

# Neue Dialektometrie mit Methoden der stochastischen Bildanalyse

Simon Pröll (München) / Simon Pickl (Salzburg) / Aaron Spetl (Ulm) / Volker Schmidt (Ulm) / Evgeny Spodarev (Ulm) / Stephan Elspaß (Salzburg) / Werner König (Augsburg)

## 1. Gegenstand

### 1.1 Organisationsrahmen und Forschungsstand

Das Projekt *Neue Dialektometrie mit Methoden der stochastischen Bildanalyse* wurde unter Finanzierung seitens der DFG interdisziplinär vom Institut für Stochastik der Universität Ulm und dem Lehrstuhl für Deutsche Sprachwissenschaft der Universität Augsburg durchgeführt.<sup>1</sup>

Vorrangiger Zweck ist die Entwicklung neuer Methoden zur quantitativen sowie qualitativen Auswertung großer Korpora areallinguistischer Daten. Zielpunkt ist die Bereitstellung der im Projekt entwickelten und erprobten Methoden in einem kompakten und anwenderfreundlichen Softwarepaket namens *GeoLing*, das es anderen Nutzern ermöglicht, mit ihren eigenen Daten entsprechende Analysen eigenständig durchzuführen. Die Software ist komplett in Java geschrieben und nicht nur plattformübergreifend lauffähig, sondern vom Benutzer bei Bedarf auch individuell anpassbar.

Ein erster Ausgangspunkt für die Arbeiten im Projekt war die „klassische“ Dialektometrie (Séguy 1971, 1973; Goebel 1982, 1984), die bislang in der Hauptsache mittels Aggregation versucht, der Variationsvielfalt in Sprachatlanten Herr zu werden. Während diese aggregativen Vorgehensweisen (die weiterhin die Basis vieler quantitativer Zugänge zur Sprachgeographie sind, vgl. für einen Überblick etwa Heeringa 2004: 14–24; Nerbonne 2010; Nerbonne & Kretschmar 2013) zweifelsohne die „dominanten“ (d.h. „möglichst hochrangige[n]“; Goebel 1986: 43) Strukturen der Variation gut abbilden können, werden schwächere oder weniger hochrangige Aspekte der Variation ausgeblendet. Einher geht damit oftmals auch eine starre Sicht auf „Areale“, die implizit als einheitliche Dialekträume konzeptualisiert werden, die sich scheinbar klar an Grenzen voneinander scheiden, statt fließend ineinander überzugehen. Darüber hinaus blieben (und bleiben bis heute) eventuell auftretende Mehrfachbelege an einem Ort in diesen Herangehensweisen unberücksichtigt.<sup>2</sup>

Der entscheidende Ansatzpunkt, um die Gestalt solcher „schwächeren“ Variationsanteile und ihre Bedeutung für die Gesamtdialekte exakt fassen zu können und dabei den Blick weg von einer Illusion der „Reinheit“ von Dialektgebieten zu lenken, liegt in der Einnahme einer *probabilistisch* geprägten Sicht auf Variation.

---

<sup>1</sup> Die Projektleitung der mathematischen Seite oblag Volker Schmidt und Evgeny Spodarev, die der sprachwissenschaftlichen Werner König und Stephan Elspaß (Sprecher). Simon Pickl und Simon Pröll arbeiteten am Standort Augsburg, Lisa Handl, Henrik Haßfeld, Daniel Meschenmoser, Jonas Rumpf und Aaron Spetl am Standort Ulm.

<sup>2</sup> Mitverantwortlich dafür ist sicher der Umstand, dass die dort als Datenbasis genutzten Atlanten romanistischer Prägung generell weniger Mehrfachbelege aufweisen – dies ist wohl durch eine grundsätzlich andere Philosophie der Erhebung, Selektion und Kartierung begründet.

## 1.2 Korpus und Methode

Als Beispielkorpus wurde im Projektrahmen meist auf den *Sprachatlas von Bayerisch-Schwaben* (im Folgenden kurz SBS), in späteren Phasen auch auf den *Atlas zur deutschen Alltagssprache* (AdA) zurückgegriffen – prinzipiell sollen Software und Methoden aber mit beliebigen Datensätzen arbeiten können, die räumliche Sprachvariation beinhalten.

Als Basis für weitere Analysen dienen Verfahren, die aus dem Feld der räumlichen Statistik und der Bildanalyse stammen: Zum einen beinhaltet dies die Interpretation von punktuellen Daten, wie sie in Sprachatlanten und anderen areal aufgelösten Korpora enthalten sind, als Flächen. Dabei soll die Aussagekraft eines einzelnen Belegs auch von seinen Nachbarn abhängen – der Einsatz eines Intensitäts- bzw. Dichteschätzverfahrens ermöglicht dies (Abschnitt 2.1 zeigt Beispiele für Anwendungen und Ergebnisse). Zum anderen sollen Optionen geboten werden, um Variablen, die sich in ihrer Raumstruktur ähneln, aufeinander beziehen zu können. Das wird umgesetzt, indem statistische Verfahren auf raumstrukturelle Eigenschaften von einzelnen Datensätzen (d.h. Karten) angewandt werden (Beispiele folgen in Abschnitt 2.2). Schließlich ist es ein Ziel, die gemeinsamen Grundströmungen, die sich durch die Daten ziehen, exakt fassen zu können – und zwar unter Rekurs darauf, auf welche Varianten diese gemeinsamen Züge zurückzuführen sind. Zu diesem Zweck werden Faktorenanalysen eingesetzt (Details und Analysen hierzu folgen in Abschnitt 2.3).

Alle Verarbeitungsschritte, die direkt auf das areallinguistische Korpus zurückgreifen (Intensitätsschätzung und Ermittlung der Kenngrößen, Faktorenanalyse, Clustern von Karten, Visualisierung), wurden mit der dafür eigens entwickelten OpenSource-Software *GeoLing* durchgeführt, die frei zum Download steht (siehe 2.4). Für die weiterführenden statistischen Auswertungen dieser Daten wurde entsprechende Standardsoftware (hauptsächlich *Matlab*, *Excel* und *R*) genutzt.

## 1.3 Umsetzung

Die Arbeiten am Projekt begannen im Jahr 2007; die erste, grundlegende Publikation zur Intensitätsschätzung und zum grundlegenden Vorgehen folgte 2009 (Rumpf et al.; vgl. auch Pickl & Rumpf 2011), 2010 die entsprechende mathematische Dissertation Jonas Rumpfs (Rumpf 2010). Zentrale Punkte waren in der Folge die Clusterung von ähnlichen Karten mittels verschiedener Verfahren (Rumpf et al. 2010 sowie Meschenmoser & Pröll 2012a) und eine Studie zur Erkennung elliptischer Strukturen auf Karten (Meschenmoser & Pröll 2012b). Zwei sprachwissenschaftliche Dissertationen (Pickl 2013a, Pröll 2013a) widmeten sich umfassenden Anwendungen der Techniken auf die Daten des SBS und der Interpretation der Ergebnisse; die jüngsten Arbeiten galten dem Vergleich und der Validierung verschiedener Intensitätsschätzungsverfahren (Pickl et al. 2014). Im Erscheinen befindet sich ein Beitrag zur Faktorenanalyse (Pröll, Pickl & Spetzl i.E.). Eine weitere gemeinsame Publikation zur Verifizierung der statistischen Signifikanz von Sprachgrenzen befindet sich in Vorbereitung (Handl et al. i.V.). Mit der Bereitstellung der Software *GeoLing* im Sommer 2014 sind die Arbeiten am Projekt weitgehend beendet.

## 2. Ergebnisse

Während zentrale Erträge des Projekts im Erarbeiten neuer Methoden liegen, sind – bedingt durch die ebenfalls neuen Interpretationsmöglichkeiten dieser Methoden – auch eine Vielzahl an konkreten Ergebnissen vorzuweisen, die zum Teil bestehende Hypothesen empirisch untermauern, zum Teil aber auch gänzlich neue Aspekte ins Spiel bringen. Wir stellen im Folgenden einige dieser Erträge anhand von Beispielen dar; dabei gehen wir von der

Betrachtungsebene der einzelnen Varianten (2.1) über die Ebene der Variablen/Karten (2.2) hin zur Varietätenebene (2.3). Zum Abschluss (2.4) rückt das Softwareinstrument in den Fokus, das gezielt für diese Analysetypen entwickelt wurde.

## 2.1 Variantenebene

So ist z.B. neu, dass die Methoden es erlauben, Übergangsphänomene als *konstitutives* Element von räumlicher Variation anzusehen, nicht als randständiges: Auch in Untersuchungen, die bislang als Beispiel für die Miteinbeziehung von Übergängen gelten – wie etwa die Einteilungskarte zur räumlichen Gliederung der deutschen Dialekte durch WIESINGER (1983) – ist ein Übergang schließlich meist als eine Art „Mischgebiet“ angesehen, als Raum, in dem sich scheinbar „reine“ Varietäten mischen.

Das Modell, das wir alternativ vorschlagen, ist in der Theorie nicht neu, wurde aber bislang nicht hinreichend in die Empirie übertragen. Zunächst zur zugrundeliegenden Theorie auf Variantenebene: Abbildung 1 stellt den Übergang zweier sprachlicher Varianten *probabilistisch* dar (vgl. Pröll 2014).<sup>3</sup> Auf einer (imaginären) Wegstrecke durch die Orte A, B, C und D durchläuft man ein Übergangsszenario der Gebrauchsfrequenzen beider Varianten: Im Ausgangspunkt A ist ausschließlich Variante 1 in Gebrauch; mit Nähe zu Ort B treten auch Äußerungen von Variante 2 auf, so dass sich für den Messpunkt ein Verhältnis von 70 zu 30 Prozent ergibt. Im Transit zu Ort C schlägt die Dominanz von Variante 1 zu Variante 2 um, am Messpunkt lässt sich ein Verhältnis von 15 zu 85 Prozent feststellen. Erreicht man Ort D, ist lediglich noch der Gebrauch von Variante 2 dokumentiert.

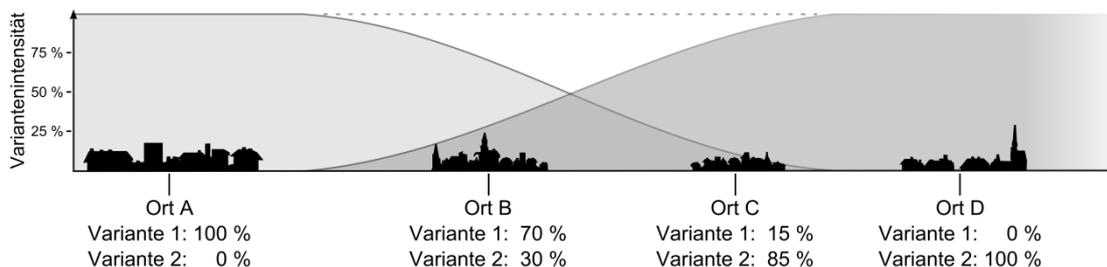


Abbildung 1: Probabilistisches Modell von Variantenübergängen.

Die Datenerhebung, die sprachlicher Empirie zu Grunde liegt, kann die kontinuierliche Natur dieses Übergangs nur gerastert wiedergeben: Pro Ort wird eine mehr oder minder gute Stichprobe aus dieser Verteilung genommen – als Störquellen, die die Qualität dieser Stichprobe mindern, kommen dabei eine Vielzahl an Einflüssen in Frage (vgl. etwa König 2010). Durch die Intensitätsschätzung wird der quantitative Übergang zwischen Varianten, der durch die Erhebungssituation quasi in groben Schritten gerastert und von statistischem Rauschen beeinflusst wurde, in Abhängigkeit der Gesamtdatenlage für jede Karte wieder einzeln rekonstruiert.

Grundmotivation für die Intensitätsschätzung der Daten ist also eine kritische Hinterfragung der Repräsentativität des einzelnen Belegs: Nachdem dieser zwangsläufig nur eine Stichprobe aus der Gesamtheit sein kann, ist seine Aussagekraft – isoliert betrachtet – nicht besonders hoch. Man bezieht daher sein Umfeld ein, um seinen Wert bestimmen zu können: Sind z.B. alle benachbarten Belege identisch, ist er als belastbarer einzustufen, als wenn alle umliegenden Orte völlig andere Belege aufweisen. Es liegt also nahe, solche Belege als „Ausreißer“ zu werten und durch die umliegenden Belege zu dämpfen. Gleichzeitig wäre es jedoch unangebracht, eventuell interessante „Ausreißer“ schlicht zu tilgen; daher wirken

<sup>3</sup> Dieses Modell ist selbstredend sozial unterdeterminiert – wir gehen für dieses Modell der Einfachheit halber von der Illusion aus, dass es möglich sei, rein geographische Variation isoliert betrachten zu können.

diese vereinzelt Datenpunkte auch auf die Masse zurück, indem sie deren Dominanz schwächen. Als Resultat der Intensitätsschätzung erhält man pro Ort die Wahrscheinlichkeit einer Variante statt dem direkt erhobenen Wert.

Das Ergebnis einer Intensitätsschätzung hängt von dem Abstand zwischen den Orten, der Bandbreite (d.h. der Reichweite der Glättung) sowie (in schwächerem Grad) von der Art des Kerns (d.h. die Art der Glättung) ab. In Pickl et al. (2014) werden im Detail zwei Typen des Abstands verglichen, der rein *geographische* Abstand („Luftlinie“) zwischen den Orten sowie der *linguistische* Abstand.<sup>4</sup> Bei diesem wird für jedes Ortspaar ermittelt, wie oft sie korpusweit jeweils dieselbe Variante aufweisen – je öfter zwei Orte übereinstimmen, desto „näher“ sind sie sich sprachlich. Dieser Abstandstyp ist erwartungsgemäß deutlich besser dazu geeignet, scharfe Grenzen und Dialektinseln abzubilden, die unter Einsatz des geographischen Abstands schnell unerwünscht stark geglättet werden können (vgl. Pickl et al. 2014).

Abbildung 2 zeigt eine Punktsymbolkarte, wie sie für neuere germanistische Dialektatlanten typisch ist: Sie ist Band 7.1 des SBS entnommen und visualisiert pro Belegort die Realisierung von mhd. *t* im Lemma *Werktag*. Die Symbole sind zu einem gewissen Grad ikonisch: Die dünnen senkrechten Striche stehen für Lenisrealisierungen, die mittelstarken Rechtecke für einen Zwischenwert, die breiten Rechtecke für Fortisrealisierungen des Plosivs. An einigen Orten, z.B. in Wertingen (72), Kellmünz (176), Salgen (180) und Kaufbeuren (240), sind mehrere Artikulationsmöglichkeiten belegt.

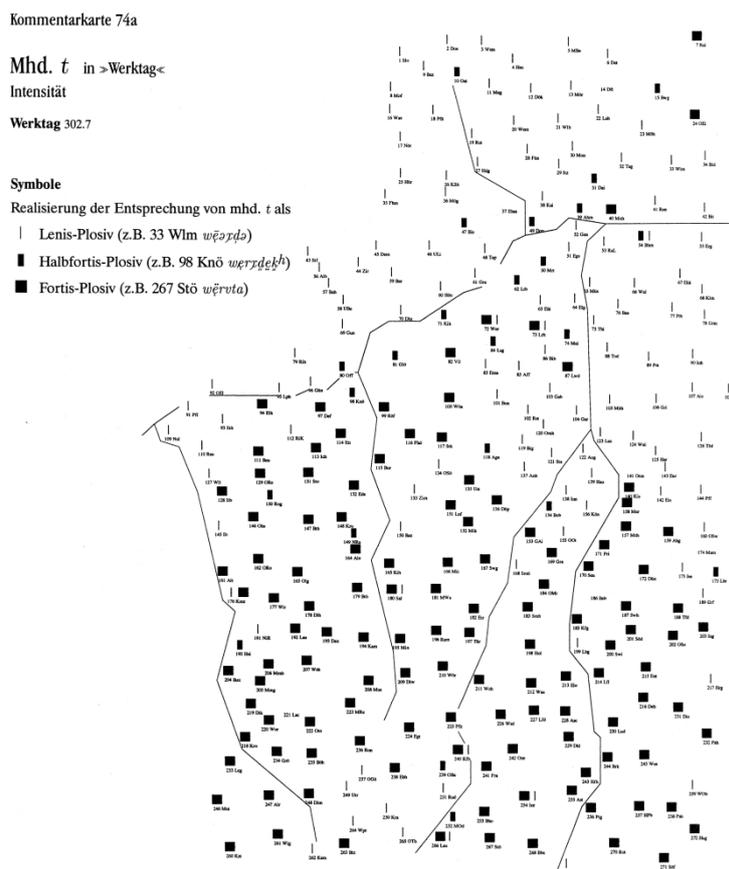


Abbildung 2: Punktsymbolkarte des SBS (Band 7.1, Karte 74a, „mhd. *t* in *Werktag*“)

<sup>4</sup> Für weitere Typen dialektometrisch nutzbarer Abstandsmaße sei z.B. auf Gooskens (2004, 2005) sowie Szmrecsanyi (2008, 2011) verwiesen. Die Software *GeoLing* ermöglicht es, beliebige eigene Distanzmaße in tabellarischer Form einzuspeisen.

Für Abbildung 3 wurde dieser Datensatz auf zwei unterschiedliche Arten mittels Intensitätsschätzung geglättet – dazu wird zunächst die Fläche des Untersuchungsgebiets in ein sogenanntes *Voronoi-Diagramm* zerlegt, wobei jedem Ortspunkt diejenigen Flächenbereiche zugeordnet werden, die ihm am nächsten sind. Jeder Variante wird eine Farbe zugeordnet; die Voronoi-Zellen erhalten die Farbe der Variante zugeordnet, die dort den stärksten Wert hat (d.h. „dominant“ ist). Nur Lenis- (blau) und Fortisrealisierungen (rot) werden dominant, die insgesamt seltener und im Kartenbild sporadischer belegten Zwischenformen dagegen nicht. Ihre Anwesenheit „schwächt“ aber die Dominanz der stärkeren Variante, was durch Aufhellung der Zellen visualisiert wird.

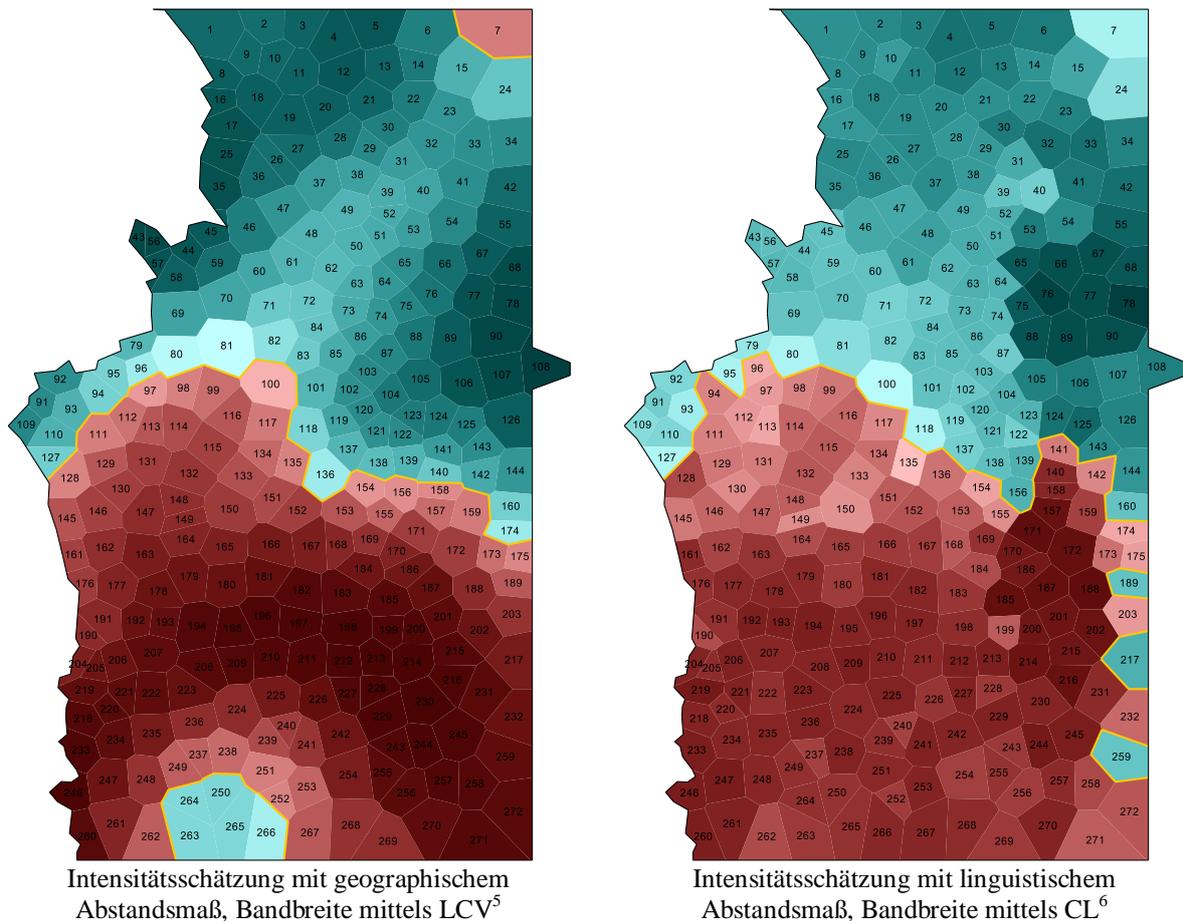


Abbildung 3: Mittels *GeoLing* intensitätsgeschätzte Umsetzungen der Daten zur Karte in Abbildung 2.

Die Karte auf der linken Seite basiert auf rein geographischen Abständen zwischen den Orten – auf der rechten Seite sind dagegen Abstände eingesetzt, die aus dem Korpus aller Lautkarten ermittelt wurden. In Abhängigkeit des Abstandsmaßes ändern sich u.a. die Varianten, die am südöstlichen Rand des Untersuchungsgebiets als dominant bestimmt werden: Während die Werte mancher Orte mit Lenisrealisierungen auf der Originalkarte unter Nutzung rein geographischer Abstände geglättet und in das größere Gebiet der Fortisrealisierungen integriert werden, scheinen sie – da die betroffenen Belegorte im Gesamtkorpus der Lautungen oft genug von ihrem Umfeld abweichen – als eigenständige Inseln durch.

<sup>5</sup> Für *likelihood cross-validation*: Die Bandbreite wird so gewählt, dass die beobachtete Variante am besten vorhergesagt wird (vgl. Pickl et al. 2014: 35–37).

<sup>6</sup> Für *Komplexität* (kurz *C*) / *Kompaktheit* (kurz *L*), ein kosten-/nutzenoptimierendes Verfahren, das einen Kompromiss zwischen möglichst zusammenhängenden Gebieten und möglichst guter Wiedergabe der Originaldaten anstrebt (vgl. Pickl 2013a: 112–113).

Darüber hinaus ist die Interpolation von Vorkommenswahrscheinlichkeiten der Varianten auch für nicht durch die Datenerhebung erfasste Punkte im Raum möglich, so dass feinmaschigere Karten gerendert werden können, die geschätzte Intensitäten „zwischen“ den Belegorten zeigen. Dies ist nur bei der Verwendung von geographischen Abständen umsetzbar, da für die geographischen Punkte zwischen den Erhebungsorten keine linguistischen Abstände ermittelt werden können. In Abbildung 4 ist das Ergebnis einer solchen Interpolation für den SBS-Datensatz zur Karte 384 in Band 9, *Vergleichspartikel (größer ... ich)*, direkt in der *GeoLing*-Oberfläche dargestellt (rot = *wie*, grün = *als*, blau = *als wie*).

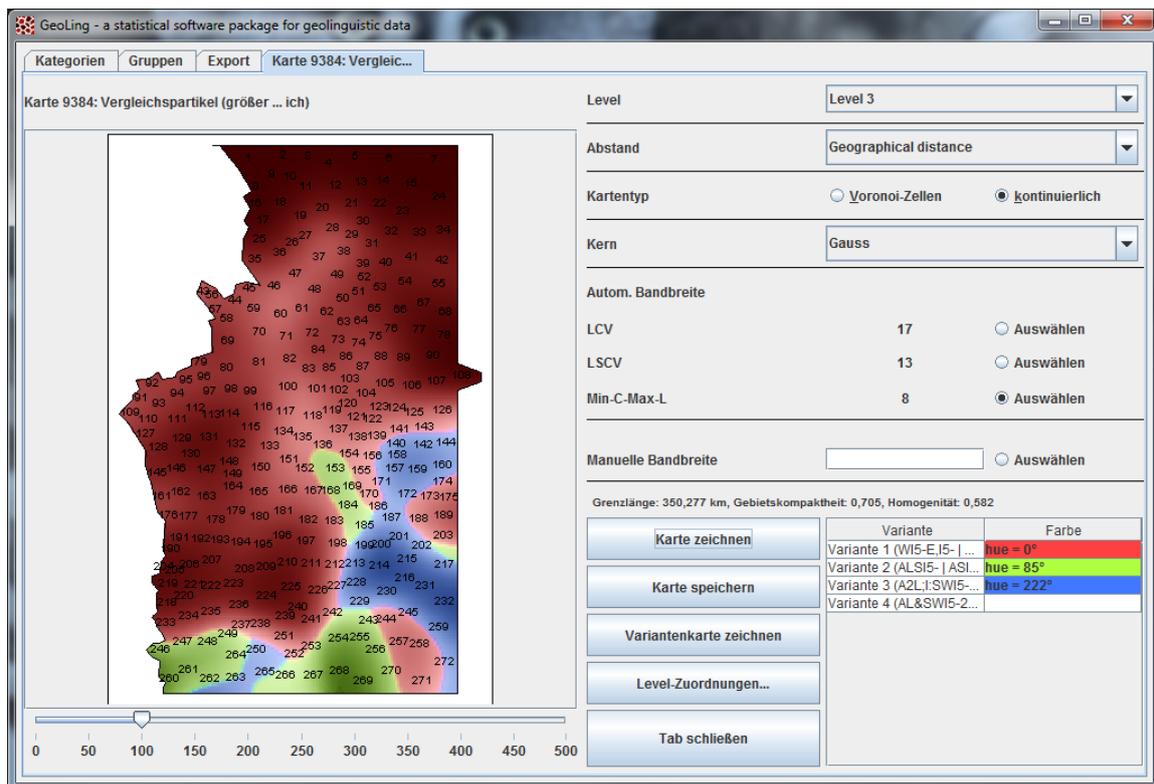


Abbildung 4: *GeoLing*-Screenshot, Interpolation einer Karte.

## 2.2 Variablenebene

Aus jeder intensitätsgeschätzten Karte lassen sich drei geostatistische Werte ermitteln, die weitere Analysen auf Variablenebene ermöglichen. Als *Komplexität* (im Folgenden kurz *C*) bezeichnen wir die Gesamtlänge der Grenzen, die sich zwischen Gebieten dominanter Varianten ergeben; die *Kompaktheit* (kurz *L*) ist ein Maß dafür, wie gut die intensitätsgeschätzte Karte die ihr zugrundeliegenden Daten wiedergibt; die *Homogenität* (kurz *B*) stellt dar, wie durchmischt die Regionen der Karte sind – klare Dominanz einer Variante führt zu hohen Werten, starke Konkurrenz der Varianten untereinander zu niedrigen.<sup>7</sup>

Basierend auf diesen Maßzahlen wertet Pickl (2013a, 2013b) die Wortschatzkarten des SBS aus (zur Einteilung der Einzelkarten in semantische Gruppen wird dafür der internen Gliederung der Printbände des SBS gefolgt): Mittels Monte-Carlo-Tests lässt sich im Vergleich der Wortarten zeigen, dass insbesondere Interjektionen geringe Komplexität bei hohen Homogenitäts- und Kompaktheitswerten aufweisen, was für große, einheitliche

<sup>7</sup> Rumpf et al. (2009) sowie Pickl & Rumpf (2011) motivieren und erläutern diese Größen im Detail.

Variantenflächen spricht. Analysiert man semantische Kategorien der Wortschatzkarten, so zeigt sich dies insbesondere für die Bereiche *Düngung, Wald und Holz, Geflügelhaltung und Imkerei* sowie in schwächerem Maße auch für *die menschliche Gesellschaft*. Genau umgekehrte Tendenzen, also Kleinräumigkeit, weisen *Wettererscheinungen, Bauern und Arbeitskräfte* sowie schwächer auch *das Bauernhaus* und *Kinderspiele* auf (vgl. Pickl 2013a: 128–135). Dagegen findet sich kein signifikanter Zusammenhang zwischen der Gebrauchsfrequenz einer Variante und ihrer Raumstruktur (Pickl 2013a: 120–123). Auch zwischen starken, schwachen und besonderen Verben lassen sich keine Unterschiede aufdecken; zwischen verschiedenen Klassen von Nomina dagegen schon (Pröll 2013a: 133–135). Klarer sind Strukturunterschiede im Vokalismus: Berechnet man jeweils für die Entsprechungen eines mittelhochdeutschen Phonems die Durchschnittswerte für *C, L* sowie *B* und interpretiert diese Werte als *Rot* (für Kompaktheit), *Grün* (Homogenität) und *Blau* (Komplexität) im RGB-Farbraum, ergibt sich ein Vokaltrapez wie in Abbildung 5.

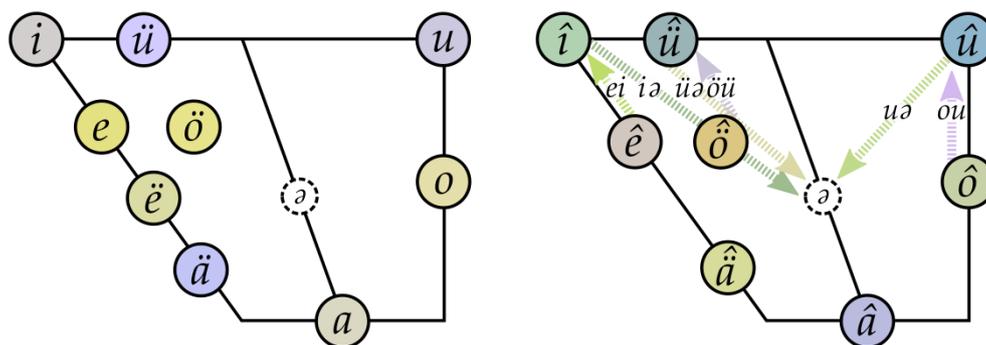


Abbildung 5: Kennwerte der Vokalkarten des SBS (links Kurzvokale, rechts Langvokale und Diphthonge).

Lila- und Blautöne indizieren relativ hohe Komplexität, während Ocker- und Gelbtöne genau dem Gegenteil entsprechen. Homogenität und Kompaktheit sind dabei etwa gleich hoch. Verschiebt sich der Farbton ins Grünliche, ist das durch erhöhte Homogenität bedingt. Im vorliegenden Fall führt das zur Interpretation, dass die kurzen Hochzungenvokale (und der Sekundärumlaut ä) räumlich komplexer sind, d.h. generell kleinere Raumareale aufweisen, während die Mittelzungenvokale zusammenhängender sind. Betrachtet man das System der Diphthonge, so findet sich dieser Gegensatz auch zwischen fallenden (üä, uä) und steigenden (öü, ou) Diphthongen. Die strukturelle Nähe der langen Hochzungenvokale ist zwar weniger deutlich ausgeprägt als bei den Kurzvokalen, eine klare Tendenz (in Form relativ hoher Homogenität) zeigt sich aber dennoch.

Interpretiert werden kann dies nicht nur als Reflex von Reihenschritten – bemerkenswert ist insbesondere, dass sich diese Muster noch in den geostatistischen Eigenschaften der Vokalkarten nachweisen lassen, nach einer Vielzahl von Eingriffen in die Daten (Exploration, Transkription, Kartierung, Dichteschätzung, Kennwertmessung, Mittelwertbildung).

Weitet man den Skopus der Kartengruppen aus und stellt gebündelt die geostatistischen Werte der Karten des SBS-Wortschatzes, der Lautung und der Formen nebeneinander, dann belegen H-Tests, dass der Wortschatz klar von den anderen beiden Subkorpora abweicht – Lautung und Formen unterscheiden sich dagegen nicht signifikant (Pröll 2013a: 152–155).

Ferner lassen es statistische Werte zu Raumeigenschaften der Karten auch zu, automatisiert und objektiviert Gruppen ähnlich strukturierter Karten zu finden. Dabei kehrt man quasi die Blickrichtung um: Statt Werte vorgegebener Kategorien zu suchen, ermittelt man über die Werte zunächst Gruppierungen von Karten mit rein räumlichen Gemeinsamkeiten. Erst im Nachhinein überprüft man, ob diese raumstrukturell ähnlichen Karten auch sprachliche Ähnlichkeiten aufweisen. Im Projekt wurden zwei verschiedene Vorgehensweisen entwickelt, die jeweils über Clusteranalysen möglichst distinkte Gruppen aus den Kartenkorpora

ermitteln können (Rumpf et al. 2010; Meschenmoser & Pröll 2012a). Beide Pilotstudien (durchgeführt an Lexikdaten) können zeigen, dass sich Karten, die raumstrukturell ähnlich sind, oft auch inhaltlich nahestehen: Insbesondere der Themenbereich *Getreide* ist in dieser Hinsicht auffällig; mit Pröll (2013b) lassen sich darüber hinaus die Charakteristika der ermittelten Cluster in Analogie zur Prototypentheorie als „Raumprototypen“ fassen.

Abbildung 6 (nach Pröll 2013a) visualisiert das Ergebnis einer Aufteilung der formengeographischen Karten des SBS (Bände 6 und 9, insgesamt 541 Karten) auf 16 Cluster. Als stellvertretender Wert für die einzelnen Karten wurde ihre jeweilige *empirische Kovarianzfunktion* eingesetzt (vgl. zum exakten technischen Vorgehen Meschenmoser & Pröll 2012a). Durch ein Quadrat wird angezeigt, wenn diese Häufigkeit überzufällig hoch ist. Die Grauskala verdeutlicht den Grad an Überdurchschnittlichkeit (gemessen an Kategorien- und Clustergröße) der Kartenzahl einer Kategorie, die dem jeweiligen Cluster zugewiesen wird.

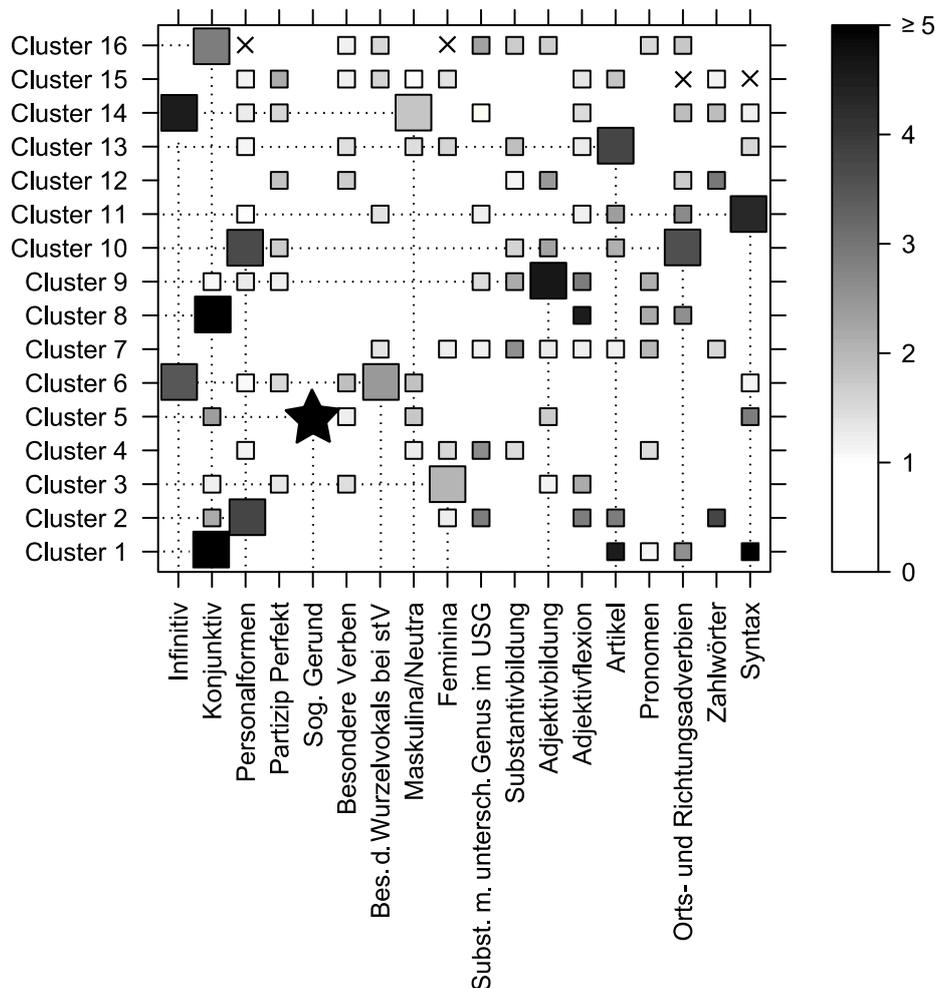


Abbildung 6: Clusterzuweisungen der formengeographischen Karten (Bände 6 und 9) im SBS (n = 541).

Ist die Überdurchschnittlichkeit signifikant (ermittelt mit Hilfe von Binomialtests), ist dies wie folgt gekennzeichnet: Ist ein Wert auf einem lokalen Alphaniveau von  $p = 0,05$  signifikant hoch, ist das Quadrat vergrößert, ist er signifikant niedrig, steht dort ein  $\times$ . Um sicherzugehen, dass die positiven Treffer nicht allein auf multiples Testen zurückzuführen sind (vgl. zum Phänomen der *Alphafehlerkumulierung* Pospeschill 2006: 178), wurde per Bonferroni-Korrektur für ein globales Alphaniveau von  $p = 0,05$  ein lokales korrigiertes Niveau von  $p < 0,000174$  errechnet. Auch dieses wird (genau einmal) erreicht, in Abbildung 6 durch einen Stern visualisiert.

Durch die Darstellung wird deutlich, welche Teilbereiche der Formenkarten sich räumlich eher entsprechen und welche nicht. Die Karten zu Infinitiv-, Konjunktiv- und Gerundformen werden relativ klar einigen wenigen Clustern zugewiesen. Auch für die Kategorien *Personalformen des Verbs*, *Orts- und Richtungsadverbien* sowie *Syntax* zeigt sich, dass sie jeweils in einem Cluster besonders stark vertreten sind. Das lässt sich dahingehend interpretieren, dass sie offenbar zu einer Raumstruktur neigen, die von der anderer Kategorien unterschieden ist. Unter den Clustern gibt es größere Cluster, in denen sich Zuweisungen nahe am Durchschnitt sammeln (z.B. die Cluster 7, 12 und 15) – sie bündeln offenbar Grundstrukturen, die tendenziell häufiger sind, also typischere Formen der Raumstruktur in der Morphologie darstellen. Gleichzeitig lassen sich aber auch kleinere Cluster mit wenigen überdurchschnittlichen Ergebnissen (vgl. etwa Cluster 1, 5 oder 8) nachweisen. Dort sind diejenigen Raumstrukturen versammelt, die quantitativ weniger auffällig sind, sich aber eher vom räumlichen Normalfall der morphologischen Raumverteilung unterscheiden.

### 2.3 Varietätenebene

Über diese Möglichkeiten der einzelvariablenbezogenen Analysen hinaus ist es ein weiteres Projektziel, Analysen auf Varietätenebene durchzuführen, ohne dabei die Bedeutung der Einzeldaten durch einen Aggregationsprozess zu schmälern (ein Anliegen, das bislang insbesondere von Nerbonne 2006 und Spruit, Heeringa & Nerbonne 2009 verfolgt wurde).

Hierzu wird auf das Verfahren der sogenannten *Faktorenanalyse* zurückgegriffen, einer ursprünglich für sozialwissenschaftliche Belange konzipierten Methode zur Dimensionsreduktion in Daten (vgl. etwa Bortz 1993: 472–521 für eine technische Einführung). Biber (1985, 1991) popularisierte sie im Rahmen korpuslinguistischer Arbeit; eine erste Anwendung in der Dialektologie erfolgte bereits durch Werlen (1984), neuere Entwicklungen sind vor allem durch Nerbonne (2006) dokumentiert. Vogelbacher (2011) leistete die Anpassung auf die spezifischen Voraussetzungen, die eine areal aufgelöste Datenbasis wie unsere mit sich bringt.

Im Rahmen einer Faktorenanalyse werden Kookkurrenzen in den Daten gesucht – wenn Varianten unterschiedlicher Variablen jeweils ähnliche räumliche Verteilungen aufweisen, muss nicht jedes Einzelvorkommen erfasst werden: Es ist ausreichend, den Zusammenhang zwischen den Einzelvorkommen mittels eines Faktors zu erfassen. Faktoren sind also distributionelle Gemeinsamkeiten, Korrelationen in den Daten (vgl. etwa Bortz 1993: 473). Hervorzuheben ist insbesondere, dass die Faktorenanalyse durchweg *data-driven* ist. Die Strukturen, die sie aufdeckt, sind somit rein durch die zugrundeliegenden Daten determiniert, nicht durch externe Hypothesen.

Die dominanten Strukturen, die im Rahmen der Analysen hervortreten, sind – mit Einschränkungen – auf den ersten Blick mit den Ergebnissen von klassisch-dialektometrischen Clusteranalysen vergleichbar. Sie sind aber gleichzeitig wesentlich differenzierter: Der besondere Mehrwert der Faktorenanalysen liegt darin begründet, dass mehr Information als nur die jeweils ‚dominante‘ verfügbar ist. Wie in der Intensitätsschätzung von Einzelkarten<sup>8</sup> zeigen die Faktorenkarten per Sättigung an, wie deutlich ein Faktor an einem Ort dominiert. Zusätzlich wird eine Vielzahl schwächerer Faktoren dokumentiert, die von den starken Faktoren überdeckt werden. Wir sprechen in diesen Fällen von *latenten Strukturen* (vgl. Pickl 2013a; Pröll, Pickl & Spetl i.E.). Die Strukturen, die der Blick auf diese unterschwelligen Anteile der Variation freilegt, bleiben mit den üblichen dialektometrischen Verfahren verborgen. Erstaunlich ist nicht nur, welche klaren

---

<sup>8</sup> Die Intensitätsschätzungen dienen der Darstellung einzelner Datensätze im Sinne einzelner Variablen; die Faktorenanalysen greifen auf die Originaldaten zu, beruhen also nicht auf geglätteten Daten.



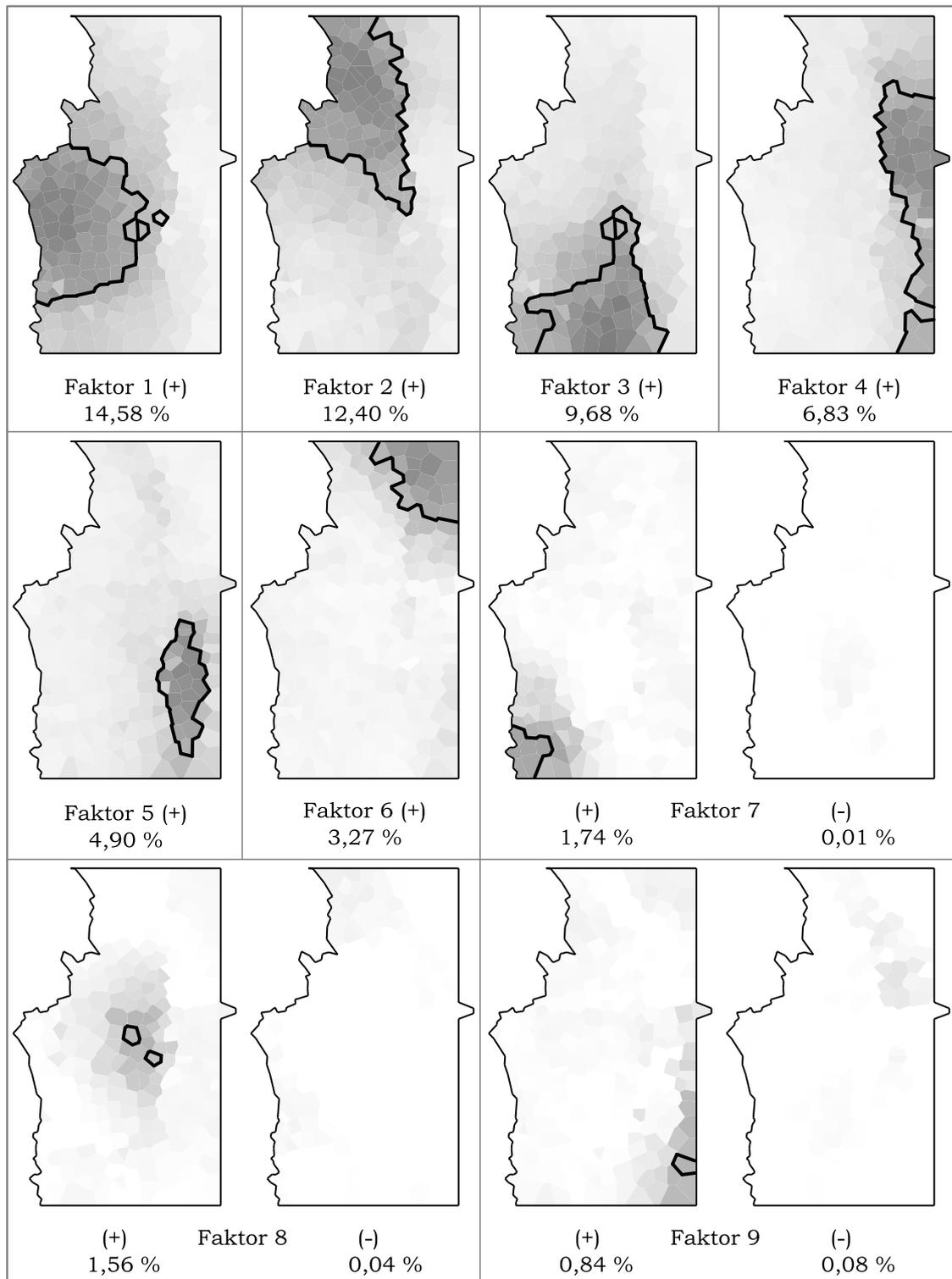


Abbildung 8: Einzelkartierungen der Faktoren 1–9.

Die stärksten Faktoren sind mit traditionellen Dialekteinteilungen in Einklang zu bringen: Faktor 1 kann man als *Mittelostschwäbisch*, Faktor 2 als *Nordostschwäbisch* charakterisieren, Faktor 3 entspricht einem Großteil des *Allgäus*. Diese drei Faktoren gehen (süd-)westlich von Augsburg relativ fließend ineinander über, wobei (meist lokal latent) Formen interferieren, die sich in Faktor 8 zeigen. Östlich des Lechs, im Norden Augsburgs, dominieren abrupt mittelbairische Formen (Faktor 4), die im Norden mit einem gewissen Übergangsstreifen zum *Nordbairischen* (Faktor 6) umschlagen. Südlich zeichnet sich dagegen klar der *Lechrain*

(Faktor 5) ab. Faktor 7 ergänzt das Allgäu, die Faktorenanalyse spaltet diese landschaftliche Einheit aus sprachlicher Sicht also in zwei Teile.

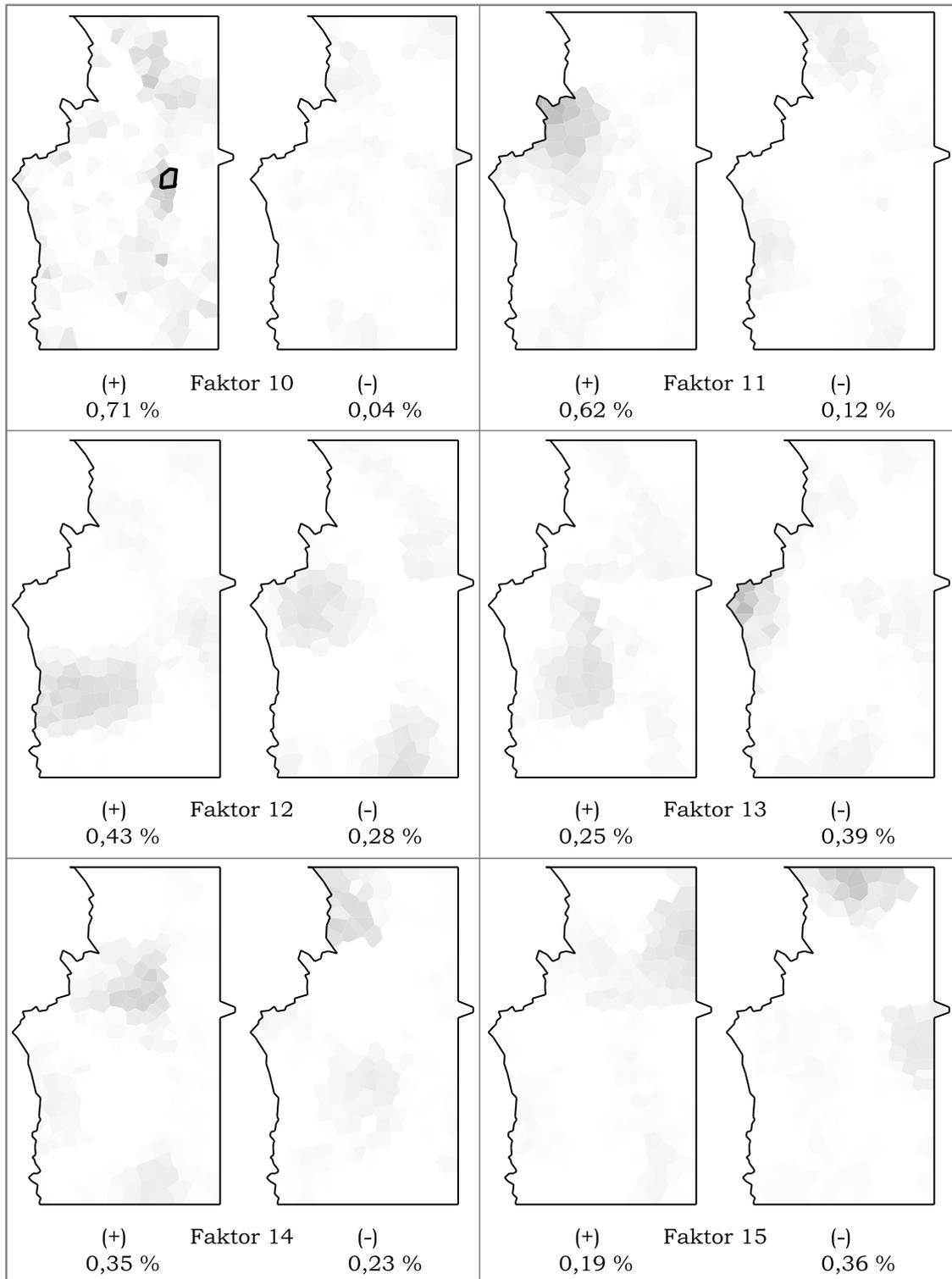


Abbildung 9: Einzelkartierungen der Faktoren 10–15.

Von den schwächeren Faktoren in Abbildung 9 wird nur noch Faktor 10 punktuell dominant. Durch diesen Faktor werden offenbar die dichter besiedelten Regionen zusammengefasst: Die Varianz, die mittels dieses Faktors beschrieben wird, korreliert mit  $r = 0,49$  (Stand 1871) bzw.  $r = 0,55$  (Stand 1970) mit der jeweiligen Einwohnerzahl pro Ort (die Daten stammen aus

SBS-Band 1). Betrachtet man die Varianten, die zu diesem Faktor beitragen, so wird deutlich, dass es sich hier überwiegend um Standardformen handelt: Faktor 10 ist demnach ein direktes Zeichen für den Kontakt zwischen Standard und Dialekt, der hier im Rückgang der angestammten Dialektvarianten resultiert. Die weiteren, allesamt latenten Faktoren bündeln quasi Tendenzen „unterhalb der Dominanzschwelle“: So stellen die Faktoren 11 und 12 die Einzugsgebiete der Marktstädte Lauingen bzw. Memmingen heraus. Faktor 13 ist möglicherweise durch die Erhebungen verursacht – er bricht scharf an einer Exploratorengrenze ein. In Faktor 14 steht der sog. „Holzwinkel“ gegen die ehemalige Grafschaft Oettingen, eine im überwiegend katholischen Bayerisch-Schwaben evangelische Region, die „vom 14. Jahrhundert an bis zu ihrer Mittelbarmachung Anfang des 19. Jahrhunderts in vergleichsweise konstanter Form existierte“ (Pickl 2013a: 153). In Faktor 15 schließlich zeichnen sich Gegensätze des nordbairischen und des fränkischen Dialektraums ab.

Mittels der Faktorenanalyse wird die Erfassung von Übergängen, die auf Variantenebene durch die Dichteschätzung auf Einzelkarten realisiert wurde, auch auf Varietätenebene möglich. Besonders deutlich wird dies anhand von Querschnitten durch die lokale Variation pro Ort, die wir hier mittels der beiden folgenden Abbildungen am Beispiel erläutern wollen. Rechts neben den Querschnitten ist der Pfad durch das Untersuchungsgebiet angezeigt.

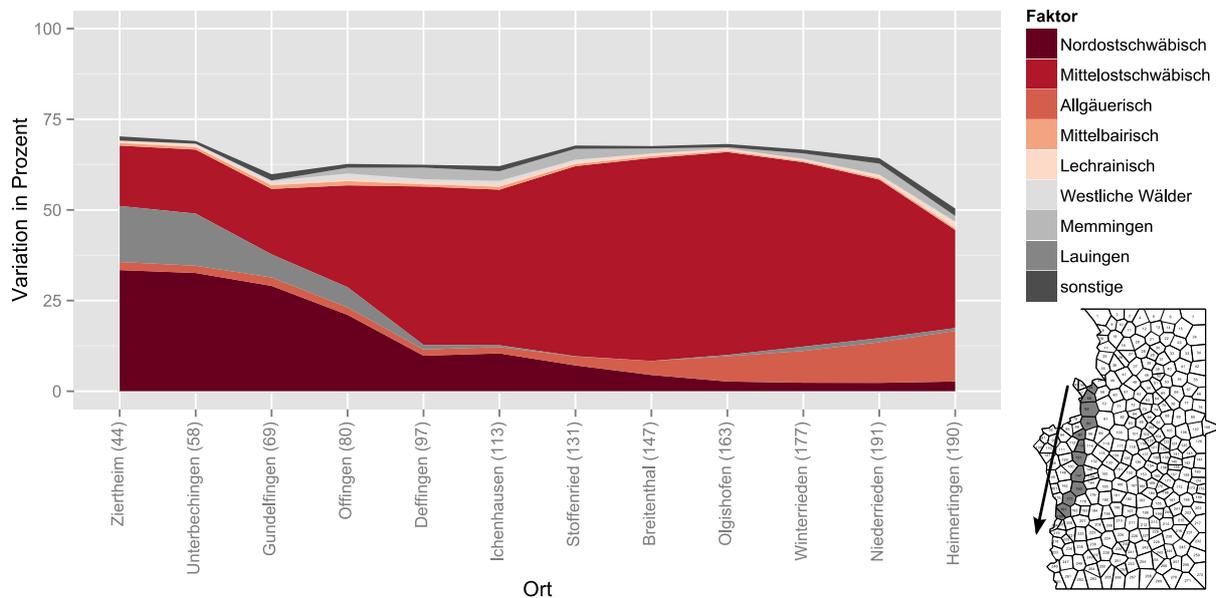


Abbildung 10: Varietätenübergang, Beispiel 1 (Kontinuum im Ostschwäbischen).

Das erste Beispiel (in Abbildung 10) beinhaltet einen Nord-Süd-Schnitt, der in Ziertheim (44) beginnt und in Heimertingen (190) endet. Die einzelnen Faktoren, die sich überlappen, wurden wieder mit Namen versehen, die auf entsprechende (auch außersprachlich gängige) Landschafts-/Kulturräume referieren. Man sieht deutlich, dass diese Entitäten fließend ineinander übergehen: So nimmt der Faktor, der mit dem Raum Lauingen in Verbindung gebracht werden kann, von Ziertheim weg kontinuierlich ab und ist ab Deffingen (97) nicht mehr relevant; in vergleichbarer Weise wird auch der „nordostschwäbische“ Faktor schwächer, während der „Allgäuer“ Faktor an Bedeutung zunimmt. Die für das „Mittellostschwäbische“ charakteristischen Formen werden insbesondere ab der Mitte der Wegstrecke stark. Die Art fließender, kontinuierlicher Übergänge der Einzelfaktoren entspricht dem, was klassischerweise mit dem Modell eines *Dialektkontinuums* beschrieben wird, erlaubt jedoch zusätzlich die Benennung unscharf ausgedehnter sprachräumlicher Einheiten.

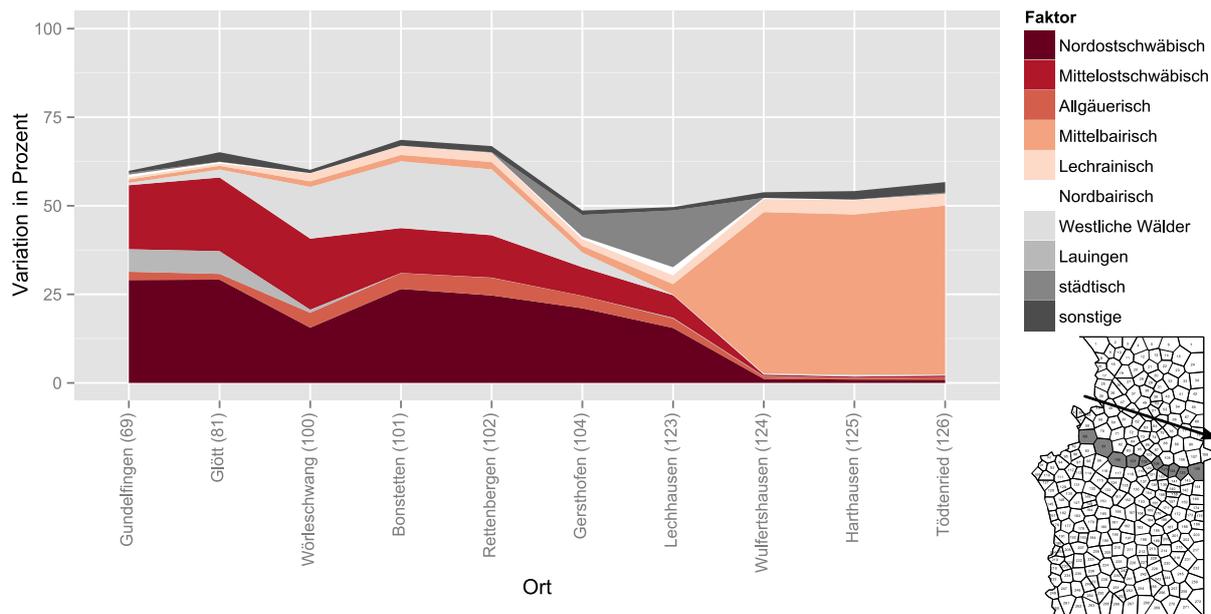


Abbildung 11: Varietätenübergang, Beispiel 2 (Bruch zwischen Alemannisch und Bairisch).

In Abbildung 11, die einen Schnitt durch das Untersuchungsgebiet zeigt, der von Westen nach Osten verläuft, fällt der Eindruck dagegen deutlich anders aus: Zunächst sind auch hier kontinuierliche Übergänge festzustellen. Das Mittellostschwäbische nimmt nach und nach ab. Ein Faktor, der typisch für ein Areal ist, das ziemlich genau dem Naturpark „Westliche Wälder“ entspricht, nimmt dagegen erst zu und wird schließlich im städtischen Umfeld von Augsburg (Ortspunkt Lechhausen) wieder schwächer; dort setzt sich ein typisch „städtischer“ Faktor durch. Dann jedoch tritt ein deutlicher Einschnitt auf: Abrupt brechen die dem Schwäbischen zuzuordnenden Faktoren ab, die vorher weitestgehend irrelevanten Formen des mittelbairischen Raums werden klar dominant. Schnittpunkt dieses plötzlichen Bruchs im Kontinuum ist der Lech.<sup>10</sup>

## 2.4 Software *GeoLing*

Zentraler Ertrag des Projekts ist, dass alle erarbeiteten Techniken der wissenschaftlichen Gemeinschaft in der Form eines kompakten und anwenderfreundlichen Softwaresystems zur Verfügung stehen: Dieses Programm mit dem Namen *GeoLing* ist unter <http://www.geoling.net> kostenlos herunterladbar. Im Download enthalten ist der komplette Satz an SBS-Daten, sowohl als Beispielkorpus als auch für die Nutzung zu eigenen, weiterführenden Analysen.

Für den Import eigener Daten ist eine grafische Oberfläche nutzbar, die aus tabellarischen Informationen eine Datenbankstruktur erzeugt (vgl. Abbildung 12). Dazu sind folgende Informationen nötig:

- eine Liste mit allen Ortspunkten und ihren Koordinaten
- eine Liste mit den Karten-/Variablennamen
- eine Liste mit Koordinaten einer Kartenaußenbegrenzung
- eine Tabelle mit den Varianten, geordnet nach Ort (Zeilen) und Variable (Spalte) und

<sup>10</sup> Damit soll nicht impliziert werden, dass der Lech rein in seiner Wirkung als naturräumliche Barriere relevant für diese scharfe Grenze sei. Es ist anzunehmen, dass der Lech lediglich eine grenzversteifende Wirkung hat, ursächlich für die beiderseitigen Unterschiede sind hingegen eher politische und ökonomische Ausrichtungen in die entgegengesetzten Richtungen.

- e) Zuordnungen der einzelnen Karten zu Kategorien und Gruppen (wie etwa Bände eines Atlas)

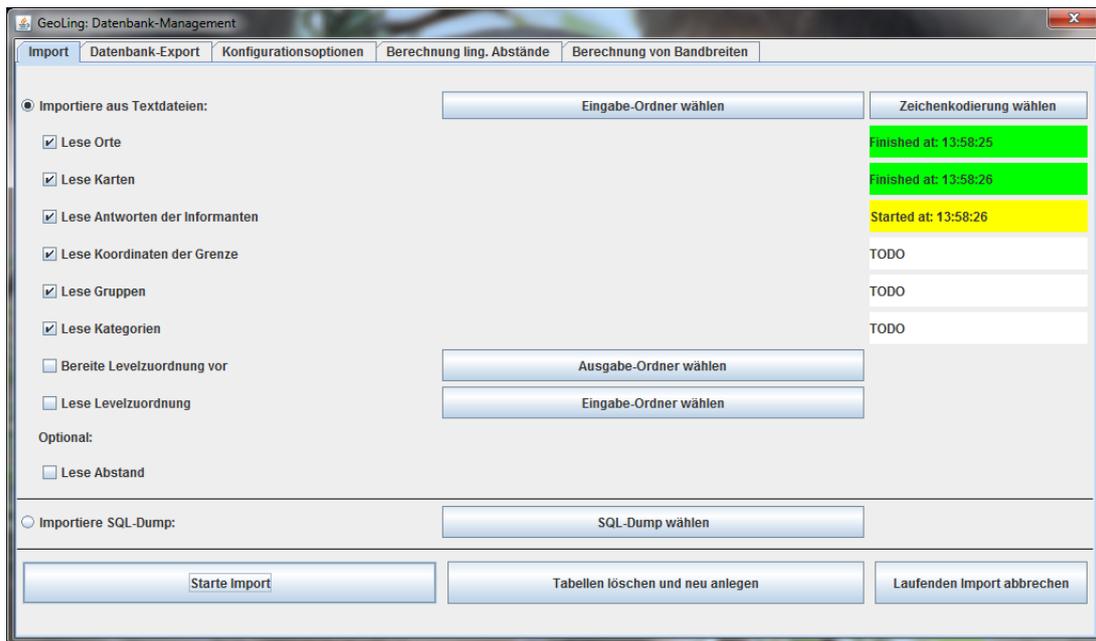


Abbildung 12: *GeoLing*-Screenshot, Import von Datentabellen.

Sämtliche Verarbeitungsschritte eigener Daten können bequem über eine klare Benutzeroberfläche durchgeführt werden (siehe Abbildung 13 mit einem Screenshot einer SBS-Faktorenanalyse). Trotzdem ist es nicht nur möglich, Karten (einzeln oder in Gruppen) in der Form von Vektorgrafiken (\*.eps) zu exportieren, sondern auch die zugehörigen Rohwerte direkt als XML-Tabellen auszugeben. Das erlaubt es, bei Bedarf weiterführende statistische Analysen oder Kartierungen auch mit anderer Software durchzuführen. Da es Distribution und Lizenz erlauben, haben erfahrene Anwender mit Programmierkenntnissen dagegen die Möglichkeit, auch direkt im Quellcode Änderungen vorzunehmen und das Programm somit individuell anpassen zu können (vgl. ausführlicher Spettl et al. i.V.).

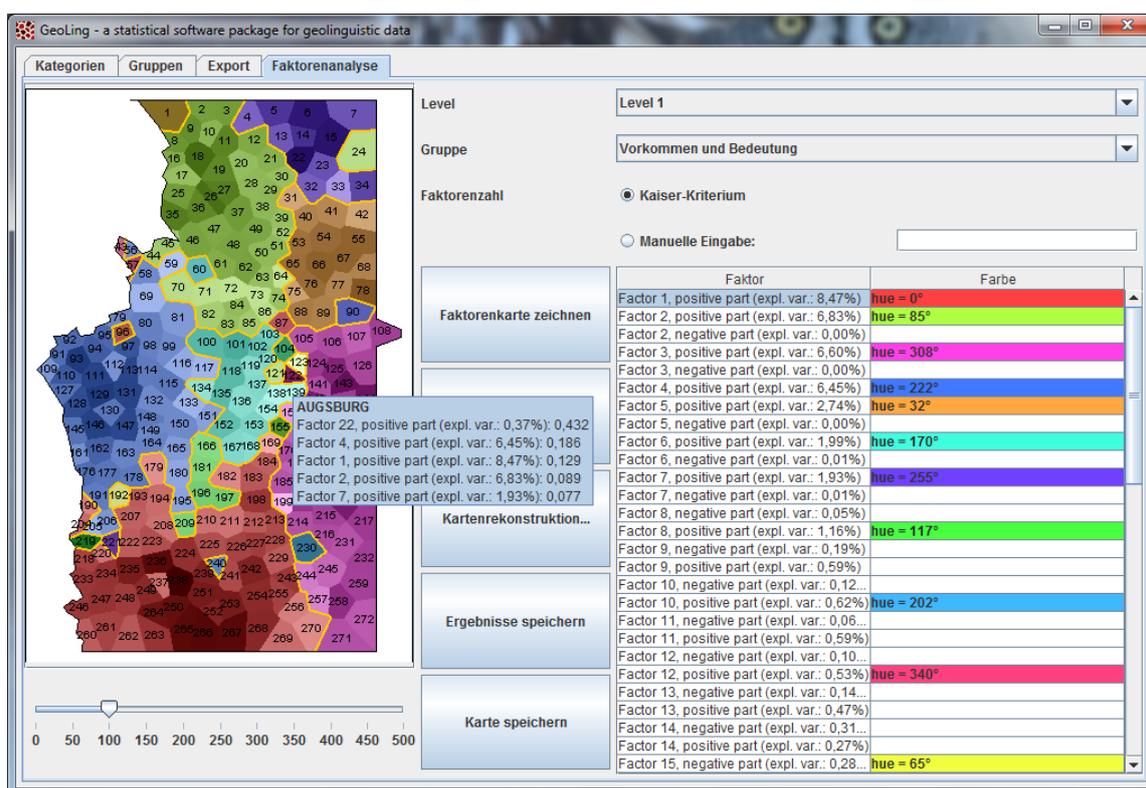


Abbildung 13: GeoLing-Screenshot, Faktorenanalyse der SBS-Karten zu Vorkommen und Bedeutung.

Teil des Onlineangebots auf <http://www.geoling.net> ist ein Handbuch, das detailliert in die Anwendung des Softwarepakets einführt.

### 3. Ausblick

Die durch das *Neue Dialektometrie*-Projekt entwickelten Verfahren ermöglichen es unserer Einschätzung nach, gezieltere Fragen zu Gestalt und Struktur räumlicher Sprachvariation zu stellen, als dies bislang möglich war, und auf neue wie alte Fragen differenziertere Antworten zu bekommen. Die Methoden sind dabei nicht als Abkehr von der bisherigen quantitativen Dialektologie und Dialektometrie anzusehen, sondern als Erweiterung des zur Verfügung stehenden Methodenrepertoires.

Ein relevanter zukünftiger Schritt wird die Klärung der Frage sein, in welchem Umfang die Datensätze räumlich anders strukturiert sein dürfen, um eine Anwendung zu ermöglichen – zwar bestehen bereits zwischen SBS und AdA deutliche Unterschiede,<sup>11</sup> aber gerade Korpora mit größeren räumlichen Lücken wären ein gutes weiterführendes Testfeld, insbesondere für die Interpolation von Werten. Da sich in den bisherigen Ergebnissen abzeichnet, dass die Dichteschätzung bei größeren Distanzen in den Daten eher dazu neigt, Feinheiten zu glätten, würde sich eine grundflächenbasierte Anpassung der Bandbreite anbieten. Auch an sprachhistorische Korpora, die notorisch lückenhaft sind, ist dabei zu denken.

Weiterhin gehen die Einsatzmöglichkeiten der Faktorenanalyse aber über einen rein areallinguistischen Bereich hinaus und eröffnen neue Perspektiven für die simultane Analyse von geographischer und sozialer Variation: Nutzt man von vornherein ein Korpus, das soziale Diversität aufweist – mehr Sprecher pro Ort, größere Altersstreuung, verschiedene soziale

<sup>11</sup> So ist das SBS-Ortsnetz aus Erhebungsgründen regelmäßiger, beim AdA sind dagegen nicht nur die Maximaldistanzen zwischen Orten größer, sondern auch die Areale mit höherer Bevölkerungsdichte als (engmaschigere) Ballungsräume wiederzufinden.

Gruppen – dann ermitteln die Analysen gleichzeitig Kookkurrenzen in den Daten, die nicht (nur) auf geographischer, sondern (auch) auf soziobiographischer oder situativer Variabilität beruhen. Es wird dann einfacher, über die (in der Soziolinguistik bereits genutzte) Faktorenanalyse sozialer und situativer Variation auch eine räumliche Dimension zu berücksichtigen: So lässt sich das Zusammenspiel sprachlicher Wirkfaktoren sozialer und situativer Natur im Raum identifizieren. Damit wird etwa für eine Region oder Stadt genauer quantifizierbar, in welchem Grade soziale Parameter (Geschlecht, Alter, Ausbildung, Situation u.ä.) rein geographischen gegenüberstehen.

## Literatur

- AdA = Elspaß, Stephan & Robert Möller (2003– ): *Atlas zur deutschen Alltagssprache*. <http://www.atlas-alltagssprache.de> (20.09.2014).
- Backhaus, Klaus, Bernd Erichson, Wulff Plinke & Rolf Weiber (2011): *Multivariate Analysemethoden*. Berlin: Springer.
- Biber, Douglas (1985): Investigating macroscopic textual variation through multifeature/multidimensional analyses. *Linguistics* 23 (2), 337–360.
- Biber, Douglas (1991): *Variation across speech and writing*. Cambridge & New York: Cambridge University Press.
- Bortz, Jürgen (1993): *Statistik für Sozialwissenschaftler*. Berlin & Heidelberg: Springer.
- Fahrmeir, Ludwig, Rita Künstler, Iris Pigeot, Gerhard Tutz (2009): *Statistik*. Berlin: Springer.
- Goebel, Hans (1982): *Dialektometrie. Prinzipien und Methoden des Einsatzes der Numerischen Taxonomie im Bereich der Dialektgeographie*. Wien: Verlag der Österreichischen Akademie der Wissenschaft.
- Goebel, Hans (1984): *Dialektometrische Studien. Anhand italoromanischer, rätoromanischer und galloromanischer Sprachmaterialien aus AIS und ALF. 3 Bände*. Tübingen: Niemeyer.
- Goebel, Hans (1986): Muster, Strukturen und Systeme in der Sprachgeographie. Explikationen zur Dialektometrie. In Guntram A. Plangg & Fabio Chiochetti (Hrsg.), *Studi ladini in onore di Luigi Heilmann nel suo 75. compleanno*, 41–71. Vich de Fascia: Istitut Cultural Ladin.
- Goebel, Hans (2005): Dialektometrie. In Reinhard Köhler, Gabriel Altmann, Rajmund G. Piotrowski (Hrsg.), *Quantitative Linguistik. Ein internationales Handbuch*, 498–531. Berlin & New York: de Gruyter.
- Gooskens, Charlotte (2004): Norwegian dialect distances geographically explained. In Britt-Louise Gunnarsson, Lena Bergström, Gerd Eklund, Staffan Fridell, Lise H. Hansen, Angela Karstadt, Bengt Nordberg, Eva Sundgren & Mats Thelander (Hrsg.), *Language Variation in Europe. Papers from the Second International Conference on Language Variation in Europe ICLaVE 2*, 195–206. Universität Uppsala.
- Gooskens, Charlotte (2005): Traveling time as a predictor of linguistic distance. *Dialectologia et Geolinguistica* 13, 38–62.
- Handl, Lisa, Aaron Spetzl, Simon Pickl, Simon Pröll, Christian Hirsch, Stephan Elspaß, Werner König, Volker Schmidt (i.V.): Statistical tests for the verification of dialect boundaries.
- Heeringa, Wilbert Jan (2004): *Measuring dialect pronunciation differences using Levenshtein distance*. Dissertation, Rijksuniversiteit Groningen.
- König, Werner (2010): Investigating language in space: Methods and empirical standards. In Jürgen Erich Schmidt & Peter Auer (Hrsg.), *Language and Space Volume 1: Theories and Methods*, 494–511. Berlin & New York: de Gruyter.

- Meschenmoser, Daniel & Simon Pröll (2012a): Using fuzzy clustering to reveal recurring spatial patterns in corpora of dialect maps. *International Journal of Corpus Linguistics* 17 (2), 176–197.
- Meschenmoser, Daniel & Simon Pröll (2012b): Automatic detection of radial structures in dialect maps: determining diffusion centers. *Dialectologia et Geolinguistica* 20, 71–83.
- Nerbonne, John (2006): Identifying linguistic structure in aggregate comparison. *Literary and Linguistic Computing* 21 (4), 463–476.
- Nerbonne, John (2010): Mapping aggregate variation. In Alfred Lameli, Roland Kehrein & Stefan Rabanus (Hrsg.), *Language and Space Volume 2: Language Mapping*, 476–495. Berlin & New York: de Gruyter.
- Nerbonne, John & William A. Kretzschmar (2013): Dialectometry++. *Literary and Linguistic Computing* 28 (1), 2–12.
- Pickl, Simon (2013a): *Probabilistische Geolinguistik. Geostatistische Analysen lexikalischer Distribution in Bayerisch-Schwaben*. Stuttgart: Steiner.
- Pickl, Simon (2013b): Lexical meaning and spatial distribution. Evidence from geostatistical dialectometry. *Literary and Linguistic Computing* 28 (1), 63–81.
- Pickl, Simon & Jonas Rumpf (2011): Automatische Strukturanalyse von Sprachkarten. Ein neues statistisches Verfahren. In Elvira Glaser, Jürgen Erich Schmidt & Natascha Frey (Hrsg.), *Dynamik des Dialekts – Wandel und Variation. Akten des 3. Kongresses der Internationalen Gesellschaft für Dialektologie des Deutschen (IGDD)*, 267–286. Stuttgart: Steiner.
- Pickl, Simon, Aaron Spettl, Simon Pröll, Stephan Elspaß, Werner König & Volker Schmidt (2014): Linguistic distances in dialectometric intensity estimation. *Journal of Linguistic Geography* 2 (1), 25–40.
- Pospeschill, Markus (2006): *Statistische Methoden. Strukturen, Grundlagen, Anwendungen in Psychologie und Sozialwissenschaften*. München: Elsevier.
- Pröll, Simon (2013a): *Raumvariation zwischen Muster und Zufall. Geostatistische Analysen am Beispiel des Sprachatlas von Bayerisch-Schwaben*. Dissertation, Universität Augsburg.
- Pröll, Simon (2013b): Detecting structures in linguistic maps – Fuzzy clustering for pattern recognition in geostatistical dialectometry. *Literary and Linguistic Computing* 28 (1), 108–118.
- Pröll, Simon (2014): Stochastisch gestützte Methoden der Dialektdifferenzierung. In Dominique Huck (Hrsg.), *Dialekte im Kontakt*, 229–242. Stuttgart: Steiner.
- Pröll, Simon, Simon Pickl & Aaron Spettl (i.E.): Latente Strukturen in geolinguistischen Korpora. Erscheint in Michael Elmentaler & Markus Hundt (Hrsg.), *Deutsche Dialekte. Konzepte, Probleme, Handlungsfelder. 4. Kongress der Internationalen Gesellschaft für Dialektologie des Deutschen (IGDD), Kiel vom 13.–15. September 2012*. Stuttgart: Steiner.
- Rumpf, Jonas (2010): *Statistical models for geographically referenced data. Applications in tropical cyclone modelling and dialectology*. Dissertation, Universität Ulm.
- Rumpf, Jonas, Simon Pickl, Stephan Elspaß, Werner König & Volker Schmidt (2009): Structural analysis of dialect maps using methods from spatial statistics. *Zeitschrift für Dialektologie und Linguistik* 76 (3), 280–308.
- Rumpf, Jonas, Simon Pickl, Stephan Elspaß, Werner König & Volker Schmidt (2010): Quantification and statistical analysis of structural similarities in dialectological area-class maps. *Dialectologia et Geolinguistica* 18, 73–100.
- SBS = König, Werner (Hrsg.) (1996–2009): *Sprachatlas von Bayerisch-Schwaben*. Band 1–14. Heidelberg: Winter.
- Séguy, Jean (1971): La relation entre la distance spatiale et la distance lexicale. *Revue de linguistique romane* 35, 335–357.
- Séguy, Jean (1973): La dialectométrie dans l'Atlas linguistique et ethnographique de la Gascogne. *Revue de linguistique romane* 37, 1–24.

- Spettl, Aaron, Simon Pickl, Simon Pröll, Stephan Elspaß, Werner König & Volker Schmidt (i.V.): Bottom-up dialectometry using the *GeoLing* package.
- Spruit, Marco René, Wilbert Heeringa & John Nerbonne (2009): Associations among linguistic levels. *Lingua* 119 (11), 1624–1642.
- Szmrecsanyi, Benedikt (2008): Corpus-based dialectometry: Aggregate morphosyntactic variability in British English dialects. *International Journal of Humanities and Arts Computing* 2 (1–2), 279–296.
- Szmrecsanyi, Benedikt (2011): Corpus-based dialectometry: a methodological sketch. *Corpora* 6 (1), 45–76.
- Vogelbacher, Julius (2011): *Statistische Analyse von Wortvarianten des Sprachatlas von Bayerisch-Schwaben*. Diplomarbeit, Universität Ulm.
- Werlen, Erika (1984): *Studien zur Datenerhebung in der Dialektologie*. Stuttgart: Steiner.
- Wiesinger, Peter (1983): Die Einteilung der deutschen Dialekte. In Werner Besch, Ulrich Knoop, Wolfgang Putschke & Herbert Ernst Wiegand (Hrsg.), *Dialektologie. Ein Handbuch zur deutschen und allgemeinen Dialektforschung. Band 1.2*, 807–900. Berlin & New York: de Gruyter.