Dialectologia 1 (2008), 135-156.

A REPORT ON THE INTERNATIONAL CONFERENCE: GEOLINGUISTICS AROUND THE WORLD

Chitsuko Fukushima & David Heap

Niigata Women's College (Japan) / University of Western Ontario (Canada)

fukusima@elle.nicol.ac.jp / djheap@uwo.ca

The National Institute for Japanese Language held its 14th international conference, for the first time on dialects, on 22-23 August 2007 in Tokyo. The symposium was entitled "Geolinguistics around the World" and focused on Current Trends in Geolinguistics around the World (Day 1) and Application Techniques of Linguistic Atlases (Day 2). The lecturers were Shinji Sanada (Osaka University, Japan), Hans Goebl (University of Salzburg, Austria), Takuichiro Onishi (National Institute for Japanese Language, Japan), Lee Sang Gyu (National Institute of the Korean Language, Korea), Iwata Ray (Kanazawa University, Japan), Joachim Herrgen (University of Marburg, Germany), Heinrich Ramisch (University of Bamberg, Germany), and Maria-Pilar Perea (University of Barcelona, Spain). The commentators were Chitsuko Fukushima (Niigata Women's College, Japan) and David Heap (University of Western Ontario, Canada).

The National Institute of the Japanese Language (NIJL) must be warmly congratulated on the publication of the *Grammar Atlas of Japanese Dialects* (GAJ, 1991-2006): this is an important milestone not just in Japanese geolinguistics but also for the field of linguistic geography internationally. While the GAJ undoubtedly provides material for future linguistic studies for many decades to come, today it offers us the ideal occasion to reflect on common challenges and different approaches, both in the

techniques of linguistic cartography and in our theoretical reflections on the goals of geolinguistic research. In this sense, we are doubly grateful to the NIJL: for the GAJ itself and also for this symposium, which brings together such a diverse range of dialectologists from the East Asian and Western European scholarly traditions.

The lectures given in the symposium are as follows:

(Day 1)

Takuichiro Onishi, “Mapping Japanese Dialects” (Hereafter Onishi 1)

Lee Sang Gyu, “Dialect Data Processing & Linguistic Maps" (Lee 1)

Iwata Ray, “Geolinguistics of Chinese (1): History and Recent Trends” (Iwata 1)

Shinji Sanada, “The 'Glottogram': A Geolinguistic Tool Developed in Japan” (Sanada)

Joachim Herrgen, “Dialectology - Digital and Interactive: The Digital Wenker Atlas – DiWA” (Herrgen 1)

Heinrich Ramisch, “Current Trends in British Geolinguistics Linking the Past with the Present” (Ramisch 1)

Maria-Pilar Perea, “Techniques in Catalan and Spanish Linguistic Atlases” (Perea 1)

Chitsuko Fukushima: Commentary on Day 1 lectures

(Day 2)

Takuichiro Onishi, “Analyzing Dialectal Distributions of Japanese” (Onishi 2)

Lee Sang Gyu, “Creating Dialect Maps Using Map Maker” (Lee 2)

Iwata Ray, “Geolinguistics of Chinese (2): Interpretation of Linguistic Map” (Iwata 2)

Hans Goebl, “Dialectometry: Theoretical Prerequisites, Practical Problems, and Concrete

Applications (Mainly with Examples Drawn from the 'Atlas Linguistique de la France', 1902-1910)" (Goebl)

Joachim Herrgen, "From Dialect to Variation Space: The 'Regionalsprache.de' (RE.DE) Project" (Herrgen 2)

Heinrich Ramisch, "Analysing Linguistic Atlas Data: The (Socio-) Linguistic Context of H-dropping" (Ramisch 2)

Maria-Pilar Perea, "Catalan Geolinguistics and New Technical Procedures" (Perea 2)

David Heap: Commentary on Day 2 lectures

The first part of this report is written by Chitsuko Fukushima based on the commentary for Day 1, and the second part is written by David Heap based on the commentary for Day 2.

1. Day 1

The lectures given on Day 1 were reports on current trends in geolinguistics around the world. Various ways of making linguistic maps were introduced. This variation is due to differences in aims, i.e. why we make linguistic maps, or the differences of viewpoints, i.e. what features we focus on in making linguistic maps.

1.1. *Reasons for making linguistic maps*

Why then do we make linguistic maps? Linguistic geography was born at the end of

the 19th century. There have been two answers to this question since then.

The first answer is to discover dialect boundaries or demarcate dialect areas. In western countries, linguistic geography followed comparative linguistics. Neogrammarians advocated the regularity of phonetic rules, and Wenker expected it to be reflected in dialectal distributions. He wanted to find dialect boundaries in German dialects and started a postal survey by sending a list of sentences in Standard German to teachers and asking them to translate them into local dialects. Thus the first linguistic atlas was Wenker's *Linguistic Atlas of the Rhine Province* (1978) (Herrgen 1). This expanded to the survey of the whole country. J. Gilliéron also took up a survey to compare dialects in France at the end of the nineteenth century. In Japan, a postal survey was carried out to decide the boundary of Eastern and Western dialects in Japanese. The first linguistic atlases in Japan were published in 1905 and 1906 (Onishi 1). In China, Kalgren (1915-1926) initiated the reconstruction of proto-sound-systems and surveyed the sounds of Chinese characters in dialects. The studies resulted in the classification or demarcation of dialects (Iwata 1).

In all these cases, the aim of the surveys was clearly to classify or demarcate dialects. When dialect demarcation is the aim, a single line of dialect boundary is expected on a map, but in fact bundles of isoglosses are found. We now know that each linguistic feature shows different geographical distributions.

Linguistic geography is an independent discipline in linguistics, so it has its own aims. Thus the second answer to the question "Why do we make linguistic maps?" is: to 'read' a linguistic history from geographical distributions on a map. Or you can say that it is to explain the process of language change based on the linguistic variation. Gilliéron is the linguist who did this for the first time. He used maps from his *The Linguistic Atlas of*

France to develop the principles used to interpret linguistic maps as well as the basic philosophy of linguistic geography. In Japan, there were scholars like Kunio Yanagita who read history from linguistic maps, but then Willem A. Grootaers introduced Western-style linguistic geography to Japan and Takesi Sibata established basic methodologies of linguistic geography (Onishi 1). Iwata examines the patterns and causes of language change based on each map, believing that the aim of linguistic geography is to explain the history of language (Iwata 1&2).

Grootaers and Sibata were oriented towards interpretive maps. Actually, any linguistic map is interpretive unless transcribed forms are printed verbatim on the map. You cannot make linguistic maps without classifying and ordering the original data. This procedure requires linguistic knowledge, which is the basis of linguistic analysis. Thus linguistic maps have the following purposes: 1) linguistic maps as a source of data for linguistic experimentation and 2) linguistic maps as a place for training in linguistics.

According to Herrgen, "Theoretical approaches in linguistics under discussion have at all times been tested on dialectal data, so that dialectology at the same time represents a dynamic linguistics laboratory." If dialectology is a laboratory, then map-making is experimentation. What kind of linguistic surveys are planned and how the original data are analyzed once acquired, all depend on what we consider to be language.

Lee says that "the very process and practice of creating linguistic maps are actually an extremely educational vehicle in improving the analytical capabilities" of students working on linguistic data. Map-making has educational implications and is useful in training students to analyze linguistic data. They can empirically learn the patterns of language changes, such as phonetic changes and analogy, by classifying and analyzing a variety of linguistic data from the viewpoints of phonology, morphology, lexicology,

semantics etc. But that is not the only use of linguistic maps. When you find distributions either as you expected or as you did not expect, the results stimulate further research. Most geolinguists should be familiar with this experience. When you make linguistic maps manually, you enjoy the process. Nowadays when we make linguistic maps using a computer, linguistic maps pop up on the screen, and this impact is much stronger, according to some scholars (personal communication).

1.2. *What kind of linguistic maps to make*

The next question is what kind of linguistic maps we should make. We can examine what linguistic features should be mapped, but, considering map-making techniques, here we examine how linguistic variation can be expressed. If you draw an isogloss, a dialect boundary is clearly shown. But, as Ramisch says, if you want to show the transitional zones in which language changes gradually, symbol maps are useful. The map-making system SEAL developed by one of these authors also uses symbols which can show transitional zones (C. Fukushima, 2000, "Using a personal computer to grasp dialectal variation", *Dialectologia et Geolinguistica* 8, 37-52). Some might prefer 'hatching' the zone. If more than one linguistic map shows a similar transitional zone, you can integrate them and show them in a single map, as discussed below.

Linguistic geography started with the focus on the dialects spoken by NORMs (nonmobile, older, rural males). It expanded its field to include social dialects, thus merging with sociolinguistics. Sanada's definition of linguistic geography which includes the "new" feature, sociolinguistics, is as follows:

> Linguistic geography presumes that language diffuses geographically

> and socially and explains using maps and graphs how new expressions spread or advance, what the dynamics of the contact, change, and extinction are. (Sanada)

Linguistic maps show the linguistic variation of a given generation on the map. Sanada reports about a 'glottogram', a geolinguistic tool developed in Japan. A glottogram, which shows variation as the graph of localities by age, was born when scholars pursued some kind of "thickness" such as differences in age and generation that was not shown on existing geolinguistic maps. We often want to know about the language of different generations when we are working on linguistic maps of the elderly. The linguistic survey of Itoigawa, the first and most famous small-area linguistic survey in Japan, developed not only a glottogram but also other new techniques, such as a survey of every hamlet in the area, a survey of all the residents in a key hamlet, a geolinguistic survey of the young generation (junior high school students) to compare with data from the elderly. Interest in the related items led us to the survey of meanings, words not used but only understood (i.e. passive vocabulary), and situational differences.

1.3. *Making linguistic maps using a computer*

As Lee stated, there is a big difference between map-making by hand and by computer. What then has changed?

In 1983 when one of these authors developed a system to make linguistic maps using a personal computer (C. Fukushima, 1983, *An Approach to Computer-Assisted Linguistic Geography: SEAL Users' Manual*), she thought of the merits of using computers as 1) accuracy, 2) beauty, and 3) ease of reanalysis. Manual labor often results

in errors, redrawing is troublesome, and the inclination of symbols cannot be parallel, etc. If the data and commands are saved on a computer, then reanalysis is always possible and handy.

After making linguistic maps using a personal computer for more than two decades, this author now considers linguistic data analyzed using a computer to be a "linguistic database" or "language corpus", not just "atlas data".

One of the difficulties in computer-assisted map-making was that the ideal software did not exist. Thus software had to be developed. Now we can make use of off-the-shelf software but we still need to adapt software suitable for our purpose. Lee explained the history of software for making linguistic maps. There are four steps in the process of map-making using a computer: 1) Electronic data production, 2) Sorting and mapping data, 3) Comparing, integrating, superimposing, and linking data, and 4) Publishing linguistic maps.

1.3.1. Electronic data production

Electronic data production is the first step to map-making using a computer. You have to make electronic data from paper media such as questionnaires and books, or from audio-visual media such as tapes and videotapes.

Working from text data is the most troublesome and often gives rise to transcription errors. If you input on the Internet, then the input becomes electronic data. Recently sounds and movies are easily integrated.

There are two kinds of electronic data. One is data from recent surveys, with examples such as the *Grammar Atlas of Japanese Dialects* (GAJ) in Japanese, *BBC*

Voices Project in English, and *The Corpus Oral Dialectal* (COD) in Catalan. The other is data from old surveys, such as the *Digital Wenker Atlas* (DiWA) in German, *Computer Developed Linguistic Atlas of England* (CLAE) and *English Dialect Dictionary* in English, and Antoni M. Alcover's data in Catalan. You can compare old data with recent data and trace language change in real time.

1.3.2. Sorting and Mapping Data

When you make linguistic maps based on electronic data, the data can be sorted by linguisitc, geographic or social criteria. At this stage, the collected linguistic data become a database.

Onishi says "Each linguistic map is a model in which geographic information is classified and expressed on a hypothesis chosen from possible hypotheses." A linguistic map only reflects one hypothesis. The GAJ data are published on the NIJL web page; thus anyone can use them to make new linguistic maps based on different hypotheses.

According to Ramisch,

> A new generation of linguistic atlases has come into existence which is profoundly influenced by modern computing. Two aspects seem to be particularly noteworthy in this context. First, it is possible to record and to store large amounts of data in the form of **databases**. Secondly, the data can **be searched automatically**, be processed and be visualized effectively by computer cartography. (Ramisch)

The data for geolinguistics are stored in a variety of data types within a database. Words and other linguistic forms are sorted by means of queries and maps are drawn from these results.

Perea states that the mapping techniques are applied to **corpora** not to atlases. She defines a corpus as a "complete collections of linguistic data" and a geolinguistic database without doubt constitute a corpus.

1.3.3. Comparing, integrating, superimposing, and linking data

Also included in this section are examples from lectures on Day 2. You can compare, integrate, superimpose, and link data without using a computer if the procedure is simple. However, the process can be easier and more sophisticated if you use a computer. Furthermore, it is possible to analyze massive amounts of data using computer methods.

1.3.3.1. Integrating linguistic data of different types

Linguistic maps related to phonetic change do not always show the same geographic distributions. If you integrate such maps, a transitional zone is clearly displayed. For example, in the SEAL system (C. Fukushima 2000, ibid.), the frequency of specific forms at each locality are counted and expressed as different sizes of symbols on a map (a 'contrast map'). Ramisch demonstrates this type of data integration (Ramisch 2). These are examples of lexical diffusion.

1.3.3.2. GIS (Geographic Information System) and superimposing maps

If you want to compare data from different surveys rather than from the same survey, or if you want to compare linguistic data with non-linguistic (e.g. demographic or

economic) geographic information, GIS can be a useful tool. In the Digital Wenker Atlas (DiWA) the use of GIS makes it possible to splice together the three individual map sheets seamlessly into a single map and to superimpose a Wenker map over any other cartographic representation in electronic form, thus directly compare the two using a transparency function (Herrgen 1). Onishi demonstrated superimposing a GAJ map with maps of non-linguistic geographic information (Onishi 2); this is also possible because he uses GIS.

1.3.3.3. Old materials and new materials

Comparison of old materials and new materials must be included in this discussion of superimposing maps. DiWA compares a Wenker map and maps of recent surveys drawn from younger and older generations (Herrgen 1). Perea compares Alcover's data with the *Corpus Oral Dialectal* (COD) (Perea 1). In the latest version of the SEAL system, two maps from different surveys are directly superimposed on the screen, one in gray and the other in color (C. Fukushima, 2007, "Superimposing Linguistic Maps to Trace Linguistic Changes", *Linguistic Atlantica* Nos. 27-28, 40-45).

1.3.3.4. Linking with multimedia information

If you are using paper media, it is impossible to link with multimedia information. This only became possible after digital methods were introduced. You can get text data, sound data, and even video data just by clicking a point or passing a mouse over it. In the DiWA, locations are clickable and linked with biographic information, digitalized copies

of the original questionnaires, and with sound recordings (Herrgen 1). Ramisch maps the CLAE data using PCMAP and MS Word and says, "With an ODBC it is equally possible to connect localities with other Windows applications to display text files or to play audio files (thus a 'speaking' atlas)."

1.3.3.5. Statistical analysis

Computers are particularly useful for statistical analysis. When you want to compare a whole set of linguistic maps, dialectometry is the answer. As discussed below, Goebl introduces dialectometry using the results from several linguistic atlases (Goebl).

1.3.4. Publishing linguistic maps

There are two ways of publishing linguistic maps. One is to publish them on the Web. You can make linguistic maps or publish original data on the web. Iwata and his team are making linguistic maps on the web. The DiWA is open to the public on the web. Another example is the *Linguistic Atlas of the Middle and South Atlantic States* (LAMSAS) created by Kretzschmar. The other is to publish the linguistic maps and the original data on CD or DVD, as with Perea's *Corpus Oral Dialectal* and Alcover's data. When you publish maps, you can just publish the completed maps or you can let viewers make linguistic maps freely by querying the database and projecting the results on base-maps.

1.4. *Tasks in geolinguistics*

There remain two tasks in geolinguistics:

One is to publish, share, and integrate geolinguistic data. There have been many linguistic atlases separately planned and edited. They are national or local, or sometimes international. In Europe, the *Atlas Linguarum Europae* is being published based on a common questionnaire. Onishi has proposed to make the Japanese Dialect Network (JDNet) to share and integrate Japanese geolinguistic data. Lee has advocated the idea of an East Asian Linguistic Atlas, and so have other Japanese scholars.

Some of the points below need to be considered. First, if you plan to do a survey, you have to make a questionnaire which can be used in the whole area. Second, you have to decide the transcription system. Recently Unicode has become multi-lingual, but you need to decide whether to use IPA transcriptions or only characters which you can input on a standard keyboard. Third, as Perea mentions, you have to consider in which data types and in which media you store the data. Fourth, do you have access to the software for your purpose or do you have to develop it? Can you keep versioning it up as computers develop? Fifth, as discussed above, GIS is necessary in order to superimpose maps.

The other task is to decide how we publish linguistic maps. Nowadays, maps and the data should be open not only to academics but also to the public, as well as for educational purposes (see Kretzschmar, below). We need to consider the following issues, among others: publication media (books, CDs, DVDs, Internet), publication of original datasets as well as linguistic maps, linking with multimedia information, the need for interactive interfaces for public (non-specialist) participation, and the handling of private

personal information that may be collected along with linguistic data.

2. Day two: Applications of Linguistic Maps

It is of course an impossible task to comment in detail on all of the highly informative presentations we have benefited from witnessing at this Symposium . All seven have been exemplary demonstrations by outstanding scholars who are leading specialists in their respective fields, and the most we can hope to do is to comment on some highlights from each of them. Below we try to synthesize some common themes and challenges which emerge from these different national perspectives, to draw some comparisons with other current work in my own field of Romance geolinguistics, and to offer some suggestions for future paths of common endeavour where we can hopefully continue to learn from each other's respective experiences in scholarly research.

One terminological note: in keeping with general usage in my field of experience, the terms "geolinguistics" and "linguistic geography" are used as equivalents in this report.

2.1. Onishi's second presentation ("Analyzing Dialectal Distributions of Japanese") begins by reminding us of the roots of research in linguistic geography: the goals of establishing dialect areas, isoglosses and radiation paths of linguistic change. Geolinguists have been concerned with these important issues since the times of Wenker and Gilliéron, and although we may struggle with these basic concepts and problematise some of the assumptions which underlie them, they still reflect some of the major preoccupations which concern our discipline today.

As is well-known, the structure of linguistic variation across geography is not only

determined by linear distance but also by natural topographical features as well as human transportation and communication networks. In order to better understand the complex and subtle interplay of physical geography, general human geography and the specifics of linguistic geography, we need to be able to overlay and contrast cartographic information from these very different fields. But of course the vast bodies of data available from each of these fields can quickly become overwhelming for researchers looking for links and generalizations, especially once we add the historical dimensions of human geography across time.

For this reason it is crucial to use Geographic Information System (GIS) techniques which allow us to synthesize very large amounts of data into manageable formats which can be readily overlaid and compared. Of course, a vital part of any such undertaking is the transformation of atomistic data (e.g. displays of individual language forms) into quantitative measures of linguistic similarities and differences.

In order to maximize the empirical coverage of the relevant data, it is important that significant works such as the GAJ and the LAJ not remain isolated from each other: in this sense, the DiWA and RE.DE projects as described by Herrgen are useful examples in that they put data from different dialectological surveys together in the same GIS context. With the impressive number of 400 linguistic atlas surveys published in Japan, a similar coordinated undertaking in this country could include huge amounts of data from different regions and different periods.

2.2. Lee's detailed presentation ("Creating Dialect Maps using Map Maker") of one tool for linguistic cartography also raises the broader issues of sharing data and techniques, which are vitally important for international cooperation among scholars, and

also to ensure the comparability of results and analyses.

Happily, the era of "proprietary" dialectology – where data, once collected, remained restricted to the private use of one scholar or one research team– seems to have passed, hopefully for good. Rather than jealously guarding data like some private treasure, scholars should instead be eager to share primary (raw) data as well as developed analyses with any colleagues and researchers who share an interest in linguistic geography. In this sense, important initiatives like the GAJ which place their data online as an openly accessible database point towards the future of dialectology.

One crucial step here is the transition from the display of individual language forms to the interpretation of systematic differences in linguistic systems, be they lexical, phonetic-phonological, morphological or syntactic. Such systematic comparisons can best be achieved by means of statistical analyses using quantitative measures derived from databases of linguistic forms. So, while it seems likely that various scholars and research teams will continue to use diverse data analysis and presentation tools which are appropriate to their particular scholarly projects, it should nonetheless be possible to develop common standards for data storage and access, which could in turn allow us to collectively take a large step towards international compara of geolinguistic results.

2.3. Iwata's presentation on the interpretation of linguistic maps again shows us the significance of physical topography (e.g. major rivers) in the establishment of dialect areas and the variation which leads to change in words. Here we are again reminded of such classic geolinguistic concepts as synonymic and homonymic collision, which as illustrated here are clearly related to the concepts of "fudged lects" and "mixed lects" as proposed by Chambers and Trudgill. The illustrations presented here extend these key

concepts from the area of phonetics (where they were originally proposed) to the area of lexical blends; it is to be hoped that they can also be extended further to the area of morpho-syntax. While the notions of analogical attraction and grammaticalisation clearly have great relevance for geolinguistic data, we have not yet established the equivalents of a grammatically "blended" or "fudged" form in the context of Chambers & Trudgill's typology of transitional varieties.

2.4. While Hans Goebl's plenary lecture on dialectometry is illustrated with examples from the mother of all Romance geolinguistics, Gilliéron and Edmont's *Atlas linguistique de la France*, we must remember that the techniques developed by his Salzburg school have been successfully applied to many different linguistic atlas datasets, including the *Linguistic Atlas of Italy and Southern Switzerland* (AIS) and of course Goebl's own magnum opus (which he modestly 'forgets' to mention here), the *Linguistic Atlas of Dolomitic Ladinian and neighbouring Dialects*, a 'speaking' Linguistic Atlas (with sound-files available online and on DVD).

Goebl succinctly restates the overall goal of our research paradigm as "the increased understanding of the dialectal management of space by humans." In order to achieve this ambitious but attainable objective, certain theoretical and technical bases must be established in order to permit concrete applications which can yield comparable results in an empirically reproducible manner. Goebl's Visual Dialectometry (VDM) approach has shown that these common requirements are worth the effort in that they have produced measures of dialect differences and similarities which can be compared between different national atlases. It is to be hoped that more geolinguists will take up the generous offer of the Salzburg school to apply these methods to different linguistic atlas datasets.

2.5. In the case of Herrgen's presentation ("From dialect to variation space"), the Regionalsprache.de (Re.de) project coordinates and synthesizes German data from different sources in a manner which suggests a possible approach for other countries which have multiple linguistic atlas and dialect surveys (such as the LAJ and GAJ). In the area of Romance geolinguistics, a comparable project would be the Thésoc (Online Occitan database http://thesaurus.unice.fr/) which is in the process of compiling data from across different Occitan-speaking regions, from both published and unpublished linguistic atlases, as well as other sources which include texts in the different varieties. In the case of the Re.de project, new data on the regional varieties of spoken German are being added from a novel source (emergency response recordings), which should hopefully lead to an important database of spontaneous connected speech. Such an initiative is very important in that it can provide new connected-speech data for the study of variation in morpho-syntax, one of the areas where geolinguistic comparison has traditionally been weakest.

2.6. Ramisch's illustration of H-dropping in British English very nicely demonstrates the interplay of sociolinguistic and geolinguistic factors, and again produces a picture which reflects Chambers & Trudgill's typology of transitional varieties, with "fudged lects" and "mixed lects" (ach as /j-/ or /w-/ points us in the direction of general phonological theory: as is often the case, dialect data illustrate an underlying relationship (e.g. between /h/ and glides) which would never be apparent if we only considered the data from Standard English. This point is reflected as well in Herrgen's work on t-deletion and Optimality Theory, in the case pointed out yesterday of

palatalisation in Chinese dialects, as well as in one of these authors' research on variable pronoun paradigms in Spanish dialects (Heap, David. 2002a. Split Subject Pronoun Paradigms: Feature Geometry and Underspecification. *Current Issues in Linguistic Theory*, 124-139; 2002b Morphological Complexity and Spanish Object Clitic Variation. *Romance Phonology and Variation*, 55-68): time and again non-standard varieties provide crucial data for linguistic theory which just never show up in standard languages.

2.6. The "New Technical Procedures" presented by Perea illustrate how computer technologies are crucial to each step of geolinguistic research today, from data compilation, to data processing and analysis, to the presentation of results. Two of the applications she demonstrates (Dialect stratography and voice synthesis of verbal forms) illustrate beautifully how the results of geolinguistic research can be made available to the broader public in novel and accessible ways. In terms of research results intended for more specialised audiences, her adaptation of old and new datasets (from Alcover's field notebooks and the COD) to Goebl's Visual Dialectometry is a perfect illustration of the inter-compatibility of methods for which geolinguistics should strive in the 21st century.

2.7. I would like to also mention the work of Kretzschmar, who in addition to editing the online Linguistic Atlas of the Middle and South Atlantic States (LAMSAS) is also responsible for the site which presents the results which are available electronically from the different U.S. linguistic atlases (the Linguistic Atlas Project website http://us.english.uga.edu/). In his 1999 article on "The Future of Dialectology", (*Proceedings of the Harold Orton centennial conference*. Leeds Studies in English 30, ed. by K. Wells & C. Upton, 271–87) Kretzschmar stresses not only the importance of using

relational databases and statistical tools in developing geolinguistic research:

> The key feature of the Web site is that it is an interactive resource. It is abundantly cross-linked in addition to allowing the user to ask several different kinds of questions of the database. When we have more data, it will be possible to ask questions across several different projects at once.... The Web is the research tool of the future, and we have it now (Kretzschmar 1999: 283)

But he also underlines the importance of the Internet for the dissemination of geolinguistic results in an accessible format so that we can reach not only other specialised scholars but also the interested general public:

> We need to accept as central to our purpose 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. (283)

2.8. Before concluding, I would like to briefly mention a few recent works in my own field of Romance linguistics which are particularly important for the themes dealt with at this symposium. *Lectures de l'Atlas linguistique de la France de Gilliéron et Edmont : Du temps dans l'espace* by Yves Le Berre, Jean Le Dû and Guylaine Brun-Trigaud (CTHS :2005) is a landmark study which uses computer methods to study the *ALF* data from both a geolinguistic and a diachronic point of view.

L'Atlas linguistique audiovisuel du Valais romand (ALAVAL) is a recent linguistic atlas of a "microdomain" (just 80 x 80 km! but with many distinct language varieties)

which provides two important innovations: 1) the maps and transcriptions are accompanied by video recordings of subjects speaking, so we can actually see the last generation of Francoprovençal speakers in this region using their language in their own homes, and 2) the data are all in the form of full sentences rather than isolated word forms, so we can study morphosyntactic phenomena in connected speech.

The *Atlas Multimédia Prosodique de l'Espace Roman* (or AMPER:) is a coordinated project which spans several different national domains and uses instrumental acoustic analyses to elicit data on intonation, another area too often neglected in geolinguistic studies.

In addition to providing a range of linguistic maps online, the Institute of the Galician Language (ILG**)** also allows users to access an online application which shows the cartography of family surnames in Galicia, based on census data.

The *Corpus Oral y Sonoro del Español Rural* (or COSER) is a contemporary corpus of interview texts from older rural Spanish speakers in different regions of Spain, especially useful for studying morphosyntactic variables.

The *Atlas Lingüístico de la Península Ibérica* (Linguistic atlas of the Iberian Peninsula or *ALPI)* is an older survey (mostly from the 1930s) which covers mainland Portugal and Spain as well as Andorra, the Balearic islands and Rousillon in France. To date, only one volume has been published in print (ALPI 1962) but the raw data are now distributed in facsimile format (see), and we hope they will also soon be made accessible as a retranscribed online database with GIS capability.

2.9. In conclusion, we must remember that a linguistic atlas (such as the GAJ and many others which colleagues at this symposium know intimately) is a great achievement

but should not be seen as an end-point in research but rather a beginning of many fruitful years of increasingly sophisticated analyses. As we look towards the future of linguistic geography we should keep in mind the following points from today's presentations:

- combining old data sources and new data sources to create contrastive analyses;
- ongoing searches for new and innovative methods of collecting language variation data;
- uses of new technologies, especially the Internet, to develop and disseminate research;
- sharing of datasets and general convergence towards compatible formats for data storage and access
- importance of geolinguistic results for formal theoretical linguistics, for related social disciplines, and (properly presented) for the general public.

If we keep in mind points like these and remain in contact with geolinguistic researchers in other countries and continents as new techniques and perspectives develop, then there is indeed a bright future for our discipline in the 21st century and beyond.