

## Workshop Proposal: Geolinguistics

"Geolinguistics" is a relatively new term to describe our traditional interest in the interrelationship between language and area. This has from time to time been called by different names, such as linguistic geography, dialect geography, and word geography, which are not mere synonyms. "Linguistic geography" is the most general designation, since the word "linguistic" attributes only some general quality of language to geography; it indicates that sort of areal study which is characterized by linguistic differences. Two more phrases are associated with dialectology. "Dialect geography" indicates the sort of areal study which locates specific dialects. "Word geography" gets down to particulars; it indicates study of the areal distribution of words. Some qualitative relationship between language and the land is posited in each one, but for each name a different relationship. "Geolinguistics" does not offer any additional clarity beyond the existing set of terms; it seems valuable chiefly to indicate a contemporary interest in linguistics as combined not only with maps but with cultural geography. However, the new term "geolinguistics" does suggest that we have new methods and tools available to us for consideration of language and area. This workshop will offer state-of-the-art discussion of new possibilities in three different directions. AAAAA will discuss measurement of the diffusion of linguistic change via normal contact among speakers (as opposed to conquest or settlement). BBBBB will consider what Dennis Preston has called perceptual dialectology, the perceptions that people have about variation in their language, as these perceptions are distinct from actual production differences. Finally, CCCCC will discuss language and region, as "region" suggests the association of cultural differences with different areas. Each presentation will offer clear definitions of the central terms and concepts of its approach, and will demonstrate techniques that sociolinguists can use in their own work.

### Diffusion and Aggregated Variation Data

We aim to illustrate one of the many scientific questions which confront geolinguistics, contrasting the degree to which a very bare concept of geography, essentially just distance, competes with a more complicated one, involving areas or regions, in explaining the aggregate linguistic distances among language varieties. The data is taken from various data atlas collections in six language areas and the distance is measured using a variant of Levenshtein distance, which has been demonstrated to be valid with respect to dialect speakers' judgments of linguistic difference. Pure distance models account for between 14% and 38% of the variation found in the data in straightforward regression designs. Wishing to use dialect areas independently established, the study takes Wrede's famous map of German dialect areas, adding dialect-divisions as categorical independent variables to assess their explanatory value. This increases the explained variation from 32% to 45%, indicating that geography is indeed structured more complexly than simple distances. On the one hand this study illustrates an *issue* in geolinguistics, while on the other hand it exemplifies a concrete *opportunity*. The opportunity arises from the apparent fact that the distribution of linguistic variation invariably follows a sub-linear distribution with respect to geographic distance, a

circumstance we call “Séguy’s law”, recalling the 1971 publication in which Séguy first demonstrated the sub-linear distribution of linguistic distance as a function of geography. The opportunities arise as one attempts to build further on Séguy’s empirical finding to conjecture about e.g. the determinants of the slope of the sub-linear (logarithmic) curve describing the distribution of linguistic variation as a function of geographic distance, or, as one attempts to theorize about the fundamental mechanisms of linguistic diffusion, and how these impact on the curve of aggregate differences.

## **Perceptions of Speech Differences**

The aim of “geolinguistics” is not only to search for and describe the relationship between speech production and geographic locations, but can also be a way to talk about perceptions of speech seen through the perspective of maps. Dennis Preston developed the ‘Draw-a-map’ methodology, giving the contemporary perceptual dialectology a foundation in beliefs and ideas about speech of folk speakers. Such an approach shifts focus from what we as linguists think about speech and how we describe it, to how lay speakers’ perceptions are shaped by constraints such as knowledge, distance, or experience with a particular kind of speech (or lack of it). The results of the Prestonian ‘Draw-a-map’ methodology give an impression of neat boundaries between regions of salient perceptual agreement. Those results originated from generalizations in the form of perceptual isoglosses. This presentation will propose an alternative way of computing results from ‘Draw-a-map’ methodology using data from Poland. The analysis of perceptual maps was performed using automated processes; however, the solutions presented do not require a high level of proficiency with technology. You will be shown a step-by-step procedure necessary to perform such an analysis. Instead of creating generalizations, maps drawn by respondents were combined into a “result map” depicting all areas of perceived speech varieties, with various levels of agreement among the respondents. The results suggest that our perceptions behave as a complex system, just as asserted in linguistics of speech approach (Kretzschmar 2009). Moreover, what creates a contrast to Prestonian analysis is that the majority of areas on the map display a low level of agreement about the locations of perceived speech varieties. This method allows us to show the variation in perceptions (both similarities and differences visible at the same time), as opposed to perceived saliency of high level of agreement created by perceptual isoglosses.

## **Language and Region**

The problem for sociolinguists about language and region is not the language but what to do with regions. Nothing could be more obvious, more natural to talk about than the existence of regions which subdivide a country (the United States in this essay, but really any country), and to which we commonly attribute different cultural characteristics including language. And yet our normal reference to regions gets us in trouble. We can be puzzled when somebody puts Kansas in “the Midwest” when we always thought that it was part of “the Great Plains,” or when somebody denies that Virginia is really part of

“the South.” As for language, we can be puzzled when somebody from “the South” does not seem to have the accent to go along with their homeplace, or when we realize that people from Michigan don’t all have the “Standard” accent that Michiganders believe themselves to have (or don’t all have an accent with the Northern Cities Shift, which is what American sociolinguists say they have, following Labov). Kansas, Virginia, and Michigan are, of course, American states with formal and precise boundaries.

“Midwest,” “Great Plains,” and “South” are common informal designations for American regions, and different people put different states into each them. We do all believe in regions and regional differences, but it is one thing to accept them as a part of our cultural environment and quite another to try to make regions work as a part of our sociolinguistic inquiries. This presentation will examine what is meant by "region" as a concept developed in cultural geography and discuss it with respect to linguistic systems, to the end of how we can use regions effectively in sociolinguistics. The relevance of complex systems to regional differences will be a particular target for discussion, especially the importance of scale-free networks.