Fuzzy dialect areas and prototype theory - Discovering latent structures in geolinguistic variation using factor analysis

SIMON PICKL
University of Salzburg

In their 2001 article, Heeringa and Nerbonne argue that "that both the area view and the continuum view are useful for gaining insight in the nature of the dialect landscape, which may be described as a continuum with varying slope or, alternatively, as a continuum with unsharp borders between dialect areas" (2001, 399). However, these are not only different views but also competing models of the structure of geolinguistic variation, a contrast which is reflected in dialectometric methodology. To date neither a unified theoretical model nor a unified methodology exists.

In this paper an attempt will be made to furnish a theoretical basis from which both concepts can be derived, and a corresponding method for the analysis of empirical data is introduced, along with first concrete results. Gaetano Berruto's concept of "condensation areas" in language space serves as a starting point; when applied to linguistic geography it entails that dialect areas are fuzzy and can overlap one another. The model is completed with a prototype-theoretical dimension, allowing condensation areas to be understood as abstract dialect types that are manifested to varying degrees in concrete individual dialects.

To operationalize this concept, factor analysis, a statistical technique for dimensionality reduction is employed; its mode of operation is eminently suited to identifying and quantifying fuzzy condensation areas, i.e. meaningful geolinguistic structures. At the same time, co-occurrences among linguistic variants help to identify systematic relations. Applying the method to lexical data from southern Germany reveals the potential of the method, not just for the modeling of linguistic space, but also in the search for previously unknown spatial structures.