Advanced mapping in dialectometry

Max Sondag
TU Eindhoven
Thom Castermans
TU Eindhoven
Bettina Speckmann
TU Eindhoven
Wieling, Martijn
University of Groningen

In dialectometry, (aggregate) linguistic variation is generally visualized on a geographical map. For example, linguistic differences between individual locations may be summarized in a multidimensional scaling (MDS) map by coloring geographical areas. Figure 1 (left) visualizes an MDS map of English lexical variation (collected per postcode *area*). Figure 1 (right) shows an MDS visualization of Dutch pronunciation variation data collected at individual *points* (i.e. locations). Here, Voronoi tesselation was used to generate a subdivision into artificial areas.
While these maps are insightful, they are potentially misleading. The human eye visually integrates over areas and perceives larger areas as more important. Furthermore, the current maps are limited to showing only the geographical distribution of a single variable. If one is interested in assessing the influence of an additional (social) predictor – such as population size – on geographical dialect variation, one needs to turn to quantitative techniques (see Wieling and Nerbonne, 2015 for an overview).

Here we propose two novel visualization techniques to dialectology which are able to create areas of specified sizes and allow for the visualization of an additional quantitative variable Q. The first technique, a mosaic cartogram (Cano et al., 2015) is applicable to area data and distorts the original geographical shape so that the size of each area now reflects the value of Q (see Figure 2a). The second technique, suitable for point data, is a Voronoi treemap (Balzer et al., 2005) which creates artificial areas that reflect the value of Q.

In the presentation, we will discuss the details of these techniques, and give a demonstration of their effect using the (currently under development) built-in functionality of Gabmap (www.gabmap.nl).

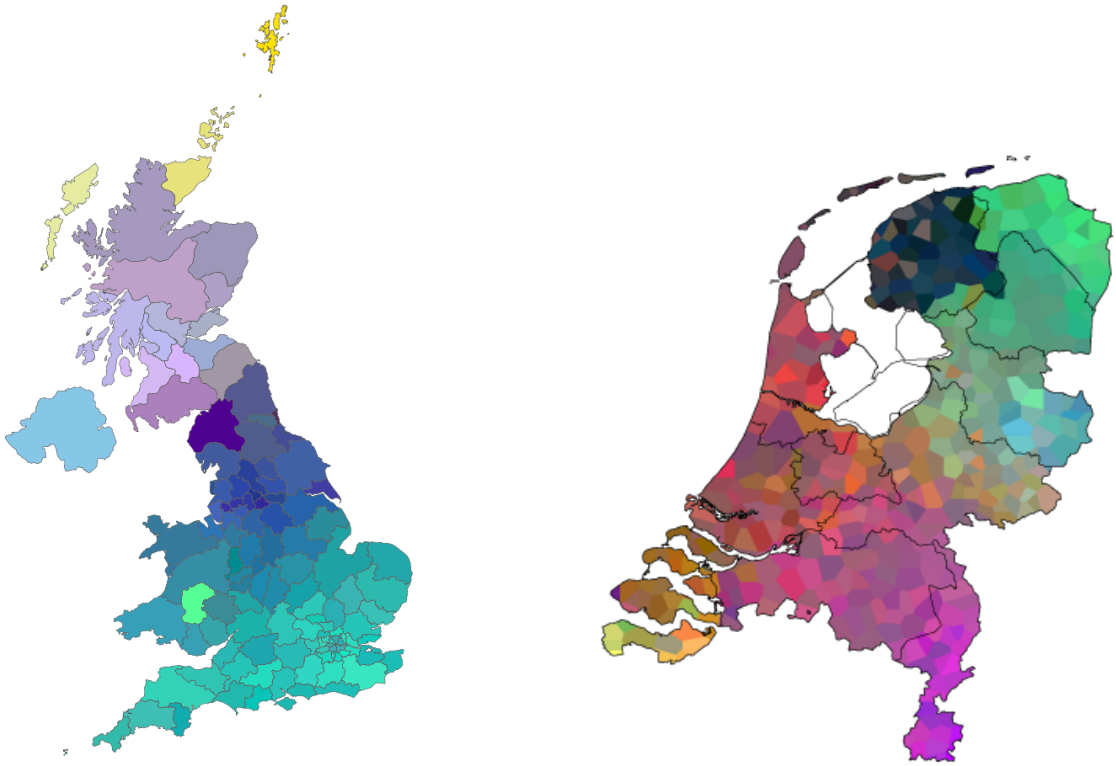

Figure 1. MDS visualizations of English lexical (left) and Dutch pronunciation variation (right)

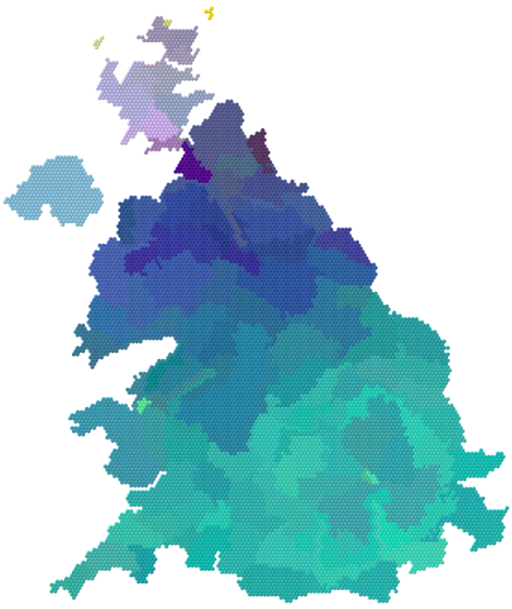

Figure 2. Mosaic cartogram (left) and Voronoi treemap (right) incorporating *population size*