The Future of Dialectology

William A. Kretzschmar, Jr.

Abstract

We rely on our teachers and admire their achievements in the first generation of dialectology, and now must decide how to proceed into the next generation. Language variation is not going away, and in order to deal with it most effectively we should consider three things: 1) our theoretical position; 2) our collection, encoding, and analysis of data; and 3) our presentation of data and results. In theory we should recognize that we are doing something different from our generativist and structuralist colleagues, and should ally ourselves with other linguists doing empirical research. We should make best use of emerging technologies in our data handling, and we should improve on first-generation procedures through use of discourse interviews, execution of valid survey research methods, and collection of speech perception as well as speech production data. We should create SGML-tagged transcripts of interviews. We should attempt to analyze our data as we go, with emphasis on quantitative methods but without reliance on any one best analytical method. Finally, we should make every effort to share our data and results with the public, which is badly in need of better information about their language.

Contemporary work in dialectology should always reflect the achievements of our teachers in the first generation of work on Linguistic Atlases. My usual work concerns the American Linguistic Atlas Project, most of whose interviews were carried out in the 1930s and 1940s, though some field work continued long after and some is even going on today. My students execute new projects on contemporary speech, but even then the past informs their best work and results (e.g., Johnson 1996). Thus it has been altogether appropriate for us to celebrate Harold Orton at this conference. The monumental labor on English dialectology of Orton and his

colleagues in England, and of Hans Kurath, Raven McDavid, and their colleagues in America, underlies what we do now. We begin with their example, and however we make our way forward we do so in respect of their monumental work.

My topic today, however, is the future of dialectology, and that kind of prediction is an unaccustomed task for me; I have therefore turned for help to our most impressive and popular contemporary vision of the future, Star Trek. I could hardly come before this audience without new field work, and the Internet provided a convenient portal to the 23rd and 24th centuries, at least as far as the Starfleet Library. I am sorry to report that the future of dialectology in the 24th century is bleak indeed. While language contact situations still do occur despite the use of a machine-based universal translator, as shown in the accented Galactic Standard English of Ensign Chekov, dialects seem almost to have disappeared. As late as the 23rd century (in the original Star Trek series), remnants of an American Southern accent still remained in the speech of Dr. McCoy, but these seem to have disappeared in the Next Generation. Scots and Irish accents seem to have become an occupational dialect for engineers like Mr. Scott, or Miles O'Brien in the Next Generation, apparently optional but preferred for job advancement. And astoundingly, what should have been dialectal French from the Next Generation's Captain Picard somehow turned into the Captain's Conservative RP – which indicates either the long-term success of the language teaching unit of Oxford University Press or some *truly* bizarre sociopolitical development in France. For dialectologists, the world of Star Trek's 24th century will not have much to offer, at least professionally; in Peter Trudgill's words, 'A world where everyone spoke the same language [or dialect] could be a very dull and stagnant place' (Trudgill 1975: 16), even among the stars.

The Star Trek Syndrome, as I name this dialect-empty prediction for my students, has its adherents in our own era. Joseph Wright felt a sense of urgency about recording English dialect features because 'pure dialect speech is rapidly disappearing from our midst' (1898: v), a sentiment shared by American dialectologists at the time and still by many people today. For instance, the 1995 annual meeting of the American Dialect Society featured a special session on 'Endangered Dialects,' and many of the papers at that session commented on 'receding' dialects, or dialects threatened with eradication. Perhaps a great many people wish that language variation *would* go away, as shown in the recent American controversy over Ebonics. However, change or even loss of dialects need not lead to the Star Trek Syndrome. William Labov has written recently about

> . . . the increasing diversity of American English [which is]

> the main finding of our research [at the Penn Linguistics Laboratory], one that violates the most commonsense expectation of how language works and is supposed to work. In spite of the intense exposure of the American population to a national media with a convergent network standard of pronunciation, sound change continues actively in all urban dialects that have been studied, so that the local accents of Boston, New York, Philadelphia, Atlanta, Buffalo, Detroit, Chicago, and San Francisco are more different from each other than at any time in the past. . . . Though the first findings dealt with sound change in Eastern cities, it is now clear that it is equally true of Northern, Western, and Southern dialects.
>
> (1997: 508)

This is good news for dialectologists, even if it does not feel like commonsense for some people (cf. Kretzschmar 1997). Language variation remains strong; however much Standard English may be preferred in the media and the schools, it has not driven out varieties that belong to local communities. We see the truth of this on a larger scale in the innumerable English voices that we come to hear as English continues its progress as *the* world language. The fact is that particular dialects change and come and go, just as particular languages change and come and go, so that what is constant is the presence of dialects and languages, not the fixed existence of any particular dialect or language. Indeed, there is evidence that salient features of many a dialect that we have taken for granted, like American Southern, are actually of recent vintage (Bailey and Ross 1992, Bailey 1997). From the situation as we observe it in the 20th century, so long as we get outside of the classrooms and the suburbs and the new towns and the newspaper columns, there seems to be little danger of progress towards the Star Trek Syndrome – and no threat to our job security.

The question of the future of dialectology, then, is not whether there will *be* any language variation, but rather what we should do about it. Whatever else we choose to do, there will be no replacement for having highly-trained field workers talk to speakers as part of a planned survey. This is a distinguishing characteristic of dialectology, as our field may be separated from the essentially structuralist approach of Labovian sociolinguistics or from the formal logic of generative linguistics (more about these soon). The original impulse of Wenker and Gilliéron in the first wave of Atlas studies was sound: there is great value in finding out what real people actually

say. And that takes real labor, as many here know all too well. Dialectology is not the kind of thing that one can do solely from a comfortable study, and stories from the field are legion. My students still laugh, for instance, when I tell them about the American field worker, Guy Lowman, who was chased from a Virginia farmhouse at the point of a pitchfork for daring to ask the woman of the house about names for male farm animals. We can look forward to yet more stories in the future. However, to go along with the basic necessity of field work, we have new technology to consider, and we form a party in the continually changing flow of ideas about dialects, and more generally about linguistics, and we need to take account of technology and theory as we imagine our future. We need to consider where we stand with regard to three major points: 1) our theoretical stance; 2) our collection, encoding, and analysis of data; and 3) our presentation of our results. I would like to address each of these points in turn.

The first thing to say about theory for dialectologists in the future is that we should claim one. In my own training with Raven McDavid, explicit talk about linguistic theory was minimal. For instance, when I once seriously raised the issue of why an isogloss should have been drawn where Raven drew it, all he replied was that that was where it ought to go. When American dialectologists have turned their minds to theory in years past, as for instance in Davis's work on diafeatures (1973), the prevailing climate did not really allow such thinking to get very far. Everybody much preferred to talk about 'methods' – and so the word 'methods' is part of the title for the triennial conference in dialectology, the International Conference on Methods in Dialectology. This is not to say, of course, that Orton, Kurath, and McDavid did not actually have any theory, only that the theory was most often left unstated. Orton and Dieth were explicitly interested in modern dialects as evidence for study of historical varieties of English, as well as in modern regional differences (Orton, Sanderson, and Widdowson 1978: Introduction). Raven McDavid confirmed this affiliation with historical linguistics when he wrote that 'Dialect geography is a venerable subfield within the new science of linguistics, and is basically a branch of historical linguistics' (McDavid et al. 1986:390). That said, the practical problems of doing the work were the most important consideration for our teachers.

Once could say that this concentration on methods and historical linguistics insulated dialectology from the mid-century ferment in structural linguistics and the somewhat later battles over the transformational paradigm. Historical linguistics has remained in large part a traditional discipline, the only area of linguistics where lists of works cited still typically include many titles from the nineteenth century. One might also say that dialectology was not insulated but isolated by that view, rendered

marginal, because it did not or could not marshal all of its tremendous weight of evidence to participate in the linguistics wars of the 1950s and thereafter. Whether or not this was the cause, we do in fact find ourselves to be isolated and marginalized, perhaps more in America than here in England or in Europe. My only regret among the successes of the previous generation of dialectologists is this apparent retreat from the main theater in linguistics, to my way of thinking without effectively having defended dialectology from attacks of irrelevance (as by Chomsky) or of faulty and inappropriate procedures (as by some sociolinguists, notably C. J. Bailey).

The sort of defense in theory that might have been mounted has been outlined by Lee Pederson, who has written that dialectology is

> a logically ordered and systematic approach that begins with common sense, proceeds through deductive cycles, and concludes in enumeration. It conducts research in a geographic context, but its research concerns a few words of a language, not the language itself and its universe of discourse. . . . American dialectologists, for example, concentrate on sorting and counting components – American English synonyms, morphs, and phones. They are not concerned with the identification of new linguistic classes, semantic, grammatical, and phonological sets established according to the scientific method. . . . In word geography, [deduction] concerns the engagement of target forms. It takes them first as contrastive lexical sets, and then carries the work forward through segmentation of self-evident morphemes, phonemes, allophones, and distinctive features, according to the needs of a descriptive problem. Taken this way, deductive word geography studies only classes and components of phonological words as they characterize speakers classified according to geographical place and analyzed according to social factors.
>
> (1995: 35-36)

What a concise description this is, and how well it captures the essentially French, Gilliéronian tradition of dialectology that always has informed the American Atlas underneath the German layer of isoglosses and dialect areas (cf. Kretzschmar 1995). I completely agree with Pederson that 'common sense itself suggests the importance of these considerations . . . because American word geographers have so far given them little attention, "disdaining as they do to turn their minds to such simple things . . ."'(35). Simple in Pederson's formulation, but crucial to express and to

defend as a way of studying language.

In future, I think that we should affiliate ourselves with the emerging area of empirical linguistics. Empirical linguistics, I would argue, is an alternative to the structural and generative paradigms, and it develops from the example of Gilliéron. To illustrate the difference, let us consider the basic question of how to make a generalization about a language from each point of view. A structuralist is interested in creating a generalization about a language or dialect, often to create a dictionary or grammar, on the basis of what real speakers *say*. However, a structuralist might interview only one real speaker, or just a few speakers, in the belief that every native speaker shares in the linguistic system of the language or dialect. Some of our colleagues, the sociolinguists, follow the structuralist model of making generalizations about a group based on evidence from a small number of speakers. A generativist, alternatively, is interested in creating a generalization about a language or dialect, usually a grammar, on the basis of what a speaker *thinks*. While the generativist would agree with the structuralist that each speaker shares in the linguistic system of the language under study, and so just one or a few speakers can serve as the basis for a generalization, the underlying rules that generate real speech are the target. The generativist needs to test examples of speech against native speaker intuitions, not usually to collect and multiply examples of real speech. So, structuralists are interested in speakers' *memory* of language, while generativists may be said to be interested in speakers' *processing* of language. Both of them make general statements about language, most often from rather restricted evidence, in accord with the axiom of systematicity of the language or dialect.

The empirical linguist, on the other hand, does not necessarily assume that each native speaker shares the same linguistic system, or conversely that speakers possess only one linguistic system or inventory, and instead wants to collect great quantities of real speech from a great many speakers in order to describe what people actually say. As Pederson earlier represented the Gilliéronian point of view, the empirical linguist is interested in 'a few words of a language, not the language itself and its universe of discourse.' Empirical linguists typically employ the grammatical categories postulated by structuralists and generativists, but they test each category empirically to assess its reality in use. Empirical linguists also test the distribution of words, whether as lexical units or as they embody morphological markers or pronunciations, not necessarily as elements in a contrastive system but for themselves, to observe the dynamics of real speech by real people in samples taken from whole regions or communities. When an empirical linguist makes a generalization, it boils down large quantities of speech from many sources, as opposed

to the structural or generative prediction of the speech of the group on the basis of one or a few individuals. Clearly, some of our colleagues in sociolinguistics follow this empirical model more closely than they do the structural or generative models. Corpus linguists are also members of the empirical group, and the study of corpora has, especially here in England, proven to have a number of industrial applications. Empirical linguistics is not a replacement for structural or generative linguistics; it is just different, and there is room for all sorts.

In the future of dialectology, I hope that we will understand that what we do is not the same as what our structuralist and generativist colleagues do – and I hope that each of our students will understand how and why it is different on theoretical grounds.

Let us turn now to collection, encoding, and analysis of data. As we enter our future, we need to take best advantage of modern recording and computing resources. Technology does not determine what we do, but it does offer particular opportunities that we cannot afford to ignore. In the last generation, the great technological advance was the portable tape recorder. Dialectologists have come to use it routinely, but, as Raven McDavid has written in his last-published book review, it has never been a cure-all (1985: 60):

> I cannot be completely objective about my own experience, but for all my early ineptitude, I think I developed a large measure of cooperation from those I interviewed, but then, I knew I was undertaking something different from my usual experience, and I dearly wanted it to work out. At the beginning, I don't think the tape recorder would have made me a better field worker. For all its convenience, it cannot substitute for interest, imagination and training, and the determination to do a good job. It can improve the work of even the best field worker by providing a permanent record and picking up far more evidence than the best can record with the unaided ear, but no student should take the tape recorder out for his interviewing without previously demonstrating the ability to get along without it.

This advice certainly applies again now, as we continue to try to cope with the computer as a technological aid, and we find ourselves in the same situation with computers as McDavid described for the tape recorder. Successful integration of very human, ethnographic data collection with computer encoding and analysis will continue to be a pressing need.

Two particular aspects of data collection are not directly linked to technology but still seem clearly to be the way of the future. First is the nature of the dialect field interview. Our colleagues the sociolinguists have convinced just about everyone, including most of us, of the value of discourse-based interviews as opposed to the long-question-with-a-short-answer method. Their arguments about interview style and its effects on results are compelling. Still, we need not abandon our wish to elicit particular words or other usages. Lee Pederson has described a modified interview technique that addresses our need for specific elicitation targets in the context of a discourse-style interview (1996a, 1996b; Pederson and Madsen 1989). He offers 'a framework for a tape-recorded interview' of approximately three-hour duration, composed of four 45-minute sections, which has already received extensive use in on-going field work in the Western states. There are 360 elicitation targets, but these are embedded under a dozen topical headings that allow the field worker to introduce the general topic and to direct conversation within it, rather than to ask pointed questions. The resulting interview should be transcribed in full, with tags in the transcription to identify elicitation targets but with full opportunity to preserve the continuous speech of the speaker for analysis of verb form frequency and other discourse features (1996a: 54-59). Interviews on this model are not 'danger-of-death' performances, but they do establish a consistently informal interview style, and they should be directly comparable with the large body of sociolinguistic discourse data obtained under similar, informal interview conditions.

Another very important aspect of data collection for dialectology is the emergence since the planning of first-generation Linguistic Atlases of reliable survey research methods. This is not the time for me to recapitulate what my colleagues and I have written about at length (Kretzschmar et al. 1993, Kretzschmar and Schneider 1996), so I will not offer any detailed argument here. Suffice it to say that I believe that we have no choice but to employ valid survey research methods. We should take no short cuts. We should not follow traditional speaker selection methods just because our teachers used them. If what I have argued earlier is true, that the distinction of dialectology from competing linguistic theories is based upon large-scale, planned, empirical survey research, then we simply must adopt the methods that are accepted, even required, in the other modern empirical sciences. We need not give up anything to do so. There is no conflict between the randomization techniques of survey research and the traditional goals of dialect surveys to achieve balanced regional and social coverage. All we have to do is learn enough about modern sampling techniques to do within the accepted parameters whatever we wanted to do anyway. And we need not give up historical comparison with the first generation of Atlas studies to do so, as

Ellen Johnson's book on the Southeastern lexicon has demonstrated with her statistically-valid paired-sample survey, matching new 1990 interviews with 1930s interviews (1996). Since we can very well apply currently-accepted scientific methods to our discipline, it seems quite clear to me that nobody would or should take what we do very seriously in future if we fail to use them.

One last aspect of data collection that I think requires comment is what we consider to be 'data.' In the future we should not only collect data about the speech produced by speakers but also data about how people perceive speech features and dialects. The breakthrough book in this area is Dennis Preston's *Perceptual Dialectology* (1989), but as Preston's forthcoming *Handbook of Perceptual Dialectology* will show (in press), there has been active work in the area for over fifty years in Japan and the Netherlands. The more we know about the distribution of individual dialect features, the more we realize that they do not pattern as neatly across the land as people (including dialectologists) expect them to. We all perceive clear differences between dialects that, in the end, we cannot document except by recourse to a small number of selected diagnostic features. Our work in future must recognize the difference between speech as produced and speech as perceived, and we need to begin collecting perceptual as well as production data. Our best work will incorporate both aspects, as Macafee's paper in this volume shows.

I have already mentioned how interviews with Pederson's new framework could be transcribed, with tags to identify elicitation targets. This, I believe, will be the future of data encoding for dialectology. I have spent many years working on a database approach to encoding questionnaire responses, and I will spend more time yet trying to get all the rest of the first-generation American Atlas data into digital form in a database structure. However, tagged transcripts should be the format for digital storage of any new surveys. My earlier choice of the database structure was conditioned by the state of computer processing at the time. In the 1980s when Lee Pederson and I were building database structures, desktop computer processing was not very good or fast for text files, but was very good indeed for databases. Moreover, highly efficient storage of responses in databases went along with the shortage of mass-storage space that existed then. Pederson designed his databases to fit on floppy disks; a few years later I designed mine for 10Mb and 20Mb hard drives. Now there is no shortage of mass storage – I have 500 times as much hard-drive space as I used to have – and processors are fast enough for full-text searches in large files. And there are substantial advantages to using tagged texts. For one thing, the automatic taggers that already exist for corpus linguistics, such as the BNC automatic tagger (cf. Aston and Burnard 1997), could be applied immediately to dialect interviews, which would render

the information in them much more accessible in more ways than ever before. As early as 1989 Pederson illustrated additional uses of tagged texts [Fig. 1; Pederson and Madsen 1989: 20]. You can see how additional information is coded in Pederson's transcription. There are unique markings to identify the elicitation targets, and to label other relevant aspects of the text such as to keep the field worker's speech separate from the informant's speech. Each type of information subsequently can be extracted automatically to form concordances or different kinds of indices. Today we should all be using SGML, the standard markup language for texts, instead of Pederson's homemade tagging, but the point is the same, and SGML allows for the customized tags that we need for the special purposes of dialectology. In the 1980s dialectologists needed to be inventive to bend emerging computer technology to our own purposes; today and in the future, we need instead to be receptive to developments that are already out there, and inventive enough to see their applications for our own work.

After collection and encoding of data, we come to the issue of analysis. The future of analysis in dialectology is like the future of theory: we need some. It is not enough merely to collect and preserve data. In the past it has been an axiom of dialectology that our goal was full and fair presentation of our data. It is understandable that dialectologists have been preoccupied with display and publication because we have had such large quantities of data, in fact so much that in America most of it never has achieved publication in any better medium than microfilm. Orton did far better with *SED* (1962-71), but even *SED* analyses have come late (e.g. Orton and Wright 1974; Orton, Sanderson, and Widdowson 1978; Upton, Sanderson, and Widdowson 1987). A neat statement of our conservative position came at the end of an article called 'Inside a Linguistic Atlas,' written by Raven McDavid and the editors of the Middle and South Atlantic Atlas project, including me (McDavid et al. 1986: 404-05):

> The business of the linguistic atlas is to provide the evidence, not verdicts. It would be silly to say that we are not interested in what may come out of our materialsñnaturally we are, and we will remain in the forefront among those who interpret themñbut we cannot afford to make interpretations before we present the data or, worse, to insert interpretations into our presentation of the materials. Those of us on the inside have a responsibility to get the data out, and this we will do, in time, as clearly, fully, and objectively as possible.

While I agree with and am still acting upon the last statement here, the irony of this

assertion is twofold. First, the failure of dialectologists to provide analysis of their materials has actually *prevented* publication of the data. After a long series of grant applications to the National Endowment for the Humanities, it has become clear to me that the proposal reviewers are not willing to fund the Atlas project just for its intrinsic merit. They have trouble seeing what the Atlas data is good for, and increasingly so the more time passes after the interviews. The second irony is that the greatest success of the American Atlas was the benchmark set of analyses by Kurath, Atwood, and McDavid (1949, 1953, 1961), which were all based on a subset of the data and were produced before publication of any but the first American regional Atlas. These studies, particularly the *Word Geography* and the *Pronunciation of English in the Atlantic States*, caught the imagination not just of dialectologists but of a wide range of readers, from philologists to cultural geographers. They established scholarly expectations about American dialects for a generation. On the basis of this experience, dialectologists in the future ought to publish working analyses early and often!

The nature of our analysis, I believe, will also be different in the future. Lee Pederson has described an essential contradiction in the goals of dialectologists and the expectations of their audience (1995: 39):

> The [Linguistic Atlas] method carries analysis through an enumeration of features and records them in lists and/or reports them in maps. Such analytic word geography ends its work at this point in a taxonomy of observed sociolinguistic facts. But the research invariably implies more than that because planners, editors, and their critics fail to characterize the work at hand. For that reason, a reader expects an identification of dialect areas and a description of dialects within those geographic divisions in a concordance of social and linguistic facts projected across space and through time. . . . Both [Hans Kurath and Harold Allen in their association of American settlement patterns with speech areas] synthesize geographic, historical, and social facts in their reorganization of evidence in an effort to meet the unreasonable expectations of linguistic geography.

The name 'dialectology' itself raises the expectation that our job is to describe and to find the borders of dialects. In the past our teachers believed that they were conducting research in historical linguistics, and well-bounded dialects are one of the tools of study in historical linguistics, which in theory is essentially a structuralist enterprise. Yet another irony in our field is that American and some British dialectologists today

regularly question the status of dialects and their borders, for example in papers by Davis, Houck, and others about Kurath's Midland region (Carver 1987; Davis and Houck 1992; Frazer 1987, 1993, 1994; Johnson 1994), or Davis, Houck, and Upton's paper from the last Methods conference about our general failure to draw convincing dialect boundaries (1997), or Davis, Houck, and Horvath's paper in this volume – all this while our sociolinguist and structuralist colleague William Labov quite happily describes well-bounded dialect areas, including the American Midland. Now and in the future, if we dialectologists believe ourselves to be engaged in empirical linguistic study, then we should have much broader scope for analysis than our teachers did. For this reason I have renamed the 'Dialectology' course at my university to 'Language Variation,' so as to enlarge the expectations of my students. In future, dialectology may lose its name if others do what I have done, and that may not be a bad thing if it would help us to be clear about what we and others expect from our work.

As my own writing reveals (e.g. 1992, 1996a; Kretzschmar and Schneider 1996), I believe that quantitative analysis will be the hallmark of future analysis in dialectology, and also in empirical linguistics more generally. I will not belabor that point here. I would like to confirm, however, that there is room in dialectology for many kinds of research and analysis. I have mentioned perceptual dialectology as an important constituent of our work, and that branch of study verges on psychology. Those who study language and identity, following LePage and Tabouret-Keller (1985), will also find friends among the dialectologists (e.g. Lanehart 1996). Qualitative research of many kinds should exploit the ethnographic side of dialectology. Finally, as much as I have suggested that we need to break free from exclusive concentration on historical linguistics, we should in future continue to contribute to that discipline (e.g. Kretzschmar 1996b). The future of dialectology should be pluralistic in its approaches to analysis.

My last point today concerns our presentation of our data and the results of our analyses. The publication of many American Atlases has been held up for years by the shear weight of the data, and it is no more practical to produce Atlases on paper today than it has been for the last fifty years. I believe that the best solution for this problem is publication via the Internet, and to that end we have created a Web site (on a server at the Georgia editorial office, http: //us.english.uga.edu) that provides a framework for comprehensive display of Linguistic Atlas data and for visualization of the data on maps – with every map created to the user's order. The problem of editing all of that data is still present, and it will take us time to get all of it onto our Atlas Web site, but when we have done that we will have a fully interactive Atlas with comprehensive coverage of the first generation of American regional projects. I believe

that, within the foreseeable future, we will have largely fulfilled Hans Kurath's dream of an American Atlas.

Let me take a few moments to illustrate the Web site for those of you who may not have accessed it. The opening screen of the Web site is followed by a clickable map of the American Atlas regions. We have made the Middle and South Atlantic project (LAMSAS) operational, with local pilot funding for creation of the site based on prior computerization of LAMSAS data (see Kretzschmar et al. 1993). Base maps for the site have come from the US government 'Tiger' Internet site, which provides the digital images for no cost. Next comes another clickable map, this time just of the LAMSAS region, from which users can find detailed information about the people interviewed for the project. A click on any state brings up a more detailed map of the state on which are plotted the locations of all the people interviewed, and a click on any person's plot reveals a full description of their particulars (age, sex, type, educations, etc.). It also allows the user to ask what that person responded to any of the questions of the survey questionnaire. Another screen shows the non-informant-based services available: 1) browsing the data, 2) searching the data, 3) generating maps of particular items, 4) a table of phonetic symbols used for the project, and 5) utilities available for downloading, notably a TrueType font with which users may view and print Atlas phonetic symbols on their own PC-compatible computers. We have also implemented a demonstration project for linguistic survey research via email. Users may browse lists of data for any survey question, or search the database for words in which they are interested. In order for users to make maps, they first select the survey question in which they are interested, and the site produces a list of all the answers; users then select the answer that they want to map. Software underlying the Web page then produces a map to order. The key feature of the Web site is that it is an interactive resource. It is extensively cross-linked in addition to allowing the user to ask several different kinds of questions of the database. When we have more data available, it will be possible to ask questions across several projects at once. We hope to automate several of the quantitative analytical methods that we have developed, so that those can be run by users online in real time. The Web is the research tool of the future, and we have it now.

Technology will certainly be important to us in presentation of our data and results, as in all aspects of data collection, encoding, and analysis, but there is another equally important point to make. I believe that we need to accept as central to our purposes the goal of informing the public, not just the scholarly community, about the facts of language variation, especially as that information can affect education and public policy. And technology can help us to achieve that goal, too. The existing

Linguistic Atlas Web site, even given the limited area for which its software is fully activated, has become the most accessible source of information about regional American English for the general public (see also http://www.ling.upenn.edu/phono_atlas/home.html, Labov's site, which is more oriented to professionals; it now offers sound samples). The scholarly works that I have mentioned have been largely technical in nature, best suited for specialists. The *Dictionary of American Regional English* (Cassidy, Hall, et al. 1985-) is much more accessible to the general public in its content, and it has sold thousands of copies to libraries, but unlike the Atlas Web site it is not available in the millions of homes in America and abroad that can access the Internet. An earlier version of the Atlas Web site averaged over 10,000 'hits' per month during its peak of use, a very large number of them from non-academic addresses. I regularly get email from high-school students who ask for help with class papers, and so far I have been able to answer every one. Many of the speech patterns documented by the first American Atlases still exist, so public and expert users can look up words and pronunciations that they have noticed in their own speech. Speakers of different ages and social circumstances, in cities as well as rural areas, participated in the American interviews, including a number of African Americans, so the Atlas surveys include a wide spectrum of American speech – and thus a great many Americans can use the Atlas Web site to find out about American words that are still relevant to them and which reflect their American cultural heritage. We hope to do better still with the site in future, now that we know that it has become such an active public resource. We have added to our new site more explanatory information keyed to the needs of our non-specialist audience. We also hope to help the public to ask appropriate questions of our data, and to allow them to ask questions using a natural-language interface in an expert system built with Artificial Intelligence tools (cf Rochester and Kretzschmar 1998). We have felt a little like the recent popular movie about a man who built a baseball diamond in an Iowa cornfield, whose credo was 'If you build it, they will come.' We built it, and they came, and because they came we now see that we have to build it better for them.

The need for this public access has been made clear to us all in the debate over Ebonics this past year. The extent of misinformation about language variation that has been displayed by politicians, pundits, and even in newspaper editorials by the normally sensible American public, has been nothing short of astounding. In large part the misinformation results from the overselling of Standard English in our educational system, the same reason that my students and many more educated people believe in the Star Trek Syndrome. The fact that a great many people have visited our Web site is a hopeful sign, a measure of the public's willingness to seek out new

information about something on which they already have definite opinions.

This curiosity on the part of the public, and the fact of extensive variation in contemporary English, both show that the future of dialectology is not bleak, no matter what things look like on Star Trek. The foundations of study laid down by our teachers in the first generation of Linguistic Atlas work still serve us well, and we have a great many technological and other kinds of improvements that we can make as we enter our own Next Generation. The most important element for our future, among all of these factors, is us ourselves. As Raven McDavid said, nothing can 'substitute for interest, imagination and training, and the determination to do a good job.' We will succeed in the next century, and beyond, to the degree that we are determined, that we can interest and train ourselves and our students, and that we have the imagination to see the way forward. I believe that we will make Harold Orton and all of our teachers proud.

REFERENCES

Aston, Guy, and Lou Burnard. 1997. *The BNC Handbook.* Edinburgh: Edinburgh University Press.

Atwood, E. Bagby. 1953. *A Survey of Verb Forms in the Eastern United States.* Ann Arbor: University of Michigan Press.

Bailey, Guy. 1997. When did Southern American English begin? In Edgar Schneider (ed) *Englishes Around the World 1*, pp. 255-75. Amsterdam: Benjamins.

Bailey, Guy, and Garry Ross. 1992. The evolution of a vernacular. In Matti Rissanen, et al. (eds) *History of Englishes*, pp. 519-31. Berlin: Mouton de Gruyter.

Carver, Craig. 1987. *American Regional English.* Ann Arbor: University of Michigan Press.

Cassidy, Frederic, Joan Hall, et al. 1985-. *Dictionary of American Regional English.* Cambridge, MA: Belknap Press.

Davis, Lawrence M. 1973. The diafeature: an approach to structural dialectology. *Journal of English Linguistics* 7: 1-20.

Davis, Lawrence M., and Charles L. Houck. 1992. Is there a Midland dialect area? – Again. *American Speech* 67: 61-70.

Davis, Lawrence M., Charles L. Houck and Clive Upton. 1997. The question of dialect boundaries: the SED and the American atlases. In Alan Thomas (ed) *Issues and Methods in Dialectology*, pp. 271-83. Bangor: Linguistics Department, University of Wales

Bangor.

Frazer, Timothy. 1987. *Midland Illinois Dialect Patterns*. PADS 73. Tuscaloosa: University of Alabama Press.

Frazer, Timothy. 1993. *'Heartland' English: Variation and Transition in the American Middle West*. Tuscaloosa: University of Alabama Press.

Frazer, Timothy. 1994. On transition areas and the 'Midland' dialect: a reply to Davis and Houck. *American Speech* 69: 430-35.

Johnson, Ellen. 1994. Yet again: the Midland dialect. *American Speech* 69: 419-29.

Johnson, Ellen. 1996. *Lexical Change and Variation in the Southeastern United States 1930-1990*. Tuscaloosa: University of Alabama Press.

Kretzschmar, William A., Jr. 1992. Isoglosses and predictive modeling. *American Speech* 67: 227-49.

Kretzschmar, William A., Jr. 1995. Dialectology and sociolinguistics: same coin, different currency. *Language Sciences* 17: 271-82.

Kretzschmar, William A., Jr. 1996a. Quantitative areal analysis of dialect features. *Language Variation and Change* 8: 13-39.

Kretzschmar, William A., Jr. 1996b. Foundations of American English. In Edgar Schneider (ed) *Focus on the USA*, pp. 25-50. Philadelphia: John Benjamins.

Kretzschmar, William A., Jr. 1997. American English for the 21st century. In *Englishes Around the World* 1, edited by Edgar Schneider, 307-24. Amsterdam: Benjamins.

Kretzschmar, William A., Jr., and Edgar Schneider. 1996. *Introduction to Quantitative Analysis of Linguistic Survey Data: An Atlas by the Numbers*. Thousand Oaks, CA: Sage.

Kretzschmar, William A., Jr., et al. 1993. *Handbook of the Linguistic Atlas of the Middle and South Atlantic States*. Chicago: University of Chicago Press.

Kurath, Hans. 1949. *A Word Geography of the Eastern United States*. Ann Arbor: University of Michigan Press.

Kurath, Hans, and Raven I. McDavid, Jr. 1961. *The Pronunciation of English in the Atlantic States*. Ann Arbor: University of Michigan Press.

Labov, William, and Sharon Ash. 1997. Understanding Birmingham. In C. Bernstein, T. Nunnally, and R. Sabino (eds) *Language Variety in the South Revisited*, pp. 508-73. Tuscaloosa: University of Alabama Press.

Lanehart, Sonja. 1996. The Language of Identity. *Journal of English Linguistics* 24:322-33.

LePage, Robert, and Andree Tabouret-Keller. 1985. *Acts of Identity*. Cambridge: Cambridge University Press.

McDavid, Raven I., Jr. 1985. Review of Edward Ives, *The Tape-Recorded Interview*. *Journal*

of English Linguistics 18: 54-63.

McDavid, Raven I., Jr., et al. 1986. Inside a linguistic atlas. *Proceedings of the American Philosophical Society* 130: 390-405.

Orton, Harold, et al. (eds). 1962-71. *Survey of English Dialects.* Intro. and 4 vols. in 3 pts. Leeds: Arnold.

Orton, Harold, Stewart Sanderson, and John Widdowson (eds). 1978. *The Linguistic Atlas of England.* London: Croom Helm.

Orton, Harold, and Nathalia Wright. 1974. *A Word Geography of England.* London: Seminar.

Pederson, Lee. 1995. Elements of word geography. *Journal of English Linguistics* 23: 33-46.

Pederson, Lee. 1996a. LAWCU project worksheets. *Journal of English Linguistics* 24: 52-60.

Pederson, Lee. 1996b. LAMR/LAWS and the main chance. *Journal of English Linguistics* 24: 234-49.

Pederson, Lee, and Michael Madsen. 1989. Linguistic geography in Wyoming. *Journal of English Linguistics* 22: 17-24.

Preston, Dennis. 1989. *Perceptual Dialectology.* Dordrecht: Foris.

Preston, Dennis, ed. in press. *Handbook of Perceptual Dialectology.* Amsterdam: John Benjamins..

Rochester, Eric, and William A. Kretzschmar, Jr. 1998. Intuitive interfaces for the retrieval of linguistic data, ACH/ALLC, Debrecen (Hungary).

Trudgill, Peter. 1975. *Accent, Dialect, and the School.* London: Arnold.

Upton, Clive, Stewart Sanderson, and John Widdowson. 1987. *Word Maps.* London: Croom Helm.

Wright, Joseph. 1898. *English Dialect Dictionary.* Vol. 1. London: Frowde.

URLs:

Labov's site: http://www.ling.upenn.edu/phono_atlas/home.html.

Linguistic Atlas site: http: //us.english.uga.edu.

Star Trek site: http://www.startrek.com (the official site; there are others which include sound samples.)

Address
William A. Kretzschmar, Jr.
Department of English
Park Hall
University of Georgia
Athens
GA 30602-6205
USA
<billk@atlas.uga.edu>