comparison applied to the Southern vowels appears to characterize the same dialect areas as analyses which aggregate over complete transcriptions. In fact the characterization agrees to a remarkable degree ($r = 0.936$).

It is clear that more needs to be done, and several alternative lines of investigation suggest themselves. We are accustomed to obtaining an aggregate distance matrix (place × place) which we analyze in different ways to understand the dialectological landscape. If alternatively we extract a (place × place) matrix, not only for the sums of phonic differences, but instead for each word in the data set, then we are in a position to calculate a correlation matrix for the words themselves based on the degree to which the place × place matrices (per word) correlate. Once we obtain the correlation matrix, we are in a position to apply numerical analysis (factor analysis) as a means of attempting to isolate the most important generalizations structuring the data set. If this is still too complex, perhaps because words themselves are too complex, we may need to attempt it in combination with the technique explored in this paper, i.e., data set restriction.

Agrawal, Imielinski and Swami (1993), together with other members of the so-called "data-mining" community, have proposed that one explore essentially all of the potential correlations between database elements. Until Agrawal et al.'s work, there was concern that the number of combinations would make such an indiscriminant procedure infeasible, but Agrawal et al. have shown that this need not be the case. Again, this sort of technique might need to be combined with some intelligent restriction on the data set.

Finally, Kondrak (2002) and Gray and Atkinson (2003) attack the historical question directly, seeking automatic means of dividing languages up using the principles of historical reconstruction in linguistics. An application of their techniques to data sets of dialectal data would seem to be straightforward, but dialectologists in the field are quite concerned to record a level of detail which neither of these works is likely to have encountered thus far.

Notes

[1]I appreciated comments on my talk by Bridget Anderson, Bill Kretzschmar, Bill Labov, Dennis Preston and Bob Schackleton. Peter Kleiweg provided the programming needed to carry out the analyses, and the programs are available at `www.let.rug.nl/~kleiweg/L04`. Wilbert Heeringa criticized a first draft of this paper very thoroughly. This work was supported by NWO grant "Determinants of Dialect Variation" 360-70-121, P.I. J.Nerbonne.

[2]Hoppenbrouwers and Hoppenbrouwers (2001), who investigate the distribution of features, is a notable exception. See also Heeringa (2004, Ch.3,7) for an evaluation of Hoppenbrouwers's work.

References

Agrawal, R., T. Imielinski and A. Swami. 1993. Mining association rules between sets of items in large databases. In Proc. ACM SIGMOD

International Conference on Management of Data eds. P.Buneman and S.Jajodia, 207–216, New York:ACM.

Almeida, Almerindo and Angelika Braun. 1986. 'Richtig' und 'Falsch' in phonetischer Transkription: Vorschläge zum Vergleich von Transkriptionen mit Beispielen aus deutschen Dialekten. Zeitschrift für Dialektologie und Linguistik LIII(2):158–172.

Chambers, J.K. and Peter Trudgill. 1998 [1980]. Dialectology. Cambridge: Cambridge University Press. Pages quoted from the 1st ed.

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.

Gray, Russell D. and Quentin D. Atkinson. 2003. Language-tree divergence times support the Anatolian theory of Indo-European origin. Nature 426:435–439. (27 Nov.)

Heeringa, W. J. 2004. Measuring Dialect Pronunciation Differences using Levenshtein Distance. Ph.D. diss., Rijksuniversiteit Groningen.

Heeringa, Wilbert and Angelika Braun 2003. The use of the Almeida-Braun system in the measurement of Dutch dialect distances. Computers and the Humanities 37:257–271. Special Issue on Computational Techniques in Dialectometry, eds. John Nerbonne and William Kretzschmar.

Heeringa, Wilbert and John Nerbonne 2002. Dialect areas and dialect continua. Language Variation and Change 13:375–400.

Hoppenbrouwers, Cor and Geer Hoppenbrouwers 2001. De indeling van de Nederlandse streektalen: Dialecten van 156 steden en dorpen geklasseerd volgens de FFM (feature frequentie methode), Assen: Koninklijke Van Gorcum.

Kleiweg, Peter, John Nerbonne and Leonie Bosveld 2004. Geographic projection of cluster composites. In Diagrammatic Representation and Inference. Third International Conference, Diagrams 2004. Cambridge, UK. eds. A.Blackwell, K.Marriott and A.Shimojima, 392–394. (Lecture Notes in Artificial Intelligence, 2980, Berlin: Springer).

Kondrak, Grzegorz 2002. Algorithms for Language Reconstruction. Ph.D. diss., University of Toronto.

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. 1994. Principles of linguistic change. Vol. 1: Internal factors, Oxford: Blackwell.

Nerbonne, John and Peter Kleiweg. 2003. Lexical variation in LAMSAS. Computers and the Humanities 37: 339–357. Special Iss. on Computational Methods in Dialectometry ed. by John Nerbonne and William Kretzschmar, Jr.

Nerbonne, John, Wilbert Heeringa and Peter Kleiweg. 1999. Edit distance and dialect proximity. In Time Warps, String Edits and Macromolecules: The Theory and Practice of Sequence Comparison, eds. David Sankoff and Joseph Kruskal 2nd ed., v–xv. Stanford: CSLI.

Nerbonne, John and William Kretzschmar. 2003. Introducing computational methods in dialectometry. Computers and the Humanities 37:245–255. Special Iss. on Computational Methods in Dialectometry ed. by John Nerbonne and William Kretzschmar, Jr.

Sokal, Robert R. and Frank James Rohlf. 1962. The comparison of dendrograms by objective methods. Taxon 11:33–40.

Stevens, S. Smith. 1975. Psychophysics: Introduction to its Perceptual, Neural and Social Prospects. New York: John Wiley.

Vieregge, Wilhelm H., A.C.M. Rietveld and Carel Jansen. 1984. A distinctive feature based system for the evaluation of segmental transcription in Dutch. In Proc. of the 10th International Congress of Phonetic Sciences, eds. M. P.van den Broecke and A. Cohen, 654–659. Dordrecht.