Visualising Typological Relations: Using Heat Maps with WALS

Richard Littauer
University of Saarland
Computational Linguistics Department
66041 Saarbrücken, Germany
richard.littauer@gmail.com

Alexis Palmer
University of Saarland
Computational Linguistics Department
66041 Saarbrücken, Germany
apalmer@coli.uni-sb.de

Rory Turnbull
Ohio State University
Department of Linguistics
Columbus, Ohio
turnbull@ling.osu.edu

Abstract

This paper presents a novel way of visualising linguistic typological data. Computational methods have only recently been applied in the formation and use of large typological databases. Many studies have since focused on discovering relations between languages using typology, often using sophisticated statistical techniques. However, few papers have provided new ways of visually presenting the resulting data. Here, we show that one can use the data from the World Atlas of Language Structures(Dryer and Haspelmath, 2011) to develop both heat maps and admixture graphs that visually show the interconnected relationships between languages and language families. We hope that the images will bring a new perspective to the data, resulting in interesting findings and illuminating areas of research.

1 Introduction

1.1 Heat maps

- brief history of typology
- WALS
- visualisation
- what visualisation can do for us
- Heat Maps and Admixture- where they come from, what they are doing here
- plan for the paper

2 Material and Methods

2.1 Linguistic material

- Information about WALS data
  - Description of typological data
  - Longitude and latitude, and how it is measured
  - language families
  - WALS-language Sparseness
- Information about Multitree(for Language Information and List), 2009)
- Information about Ethnologue(Lewis, 2009) scrape

2.2 Methods

- Cleaning WALS data
- Collapsing WALS data
- Scraping Ethnologue, Multitree, formati
- Measuring Phylogenetic distance
- Measuring Geographical distance
- Combined map
- Compiling python scripts, converting into R
- Sorting R output

3 Results

3.1 Heat maps

- Phylogenetic distance heatmaps
- Geographical distance heatmaps
- Combined maps

3.2 Heat maps

- Admixture mappings

4 Discussion

- What these tell us (map by map)
- What these tell us, overall - implications
- Warnings: sparse data, data not there, etc.
- Future work

References

Matthew S. Dryer and Martin Haspelmath, editors. 2011. *The World Atlas of Language Structures Online*. Max Planck Digital Library, Munich, 2011 edition.

Institute for Language Information and Technology (LINGUIST List), editors. 2009. *Multitree: A digital library of language relationships*. Eastern Michigan University, Ypsilanti, MI, 2009 edition.

M. Paul Lewis, editor. 2009. *Ethnologue: Languages of the World*. SIL International, Dallas, TX, sixteenth edition.