
Simon Pickl. 2013. *Probabilistische Geolinguistik. Geostatistische Analysen lexikalischer Variation in Bayerisch-Schwaben.* (Zeitschrift für Dialektologie und Linguistik Beihefte 154). Stuttgart: Franz Steiner. 232 S.¹

Das vorliegende Buch wurde als Dissertation im Februar 2012 in Augsburg verteidigt und entstammt dem von der DFG unterstützten Projekt *Neue Dialektometrie mit Methoden der stochastischen Bildanalyse*, einer Zusammenarbeit zwischen der Abteilung für Deutsche Sprachwissenschaft, Augsburg, und dem Institut für Stochastik, Ulm. Es stellt so die Frucht einer eingehenden Zusammenarbeit zwischen den zwei Institutionen dar und wäre ohne diese kaum realisierbar gewesen. Diese Rezension konzentriert sich auf die methodischen Neuerungen in dem Buch, die beachtlich und weitreichend sind.

Der Autor hält sich im ersten Drittel des Buchs damit zurück, seine Arbeit in die Forschungslinie der Dialektometrie einzuordnen, anscheinend weil er die Aggregation von Dialektunterschieden bewusst nicht von vornherein für erstrebenswert hält. Er bekennt sich aber (S. 80) explizit zur „Dialektometrie“, und zu Recht, denn seine Arbeit beschäftigt sich eingehend mit der „Messung von Dialekten(/Dialektunterschieden)“, auch wenn er eine neue Art der Aggregation in den Mittelpunkt stellt, die wir weiter unten besprechen werden. Wenn Pickl wegen dieser Innovation zurückhaltend mit dem Begriff ‚Dialektometrie‘ ist, so muss man aber doch feststellen, dass es falsch wäre, die Methoden der Dialektometrie frühzeitig zu kanonisieren; neue Techniken aus Informatik und Statistik müssen im Gegenteil willkommen sein, denn Analysen aus verschiedenen Perspektiven sind oft aufschlussreich.

Inwiefern ein Fachmann Pickls Analyse als Geostatistik einordnen würde, wird in dieser Rezension nicht problematisiert. Freilich arbeitet Pickl geographisch und statistisch, aber Geostatistik ist eine Disziplin mit eigenen Techniken (vgl. Wackernagel 2003), die z. B. Grieve (2013) auf die Dialektologie anwendet. Pickl scheint eigene, besondere Techniken zu entwickeln, z. B. um die geographischen Verteilungen von lexikalischen Variablen zu analysieren (Kap. 4.4, in dem es keine Literaturhinweise auf andere Arbeiten gibt). Vielleicht wurden Pickls Techniken eher durch die stochastische Bildanalyse der Ulmer Partner inspiriert.

Die grundlegende Neuerung des Buches ist Pickls Entscheidung, individuelle Variablenverteilungen, wie die Frequenzverteilung der lexikali-

¹ Mein Dank gilt Andrea Mathussek für Kommentare und Simon Pickl für eine Diskussion zu dieser Rezension. Selbstverständlich bedeutet dies nicht, dass sie mit dem Inhalt der Rezension einverstanden sind.

schen Realisierung des Worts für *Kartoffelkraut*, geographisch zu glätten, anstatt Erhebungsorte immer aufgrund aller aggregierten Daten zu vergleichen, so wie es in der Dialektometrie üblich ist (Nerbonne 2009). Die Variablenverteilung wird dadurch geglättet, dass die Intensität der Variable nicht allein aus lokalen Umfrageergebnissen bestimmt wird, sondern mit Rücksicht auf die Ergebnisse in naheliegenden Erhebungsorten. Unter anderem sollte ein Fokus auf individuelle Variablen der Wissenschaft die Sicht auf die (systematischen) Unterschiede zwischen Variablen eröffnen, die letztendlich auf pragmatische und semantische Unterschiede zurückzuführen seien. Die Unterschiede reagieren ihrerseits auf unterschiedliche Art und Weise auf außersprachliche Einflüsse (S. 22), oder – und linguistisch am spannendsten – sie sind „durch intralinguistische Gründe aneinander gekoppelt“ (S. 65). Pickl schlägt an späterer Stelle im Methodenkapitel vor, mittels Faktorenanalyse die statistischen Abhängigkeiten zwischen den sprachlichen Variablen auszuarbeiten. Insbesondere gebraucht Pickl eine seltene Variante der Faktorenanalyse, die auch Datenmatrizen analysieren kann, wo es (viel) mehr Variablen als Individuen (Erhebungsorte) gibt (für Details siehe Pickl, im Erscheinen). Pickls Einsatz ist im Rahmen der Dialektometrie originell genug, um seine Übertreibung überflüssig zu machen, dass die Dialektometrie wenig Interesse an individuellen Variablen habe (S. 34). In einem Übersichtsartikel besprechen Wieling & Nerbonne (2015: §3.1) mehr als zwanzig Arbeiten mit diesem Schwerpunkt.

Das Standardverfahren in der disziplinären Geostatistik, um die Verteilung einer Variable geographisch zu glätten, heißt „Kriging“ (Grieve, 2013). Pickl gebraucht ein eigenes Verfahren, das zwar zuverlässig erscheint (Gleichungen 4-5, S. 89), aber es wäre dennoch interessant, einen Vergleich zwischen dem Standardverfahren und Pickls Methode durchzuführen. In Kapitel 5.1 folgen Studien, die von den neuen Möglichkeiten Gebrauch machen, und dieses Kapitel ist das innovativste des Buches. Die erste Studie nimmt die Frage unter die Lupe, ob die Gebrauchsfrequenz bei Wörtern, deren Gebrauchsareal groß ist, tendenziell höher ist. Es überrascht angesichts vieler Arbeiten in der historischen Linguistik, dass Pickl dies nicht bestätigt, denn diese Arbeiten zeigen, dass häufig gebrauchte Wörter sich weniger schnell ändern als seltener gebrauchte (Pagel, Atkinson & Meade 2007). Dies sollte zu größeren Gebrauchsarealen führen, wenn man annimmt, dass die Sprache früher einheitlicher war. Wie Pickl die Größe des Gebrauchsareals operationalisiert, ist vielleicht ein bedenklicher Punkt. Anstatt die Erhebungsorte zu zählen, was eine einfache Vorstellung der Größe des Gebrauchsareals zum Ergebnis hätte, schaut er sich verschiedene Maße der Verteilungskarte an, z. B. die „Komplexität“, die durch die Länge der Grenzen (zwischen Varianten)

operationalisiert wird. Komplexität müsste normalerweise eine negative Korrelation mit der Größe der Gebrauchsareale aufweisen, denn man zieht viele Grenzen bei kleinräumigen Arealen. Pickls Operationalisierung ist also plausibel, könnte aber natürlich auch anders sein. Zum Beispiel erhielte man dort, wo man Inseln mit der gleichen Varietät fände (wie im Falle des Stadtfries), sicher eine komplexe Karte, obwohl das Gesamtareal relativ groß wäre. Vielleicht kommt diese Situation aber nicht oft vor. Pickls Forschung verdiente eine Fortsetzung bezüglich dieser Frage.

Die zweite Gruppe von Teilstudien betrachtet Gruppen von Wörtern, um herauszuarbeiten, ob sie sich bezüglich ihrer Verteilung einheitlich verhalten (siehe oben "Komplexität"). Pickl aggregiert über alle Wörter einer bestimmten Kategorie, wobei die Kategorie auf jede nur denkbare Weise bestimmt werden kann; man kann sich z. B. die grammatische Wortklasse Substantive anschauen oder die pragmatische Klasse der Größe oder semantische Klassen wie z. B. Wörter, die sich auf Ackerbau beziehen. Pickl kommt u. a. zu dem Ergebnis, dass Wörter für Wettererscheinungen relativ kleinräumig sind. Die Idee, Teilaggregationen zu untersuchen, ist nicht neu (vgl. Nerbonne, im Druck), doch ist diese Fragestellung innerhalb der Dialektometrie als innovativer Schritt zu betrachten.

Die dritte Gruppe der Teilstudien testet Dialektgrenzen bei Grenzkandidaten wie Flüssen, alten politischen Grenzen und Exploratorengeländen. Auch hier aggregiert Pickl auf die übliche dialektometrische Weise (auch wenn er nur einen Teil der Abstandsmatrix gebraucht, also nicht alle Ort×Ort-Paare) und prüft dann nach, ob sich die Orte an den zwei Seiten einer Grenze stärker unterscheiden, als per Zufall zu erwarten wäre. Auch diese Studien sind m. W. in der Dialektometrie neu und bereichern die Einsatzmöglichkeiten der Techniken.

In einem langen abschließenden Teil von Kap. 5 wendet Pickl die Faktoranalyse auf 735 lexikalische Variablen an und findet für viele Faktoren (siehe unten) ein geographisch zusammenhängendes Gebiet, in dem der Faktor (bzw. die Wörter mit hohen Faktorwerten) gut repräsentiert ist. Linguistisch motivierte Faktoren werden leider nicht gefunden.

Pickl betont an mehreren Stellen, dass die Dialektometrie die Klassifikation von Erhebungsorten durch Clusteranalyse hinter sich lassen sollte (S. 33, 64f., 79), insbesondere weil die Clusteranalyse (aufgrund sprachlicher Unterscheidungen zwischen Gruppen) Grenzen liefert, die kategorisch sind und deswegen eine überdeutliche Unterscheidung suggerieren. Der Autor lässt sich sogar zu Äußerungen verleiten, die die Schlüssigkeit des Begriffs Varietät (und folglich Dialektgebiet) in Zweifel ziehen (S. 64f.). Die Clusteranalyse ist schon öfter kritisiert worden (Nerbonne et al. 2008; Prokić & Nerbonne 2008), aber sie bleibt nützlich, eben weil sie

normalerweise Dialektareale als Ergebnisse produziert, die einen Vergleich zu älteren Arbeiten ermöglichen, in denen Areale die Hauptträger der räumlichen Organisation waren. Interessanter aus Pickls Perspektive wäre ein Vergleich zu Multidimensionaler Skalierung (MDS), der populärsten Technik, um geographische Strukturen in Sprachvarietäten zu analysieren (Nerbonne, Heeringa & Kleiweg 1999; Wieling & Nerbonne 2015), denen keine kategorische Klassifikation zugrunde liegt.

Pickl fügt ein interessantes letztes Kapitel „Modell“ hinzu, in dem er zeigt, dass er sehr systematisch über Sprachvariation nachdenkt, insbesondere über die Frage, wie man weitere theoretische Begriffe dialektometrisch basieren, und dadurch genauer empirisch einsetzen kann. Das Kapitel schlägt statistische Formulierungen vor, um Begriffe aus Dialektatlasdaten numerisch zu schätzen, u. a. die INNOVATIONSAFFINITÄT (eines Begriffs), die ETABLIERUNGSRATE (einer Variante) und die DIFFUSIONSAFFINITÄT (einer Variable). Mittels dieser Begriffe zeigt Pickl eine Simulation der Diffusion einer linguistischen Variablen durch eine Dialektlandschaft. Implizit in der Begriffsbildung und der Simulation ist ein Verbesserungsvorschlag für Simulationen, die auf diese Differenzierung größtenteils verzichten (Abrams & Strogatz 2003, Nerbonne 2010), aber der Beweis, dass sie eine wichtige Rolle spielen, wird man in späteren Arbeiten suchen müssen. Weiter interessant in diesem Kapitel sind Pickls häufige Verweise auf Berruto (siehe auch §2.3) einerseits und Croft andererseits (Berruto 2004 und Blythe & Croft 2012). Die Ideen von Berruto entsprechen Pickls Überzeugung, dass Varietäten am besten als kognitive Prototypen aufzufassen seien, wofür man in sprachlichen Unterscheidungen vielleicht keine robuste Grundlage findet. Croft ist bezüglich Pickls Arbeit interessant für seine Simulationen, die sprachwissenschaftlich besonders gut untermauert sind.

Aus statistischer Sicht würde ich Pickls Arbeit an drei Punkten kritisieren. Erstens testet er öfter zweiseitig, z. B. dass die Verteilung von Substantiven (bezüglich der Komplexität) anders ist als die Verteilung zufallsbedingter Variablen. Er leitet aber hiervon zwei einseitige Hypothesen ab: $H_1: \bar{C}_i < \bar{C}_j$ und $H_2: \bar{C}_i > \bar{C}_j$ (S. 127), und prüft, ob der p-Wert für die erste *oder* die zweite in den extremsten 5 % der Verteilung liegt. Normalerweise testet man die Alternativhypothese $H: C_i \neq C_j$ und sucht einen p-Wert in den extremsten 5 %, d. h. in den niedrigsten 2,5 % oder höchsten 2,5 % der Verteilung. Pickls Herangehen verdoppelt die Gefahr von Irrtümern (Abweisung von falschen Nullhypothesen). Ein ähnliches, zweites Problem tritt bei mehrfachen Tests auf (Tabellen 6-8, S. 129-131), wo die Standardempfehlung wäre, die Signifikanzschwelle α erst durch die Anzahl der Tests zu teilen (Bonferroni-Korrektur). So müsste man nicht die Signifikanzschwelle $\alpha = 0.05$ für Tabelle 8 gebrauchen, wo von mehr

als 50 Tests berichtet wird, sondern ungefähr $\alpha = 0.001$. Glücklicherweise präsentiert Pickl die p-Werte so, dass man die nötige Information der Tabelle entnehmen kann. Der dritte Kritikpunkt betrifft den Eifer, mit dem Faktoren in der Faktoranalyse interpretiert werden. Pickl interpretiert insgesamt 20 Faktoren, einschließlich drei mit Eigenwerten unter 1,0 und zwölf, die weniger als 1 % der Varianz erklären. Er beruft sich hierfür auf Wartenberg (1985), der behauptet, dass jeder geographisch zusammenhängende Faktor zu interpretieren sei, und findet tatsächlich einen geographischen Zusammenhang für 19 der 20 Faktoren. Wartenbergs Empfehlung widerspricht aber der gängigen statistischen Praxis. Vielleicht haben Pickl und Wartenberg Recht, aber eine schlüssige Argumentation wäre wünschenswert gewesen.

Im Großen und Ganzen wurde das Buch sorgfältig geschrieben, aber kleine Fehler werden dem aufmerksamen Leser dennoch nicht entgehen. Es fehlen Ziffern in Tabelle 5 (S. 115), öfter wird das griechische 'ρ' (kleines rho) für Pearsons Produktmomentkorrelationskoeffizienten gebraucht, wo das Lateinische 'r' normal wäre (das 'ρ' ist für Spearmans Rangkorrelationskoeffizienten reserviert), und gelegentlich werden formale Details unzureichend erklärt, z. B. die Bedeutung des Pfeils in Gleichung (20) (S. 112).

Zum Schluss möchte ich das Buch jedem empfehlen, der die Arbeit eines innovativen und ambitionierten jungen Dialektologen kennenlernen will. Insbesondere ist es schön zu sehen, wie Pickl neue Probleme identifiziert, zerlegt und zu lösen versucht.

Literatur

- Abrams, Daniel M., & Steven H. Strogatz. 2003. Linguistics: Modelling the dynamics of language death. *Nature* 424.6951. 900-900.
- Berruto, G. (2004). Sprachvarietät – Sprache (Gesamtsprache, historische Sprache). In: Ulrich Ammon, Norbert Dittmar, Klaus J. Mattheier & Peter Trudgill (Hg.) *Soziolinguistik. Ein internationales Handbuch zur Wissenschaft von Sprache und Gesellschaft 2*, 188-195. Berlin, New York: De Gruyter.
- Blythe, Richard A. & William Croft. 2012. S-curves and the mechanisms of propagation in language change." *Language* 88.2. 269-304.
- Croft, W. (2000). *Explaining language change: an evolutionary approach*. New York: Pearson Education.
- Grieve, Jack. 2013. A statistical comparison of regional phonetic and lexical variation in American English. In: *LLC: The Journal of Digital Humanities Scholarship* 28. 82-107.

- Nerbonne, John. 2009. Data-driven dialectology. In: *Language and Linguistics Compass* 3. 175-198.
- Nerbonne, John. 2010. Measuring the diffusion of linguistic change. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1559), 3821-3828.
- Nerbonne, John. 2015. Various variation aggregates in the LAMSAS South. In: Michael Picone & Catherine Davies (Hg.). *Language Variety in the South III*. University of Alabama Press, Tuscaloosa. 369-382.
- Nerbonne, John, Wilbert Heeringa & Peter Kleiweg. 1999. Edit distance and dialect proximity. In: David Sankoff & Joseph Kruskal (Hg.). *Time Warps, String Edits and Macromolecules: The Theory and Practice of Sequence Comparison*. Stanford: CSLI Press, V-XV.
- Nerbonne, John et al. 2008. Projecting dialect distances to geography: Bootstrap clustering vs. noisy clustering. In: Christine Preisach et al. (Hg.). *Data Analysis, Machine Learning and Applications*. Berlin, Heidelberg: Springer. 647-654.
- Pagel, Mark, Quentin D. Atkinson & Andrew Meade. 2007. Frequency of word-use predicts rates of lexical evolution throughout Indo-European history. In: *Nature* 449.7163. 717-720.
- Pickl, Simon (im Erscheinen). Fuzzy dialect areas and prototype theory. Discovering latent patterns in geolinguistic variation. In: Marie-Hélène Côté, Remco M. Knooihuizen & John Nerbonne (Hg.) *The Future of Dialects. Selected papers from Methods in Dialectology XV*, Berlin: Language Science Press.
- Prokić, Jelena & John Nerbonne. 2008. Recognising groups among dialects. In: *International Journal of Humanities and Arts Computing* 2. 153-172.
- Wackernagel, Hans. 2003. *Multivariate Geostatistics*. Berlin: Springer.
- Wartenberg, Daniel. 1985. Spatial autocorrelation as a criterion for retaining factors in ordinations of geographic data. In: *Journal of the International Association for Mathematical Geology* 17. 665-682.
- Wieling, Martijn & John Nerbonne. 2015. Advances in Dialectometry. In: *Annual Review of Linguistics* 1. 243-264.

John Nerbonne: Dept. Informatiekunde, Fac. der Letteren. Oude Kijk in 't Jatstr. 26. Rijksuniversiteit Groningen. Postbus 716. NL 9700 AS Groningen. Nederlande