

A Colorful First Glance at Data on Regional Variation

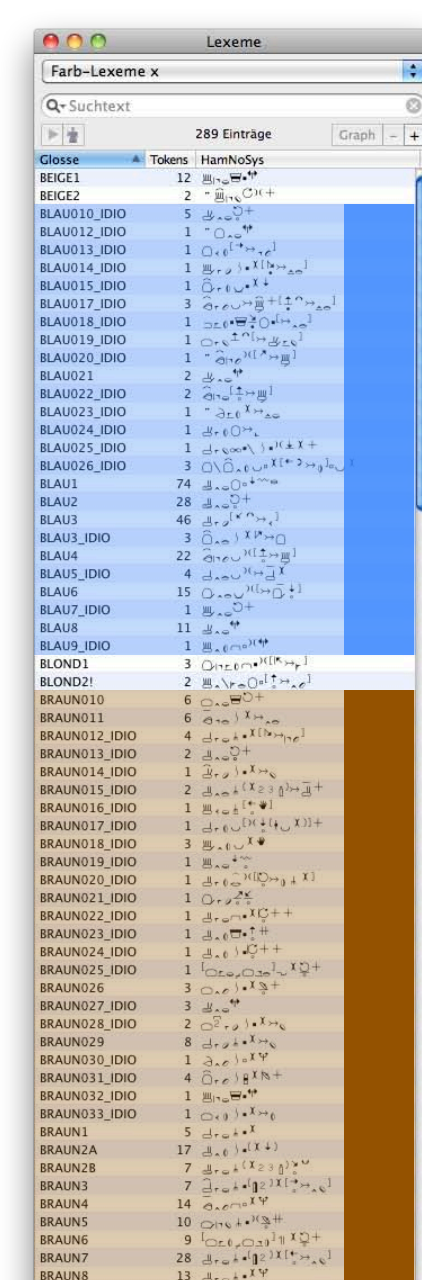
Extracted from the DGS-Corpus

– With a Focus on Procedures –

Gabriele Langer – University of Hamburg, Institute for German Sign Language and Communication of the Deaf

The DGS Corpus

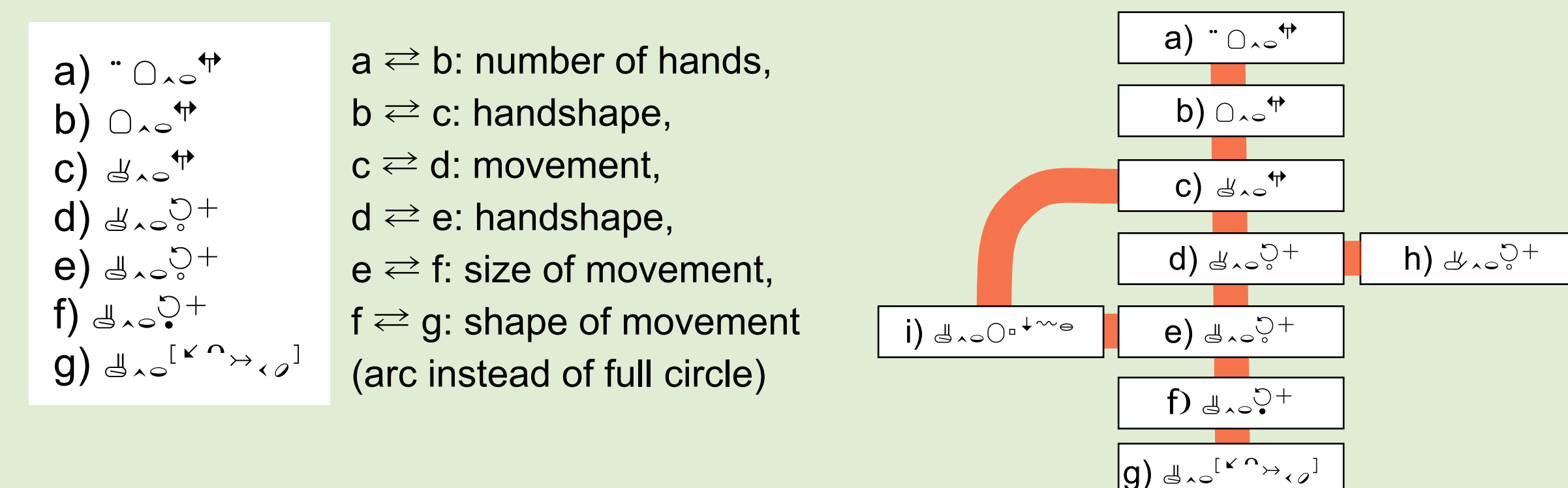
- Data
- 330 informants from 13 regions, (148 of 434 counties)
 - metadata on informants (including place of residence, place of growing up and place of schooling)
- Content
- 20 elicitation tasks, including elicitation of isolated signs
- Size
- estimated 540 hours of signed material
 - estimated 2.5 million tokens



Phonological and Lexical Variation: Same or Different Sign?

General practice:

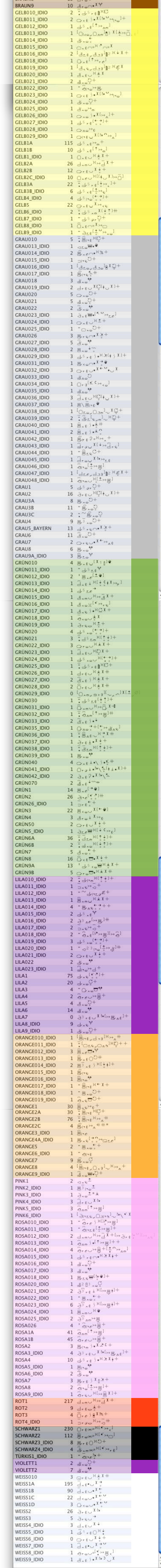
- phonological variation: phonologically related forms – similar sign forms that differ slightly (forms treated as subvariants of one sign)
 - lexical variation: phonologically unrelated forms – (forms treated as distinct signs)
- Problem: distinction is not always clear-cut, especially considering chains of similar forms with very diverging forms as opposing ends. Chains also may branch and reconnect.



chains of partly similar forms used for *blue*

Preliminary Study on Color Signs

- Data
- 156 informants (from 12 regions, 90 counties)
- Content
- task: elicitation of isolated signs for colors
- Size
- 2052 tokens
- Access
- spot transcriptions: segmentation, lemmatization
 - working environment: iLex
 - data extracted via SQL-queries
- Purpose
- testing the procedure of displaying regional distribution of signs
 - generation of distributional maps
 - first insights on patterns of regional distribution of sign variants in Germany
 - investigation of variation of color signs in DGS

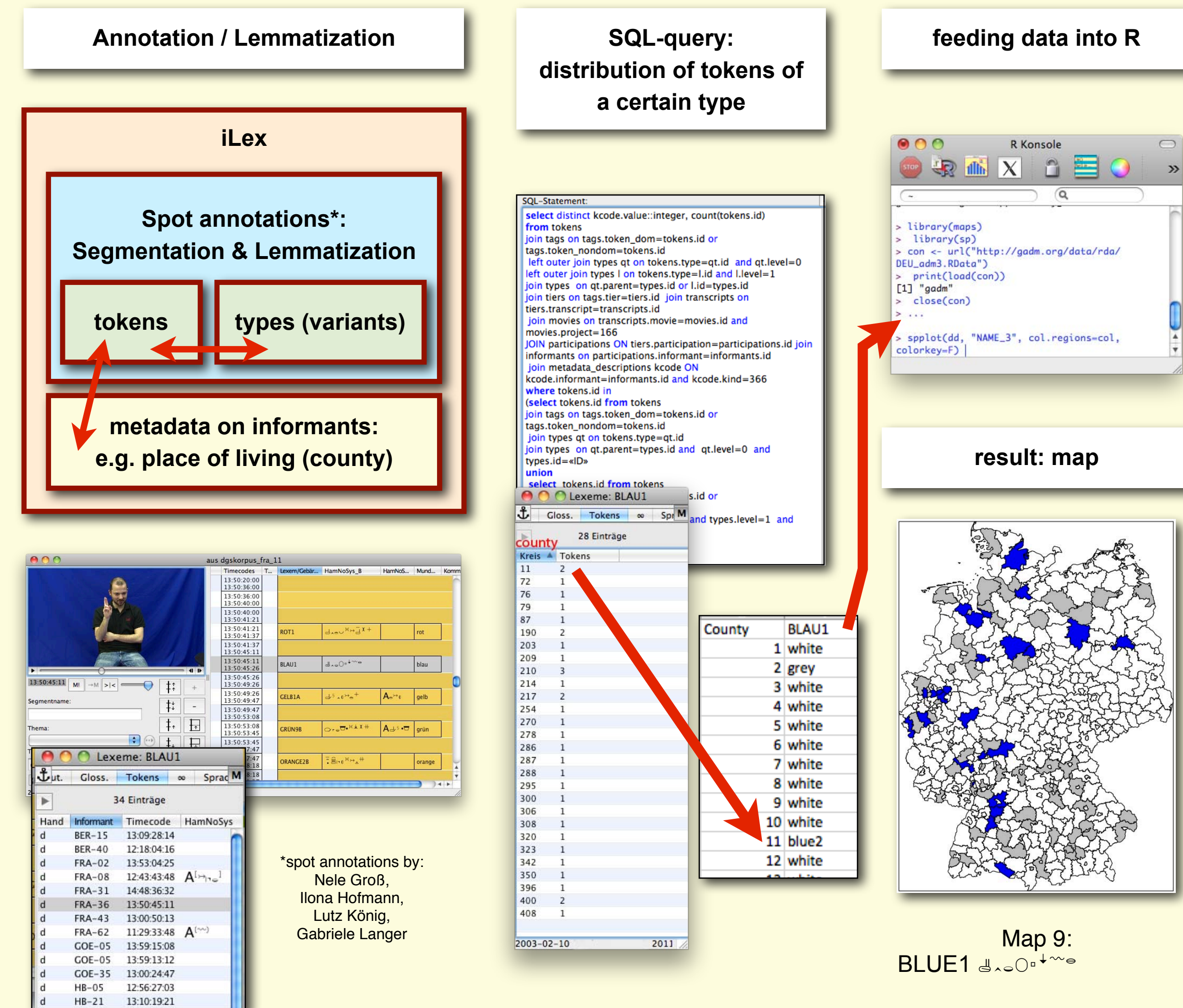


Results

- First pass of annotation:
- 256 types (forms) of signs used for color
 - 117 types (forms) with only one token
 - 45 types with 9 or more tokens (accounting for 75 % of all tokens)
 - some signs such as RED1 are used in all regions analyzed
 - other signs e.g. BLUE3, GREEN2, GREEN3, GREEN9A, BROWN4 show clear regionality of use
 - no single set of color signs for DGS
 - overlap of regional variants to various degrees
 - some evidence for dialectal regions
- These results indicate tendencies but are work in progress!
- All annotations still have to undergo the lemma revision process - this will modify the results as some forms (types) may be recategorized as deviations of other types – thus reducing the number of types presented here – while others such as BLACK1 may be split up into two forms when analyzing deviation of token forms and their regional distribution.

Generating Distributional Maps

could be produced directly from the data in the database (e.g. iLex) on demand



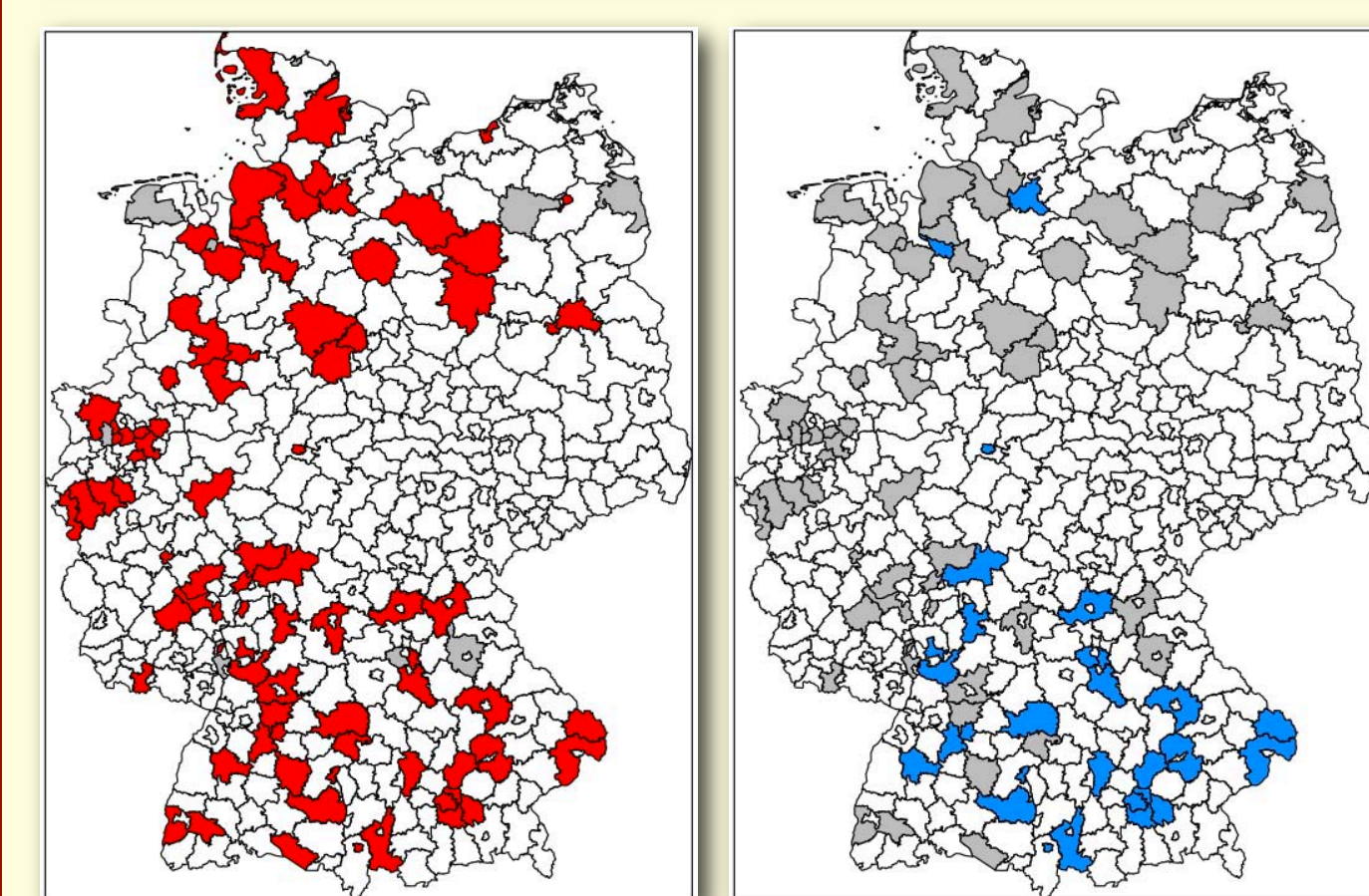
Advantages and Uses of Distributional Maps

- visualization of regional distribution of signs, variants or other phenomena: distributional patterns can be recognized at a glance
- may support the lemma revision process:
 - distribution patterns may be taken into account for decisions on whether two variants belong to the same sign (because they show comparable distribution) or whether they belong to different regional variants
 - tokens with forms that are in between two of similar forms (of competing types) may be assigned with more certainty to the corresponding type when taking regionality of informant and sign distributions into account
- may serve as basis for analysis when writing a dictionary entry (regional use of sign)
- maps may be included in a dictionary entry (in an adjusted version)
- cumulated distribution patterns of many signs can help to define dialectal areas

Distribution of Color Signs

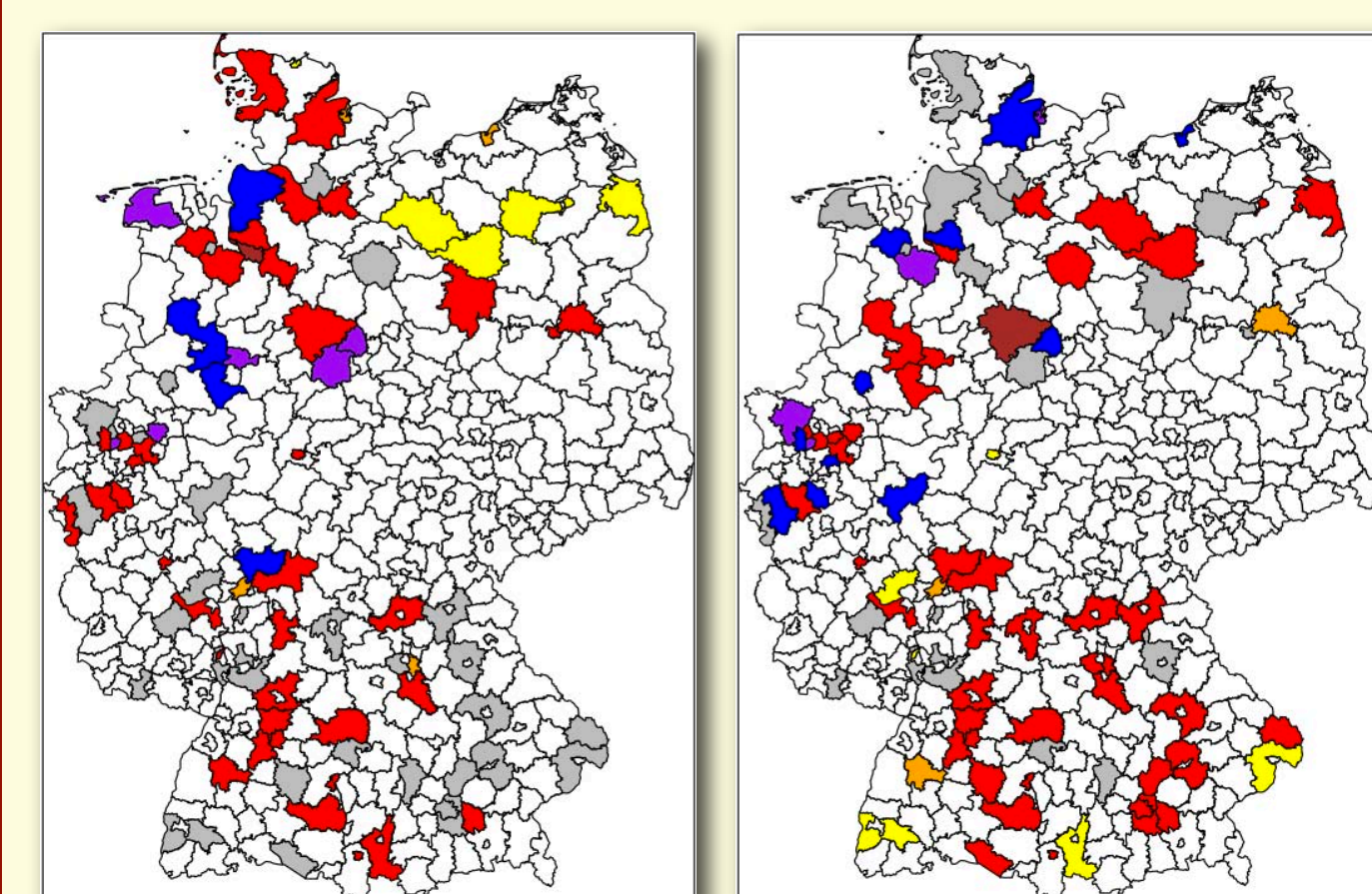
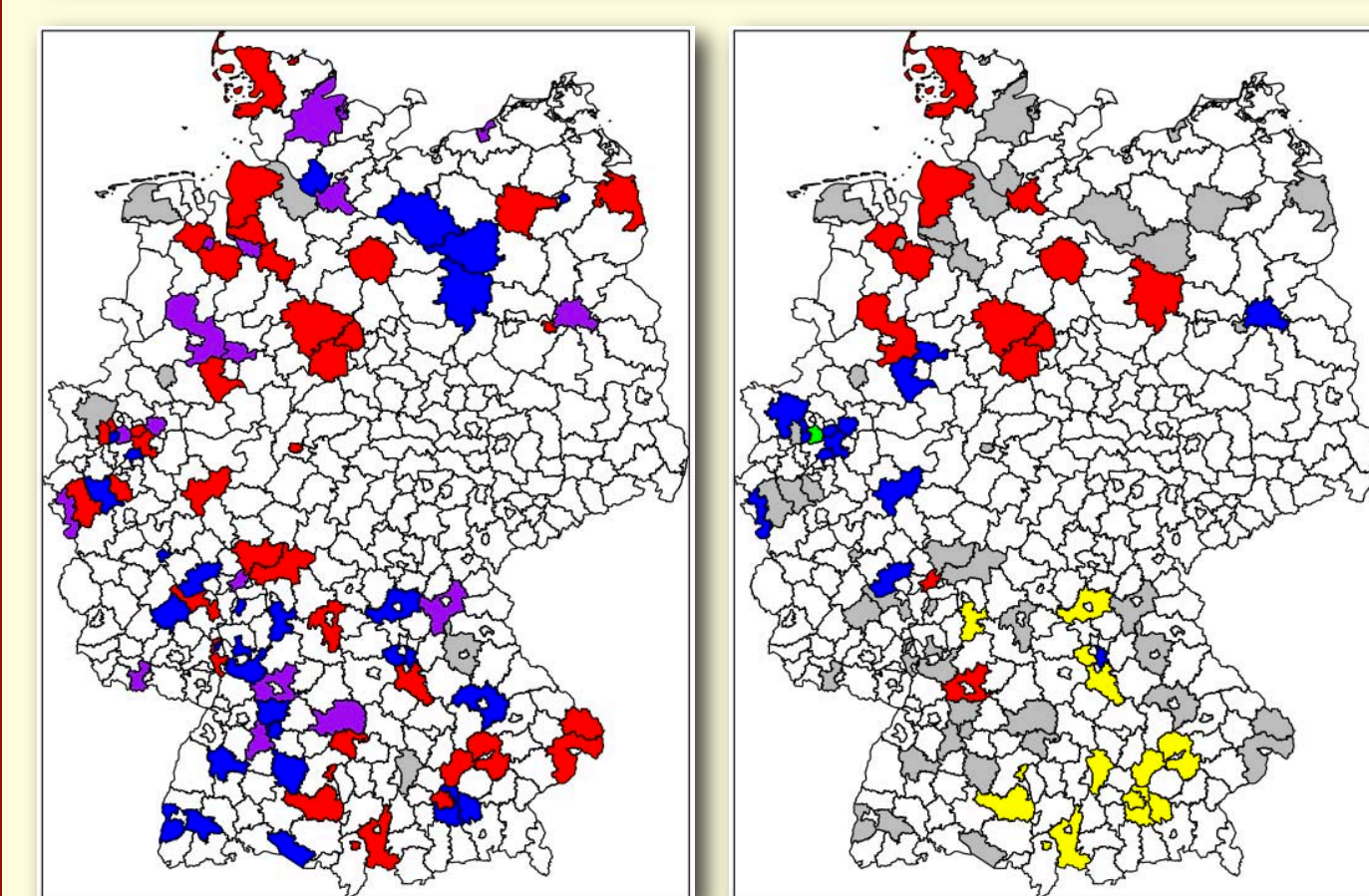
(white areas: no data, grey areas: informants, but no use of the sign(s) analysed)

display of one variant

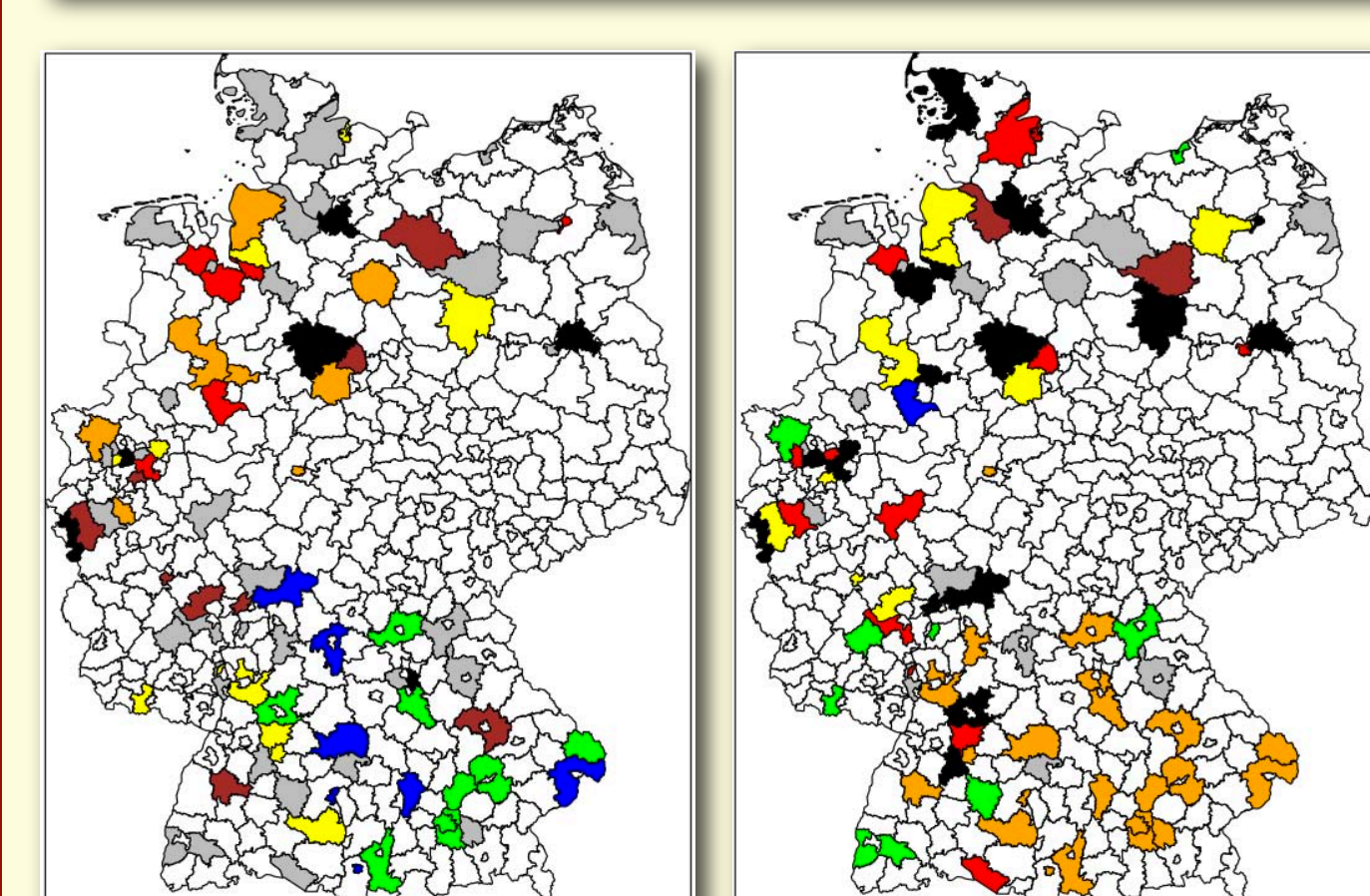


display of two or three variants

(overlapping areas of use: mixed colors)

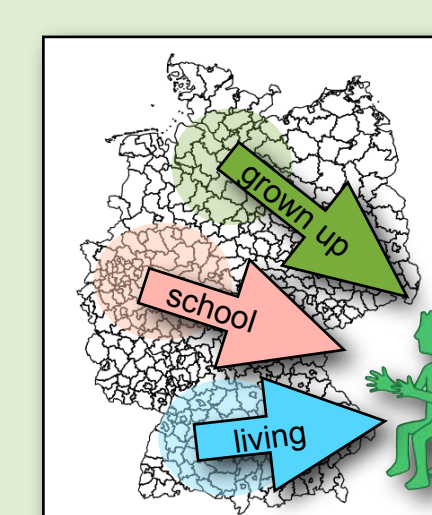


display of six variants



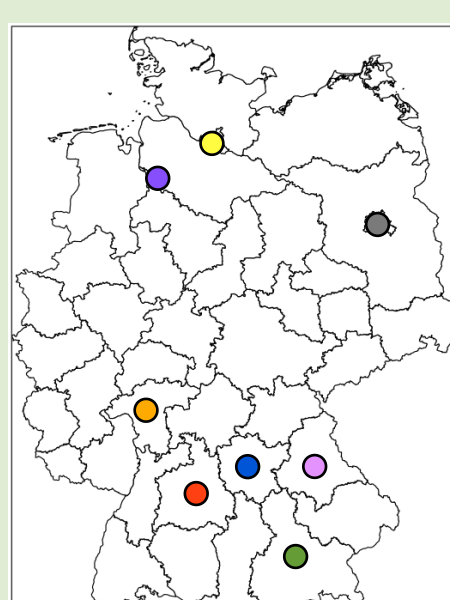
Dealing with Multiple Regional Influences on Signers

Many signers show influences of several regions in their signing.
(proposed procedure – example has been made up for demonstration purposes; areas: government districts)

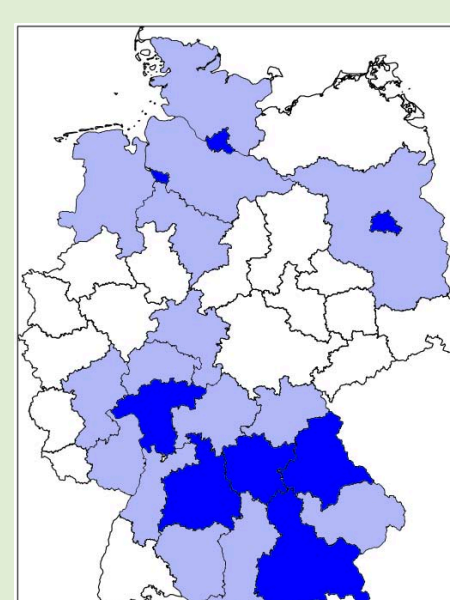


signer	several regional influences
A	
B	
C	
D	
E	
F	
G	
H	

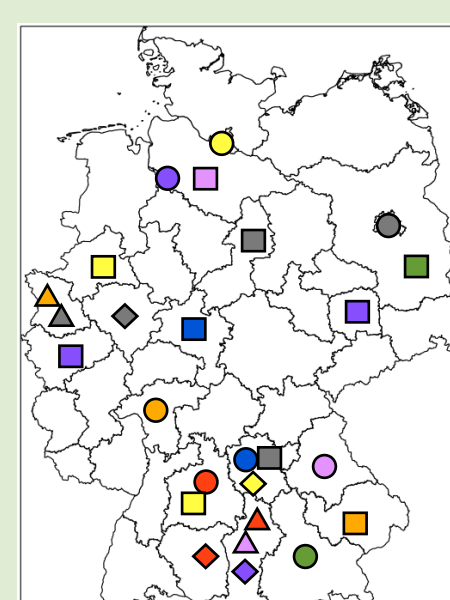
- ◇ place of growing up
- place of living
- △ place of schooling
- previous long term residence



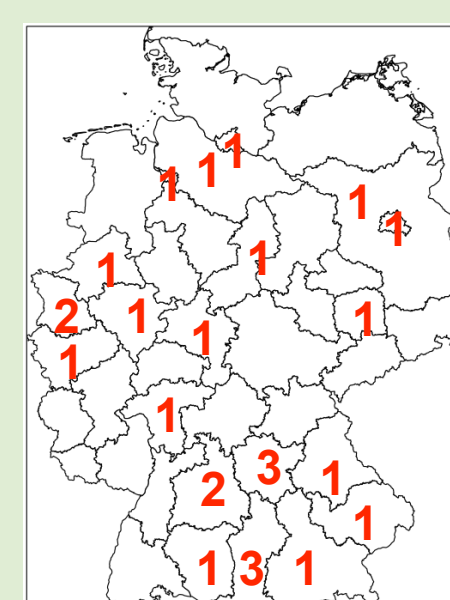
places of living of informants using a particular sign



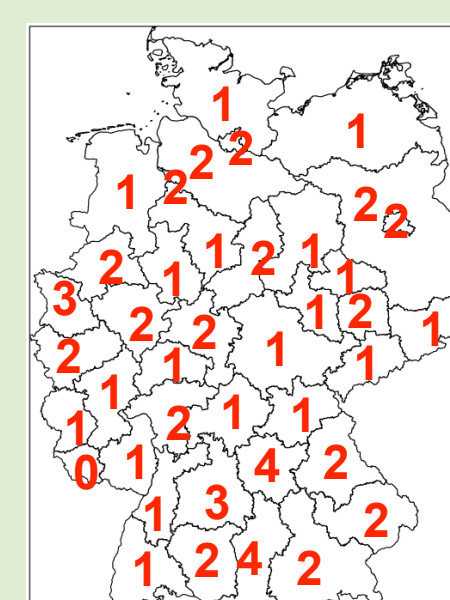
regions of attested use on basis of place of living (dark blue) and neighboring areas (light blue)



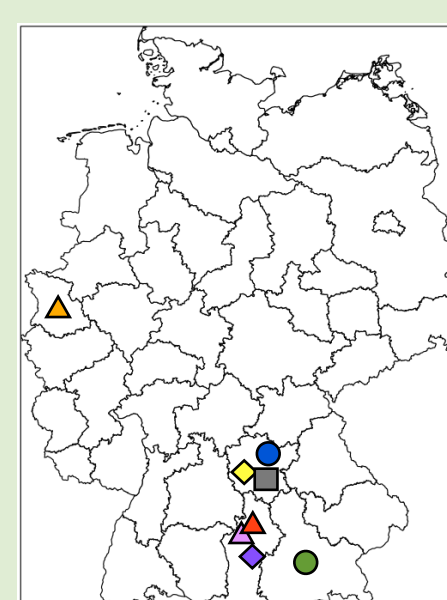
multiple geographical influences of informants using a particular sign



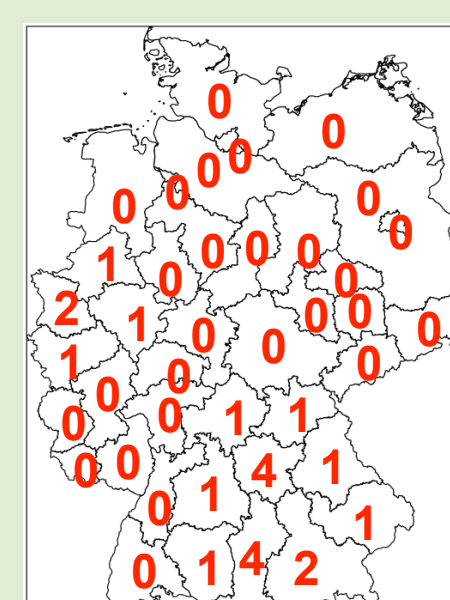
cumulated geographical influences of informants using a particular sign



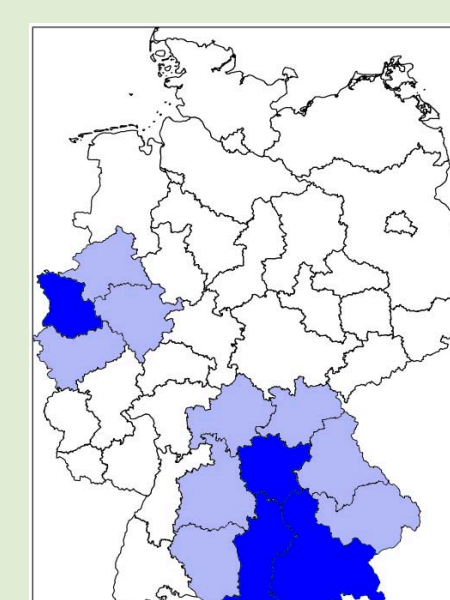
a-values: determined on basis of all known regional influences (thresholds need to be defined)



most probable geographical influence of each informant using a particular sign determined on basis of a-values and priorities* (*to be defined)



b-values: determined on basis of most probable influences



regions of attested use on basis of most probable regional influence (dark blue) and neighboring areas (light blue)