Neues aus der Salzburger Dialektometrie-Werkstatt: die *Beta*-Dialektometrie. Eine Kurzvorstellung anhand der Daten des ALD-I

Hans Goebl – Salzburg

Abstract

Es geht um die exemplarische Vorstellung der „*Beta*-Dialektometrie", die eine neue Variante der in Salzburg seit ein paar Jahrzehnten praktizierten Dialektometrie darstellt. Ihr visueller Output besteht aus optisch besonders kompakten bzw. ausdrucksstarken Choroplethen-Profilen. Dieser kartographische Sonder-Effekt wird durch eine doppelte Verrechnung der zur Verfügung stehenden Ausgangs-Daten erreicht. Während bei der (traditionellen) *Alpha*-Dialektometrie die in der jeweiligen Datenmatrix (Dimensionen: N *Messpunkte* mal p *Arbeitskarten*) enthaltenen „Ur-Daten" nur einmal hinsichtlich der in ihnen enthaltenen quantitativen Relationen verrechnet und darnach in der Ähnlichkeitsmatrix (Dimensionen: N^2) abgespeichert werden, geschieht dies bei der *Beta*-Dialektometrie zweimal: zunächst mittels eines der in Salzburg üblichen Ähnlichkeitsmaße (z. B. „Relativer Identitätswert" - RIW) und daran anschließend mittels des „Produkt-Moment-Korrelations-Koeffizienten" (PMK) von E. Bravais und K. Pearson. Die hier benützten Daten stammen aus dem Sprachatlas ALD-I und haben die folgenden Dimensionen: 217+2 *Messpunkte* mal 3.899 *Arbeitskarten* (zu allen linguistischen Kategorien). Der Artikel enthält im Anhang zehn Paare an färbigen Choroplethen-Karten, die alle einen besonderen Bezug zur brixnerisch-tirolischen Ladinia haben. Dabei wurde jeweils die obere Karte mittels der *Alpha*-Dialektometrie berechnet, während die untere Karte über eine Applikation der *Beta*-Dialektometrie entstanden ist. Die IT-technische Grundlage all dieser Berechnungen und Visualisierungen ist die neueste Version (2.24.2.3.) des von Edgar Haimerl geschaffenen Dialektometrie-Programms VDM („Visual DialectoMetry").

Part of
Videsott, P. & Videsott, R. (Eds.). (2025). *Ladin, Ladins, Ladinia. Publicazion en onour de Lois Craffonara per si 85 agn : Festschrift für Lois Craffonara zum 85. Geburtstag : Miscellanea in onore di Lois Craffonara per il suo 85° compleanno*. bu,press.
https://doi.org/10.13124/9788860462060

(cc) BY-SA

Widmung

Lieber Lois!
Unsere Lebenspfade haben sich erstmals am 18. August des fernen Jahres 1972 gekreuzt, und zwar in St. Martin in Thurn, im Garten der dortigen Bar Pineta. Damals kam es zwischen uns beiden zu einem von einer spontanen persönlichen Sympathie und einer großen wissenschaftlichen Konvergenz getragenen Gespräch, an dessen Ende der Vorsatz entstand, für die Zentralzone des *Rätoromanischen* (nach Th. Gartner) bzw. des *ladino* (im Sinne von G. I. Ascoli) einen neuen Sprachatlas zu machen. Das ist – wie allgemein bekannt – auch tatsächlich erledigt worden: in zwei Schüben und weitestgehend nach Plan: 1998 mit den sieben Bänden des ALD-I und 2012 mit den ebenfalls sieben Bänden des ALD-II. Du hast am Gelingen dieses Unternehmens einen nicht zu unterschätzenden Anteil: als generöser Mäzen in Deiner Funktion als Direktor des Istitut Ladin *Micurá de Rü* sowie als Freund und fürsorglicher Begleiter in schwierigen Stunden.

Damals – im August des Jahres 1972 – rumorte in meinem Kopf aber nicht nur die empirische Seite der Sprachgeographie, sondern auch deren theoretische Durchdringung bzw. die Frage, was bzw. welche „Tiefenstrukturen" sich denn eigentlich *hinter*, *in* oder *unter* den ungemein variablen Daten eines Sprachatlasses verbergen. Im Jahr 1973 hat ein anderer sprachgeographischer Grübler, der Tolosaner Romanist Jean Séguy (1914–1973), für solche Reflexionen den Ausdruck *dialectométrie* vorgeschlagen, den ich umgehend und voller Freude aufgegriffen habe.

Da sich aber etwa seit dem Jahr 2000 die Initiativen zur Erforschung der in Sprachatlanten verborgenen (Tiefen-)Strukturen vervielfacht und methodisch sehr ausdifferenziert haben, schien es mir angezeigt zu sein, fortan zur Charakterisierung der von mir zunächst in Regensburg (1973–1982) und darnach in Salzburg (1982 f.) entwickelten und praktizierten „Dialektometrie" (DM) den Terminus „Salzburger Schule der Dialektometrie" (*SaSDM*) zu verwenden, quasi als unverwechselbares Markenzeichen oder Logo.

Du, lieber Lois, warst auch ein effizienter Förderer dieser *SaSDM,* und zwar durch die Publikation einiger meiner ersten dialektometrischen Artikel in der von Dir im Jahr 1977 begründeten und bis zum Jahr 2001 hauptverantwortlich herausgegebenen Zeitschrift *Ladinia*. Dazu erwähne ich hier nur meine

dialektometrischen Beiträge aus den Jahren 1977, 1980 und 1981, die zudem von den vorzüglichen typographischen Standards der *Ladinia* (wie Farbdruck, Verwendung eines besonders guten Papiers etc.) ausgiebig profitieren konnten.

1. Wegmarken der *SaSDM*

Ich bin für die DM im Jahr 1970 im Rahmen einer nur einsemestrigen Assistenz an der Universität Marburg an der Lahn entflammt worden. Dies geschah vor allem in den Räumen der Bibliothek des dortigen Forschungsinstitutes „Deutscher Sprachatlas", wo Sprachatlanten aus aller Herren Länder und Zungen aufgestellt waren und tatsächlich so etwas wie einen weltweiten Überblick über die Sprachgeographie boten. In den nachfolgenden Jahren habe ich in Wien (1971–l973), Regensburg (1973f.) und Salzburg (1982f.) meine neuen Ideen methodisch weiterentwickelt, und zwar *technisch* – hinsichtlich der mit der DM eo ipso verbundenen *numerischen* und *graphischen* Herausforderungen – durch den systematischen Einbezug der damals sich allenthalben immer mehr durchsetzenden „Computer", und *methodisch* durch die vor- und umsichtige Applikation der vor allem im Rahmen der Biologie entwickelten „Numerischen Taxonomie" (auch: „Taxometrie")[1] auf die Daten romanischer Sprachatlanten wie des ALF (Frankreich) oder des AIS (Italien und Südschweiz).

Ich hatte dabei das Glück, zur Lösung diverser *IT*-Probleme – die damals noch unter dem Kürzel *EDV*[2] liefen – nacheinander vier Mitstreiter zu finden, die am späteren Erfolg der *SaSDM* entscheidend beteiligt waren. Es waren das der in Bonn tätige Kartograph Wolf-Dieter Rase, der in Münster aktive Mathematiker Hilmar Pudlatz, der an der TU Wien wirkende Mikro-Elektroniker Siegfried Selberherr[3] und – ab 1990 – der damals freischaffende Skandinavist und Mathematiker Edgar Haimerl.

Die drei Ersteren waren federführend an der Erstellung der in meiner Habilitationsschrift von 1984 präsentierten Graphiken und Berechnungen beteiligt; der Vierte kümmerte sich zunächst um die EDV-Probleme des ALD-I

1 Siehe dazu das Handbuch von H. H. Bock aus dem Jahr 1974.
2 EDV = Elektronische Datenverarbeitung.
3 Die Namen der drei Ersteren figurieren deshalb auf dem Titelblatt meiner im Jahr 1984 veröffentlichten Habilitationsschrift.

und dann, ab 1996, um jene der Dialektometrie. Edgar Haimerl kreierte dabei ein in der Folge „universell“[4] einsetzbares „Mantel“-Programm namens „Visual DialectoMetry“ (VDM), das seit damals viele Ergänzungen und Erweiterungen erfahren hat und auch relativ oft außerhalb Salzburgs mit sehr gutem Erfolg verwendet wurde bzw. immer noch wird.[5]

Die knapp vor der Jahrtausendwende von Edgar Haimerl fertiggestellte erste Version von VDM bezog sich noch auf den in meiner Habilitationsschrift[6] vorgestellten Methoden-Apparat. Größere methodische Erweiterungen kamen erst 2005 mit dem Einbau der „korrelativen DM“ dazu[7], und mit einer schon ab etwa 2007 – allerdings noch ohne Verankerung in VDM – praktizierten DM-Variante, die hier als „*Beta*-Dialektometrie“ bezeichnet wird[8]. Die ersten computativen Tests dieser zunächst rein taxometrisch interessierenden Erweiterung wurden von meinem damaligen Projekt-Mitarbeiter Slawomir Sobota durchgeführt, dem ich für seine langjähriger Treue und seine stets wache Entdeckerfreude bis heute sehr dankbar bin.

Überdies sei hier auch angemerkt, dass meinerseits an all diesen Erweiterungen nicht nur ein rein *linguistisches* Interesse, sondern auch ein gerütteltes Maß an *mathematisch-taxometrischer* Neugier beteiligt waren.

Eine dritte Erweiterung von VDM betrifft die „Areal-Bewertung“. Es ist dies ein um 2011 nach meinen Vorschlägen von Edgar Haimerl kreiertes Zusatz-Modul von VDM, das systematische Analysen an den „Grund-Bausteinen“[9] einer dialektometrischen *Datenmatrix* zulässt.[10] Letztere steht – siehe dazu die Abbildung 1, Position 2 – ganz am Beginn der dialektometrischen Methoden-Kette und entspricht formal einem zweidimensionalen Schema aus *N* Messpunkten und *p* Arbeitskarten (AK). Jede dieser (stets variationsbehafteten) AK besteht aus einer zwischen *2* und *N* schwankenden Menge an Arealen variabler Größe, wobei zu jedem dieser Areale ein eigenes „Taxat“ gehört. Die Flächen (bzw.

4 Tatsächlich kann das Programm VDM über die ganze Länge der Methodenkette der DM (siehe dazu die Abbildung 1) verwendet werden: also von der Erfassung der Ausgangsdaten bis zu den finalen Visualisierungen.
5 Siehe dazu beispielsweise Videsott (2009) und Hajek (2018).
6 Goebl (1984).
7 Goebl (2005).
8 Die derzeit größte von mir zur *Beta*-DM verfasste Publikation stammt aus dem Jahr 2024.
9 Unter diesen „Grund-Bausteinen“ versteht man in Salzburg die „Arbeitskarten“ und die Verbreitungsareale der „Taxate“: cf. dazu Goebl & Smečka (2016, auf Englisch) und Goebl (2018b, auf Deutsch).
10 Siehe dazu Goebl & Smečka (2022).

Größen) dieser „Taxat-Areale" wiederum können zwischen *1* und *N-1* (Messpunkten bzw. Polygonen) variieren. Variationslose Arbeitskarten bzw. solche, die sich nur auf ein einziges Taxat beziehen, sind dabei definitionsgemäß ausgeschlossen.

Zur Klarstellung: die AK (Arbeits-Karten) sind durch Typisierung (in Salzburg: durch „Taxierung") gewonnene Abbilder (bzw. „Messungen") der Original-Karten von Sprachatlanten, deren Inhalt in der Romanistik in aller Regel aus einigen Hundertschaften von auf einem (großen) Kartenblatt relativ gleichmäßig verteilten Lautschrift-Einträgen besteht. Diese lautschriftliche Buntheit muss – um adäquat ge- und vermessen werden zu können – vom Dialektometer zusammenfassend bzw. simplifizierend typisiert (bzw. „taxiert") werden. Es ist das ein Prozedere, das schon den Pionieren der romanischen Sprachgeographie[11] bekannt war, welche dafür die seit der Publikation des ALF bekannten und seitdem allenthalben verwendeten „stummen Karten"[12] eingesetzt haben.

Es hat sich nämlich im Lauf der Jahre gezeigt, dass die auf den Karten von Sprachatlanten feststellbaren linguistischen Formen (in Salzburg „Taxate" genannt) trotz ihrer schier unbegrenzten Vielgestaltigkeit alle zusammen gewissen statistischen Gesetzmäßigkeiten unterworfen sind (bzw. „gehorchen"), die durchaus dem entsprechen, was man bei der Analyse der Gebrauchshäufigkeiten von *Lexemen* (Wörtern), *Phonemen*, *Morphemen* und *Syntagmen* in normalsprachlichen Texten feststellen kann. Es schien daher angebracht, den räumlichen Bedingtheiten dieser großen geolinguistischen Variabilität systematisch „auf den Grund" zu gehen.

Mit Hilfe der VDM-Module zur *Korrelations-Analyse* (seit 2005) und *Areal-Bewertung* (seit 2011) konnte – auch unter Verwendung von Datenmatrizen außerroman(ist)ischer Herkunft[13] – zu den in den Sprachatlanten verborgenen Tiefenstrukturen der folgende „General-Befund" erarbeitet werden: es hat den Anschein, dass die Bewohner einer Sprachlandschaft diese in dia- bzw. basilektaler Weise ungefähr so „bewirtschaften", wie x-beliebige Sprecher das bei

11 Siehe dazu beispielsweise Jaberg (1908).

12 Zu Geschichte der „stummen Karten" in der Romanistik siehe Goebl (2018a). Man findet überdies unter dem folgenden Link eine stets *à jour* gehaltene Bibliographie aller meiner dialektometrischen Publikationen: https://goebl.ads.plus.ac.at/people/prof/goebl/dm_publi.htm

13 Dies betrifft die Deutschschweiz (Sprachatlas: SDS) und England (Sprachatlas: SED).

der Produktion normaler Texte tun, also beim *graphischen* Befüllen eines Blattes Papier oder bei der *mündlichen* Führung eines Zwiegesprächs.
In beiden Fällen hat die quantitative Linguistik schon vor vielen Jahren in der Form eigener „Text-Gesetze" die Existenz sehr stabiler Regelmäßigkeiten nachgewiesen, die sich also bei jeder Schreib- und Sprechtätigkeit – egal in welcher natürlichen Sprache – manifestieren bzw. die dabei erzeugten Produkte konditionieren: siehe dazu die Beiträge des Kapitels VI (325–409) im von R. Köhler, G. Altmann und R. G. Piotrowski im Jahr 2005 herausgegebenen Walter de Gruyter-Handbuch zur „Quantitativen Linguistik".

Auf den Kartenblättern x-beliebiger Sprachatlanten sind also Verteilungsareale zu sehen, die nach Größe und räumlicher Verzahnung zwar den Eindruck eines totalen Chaos erwecken, de facto aber einer „speziellen (reziproken) Verzahnung" (*enchevêtrement particulier, intreccio particolare, special entanglement* etc.) gehorchen, die durch das (unbewusste) Zutun ganzer Sprach- bzw. Sprecher-Gemeinschaften langsam quer durch Raum und Zeit, aber immer nach den Direktiven bestimmter mathematisch relevanter Ordnungsparameter entstanden ist und sich anhand dieser Vorgaben auch weiterhin verändert.

1.1 Vorstellung der Beta-Dialektometrie

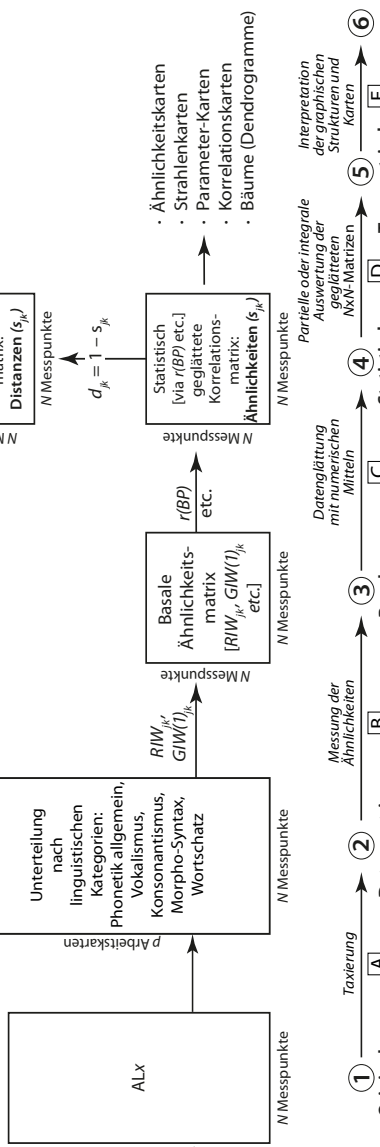


Abb. 1 – Verfahrenskette der *Beta*-Dialektometrie: vom ursprünglichen *Sprachatlas* (1) über die daraus abgeleitete *Datenmatrix* (2) und die intermediäre *Basale Ähnlichkeitsmatrix* (3) bis zu den *Statistisch geglätteten Korrelationsmatrizen* (4) und deren *Interpretationen* (5–6).

Auf der Abbildung 1 sieht man die bei der *SaSDM* der Reihe nach zu durchlaufenden Methodenschritte. Die originale Datenquelle bzw. der analysierte Sprachatlas muss zunächst taxiert werden: siehe dazu die Positionen 1, A und 2. Als Resultat dieser Taxierung entsteht eine zweidimensionale Datenmatrix (*N* mal *p*; Position 2), deren Inhalt in weiterer Folge dazu verwendet wird, anhand eines oder mehrerer Ähnlichkeitsmaße (wie *RIW*, *GIW* etc.) die reziproken Ähnlichkeiten zwischen den *N* Messpunkt-Vektoren dieser Datenmatrix zu berechnen. Im Rahmen der bis etwa 2007 exklusiv betriebenen *Alpha*-Dialektometrie wird der Inhalt der „basalen Ähnlichkeitsmatrix" (siehe die Position 3 auf Abbildung l) zur Herstellung und Interpretation verschiedener Karten-Typen[14] verwendet, die auf Abbildung 1 unter 5-E-6 aufgelistet werden.

Das Wesen der *Beta*-Dialektometrie besteht nunmehr darin, dass der metrische Inhalt der „basalen Ähnlichkeitsmatrix" (hier: Position 3) unter Einsatz quantitativer Metriken (wie *DEM*, *DMM* oder *PMK*) in eine weitere quadratische *Ähnlichkeits-* bzw. *Korrelations*-Matrix (erneut mit den Dimensionen N^2) transformiert wird. Die Methodenkette der *Beta*-DM ist also um ein Glied länger als jene der *Alpha*-DM.

Im Falle einer Visualisierung des Inhalts solcher VDM-extern errechneter Korrelationsmatrizen hat sich schon vor einigen Jahren gezeigt, dass damit kartographische Glättungs-Effekte verbunden sind, die aus rein (geo)linguistischer Sicht sehr interessant und auch gut interpretierbar sind. Mittels der von Edgar Haimerl im Jahr 2024 kreierten *Beta*-Erweiterung von VDM (Versionsnummer: 2.24.2.3.) ist es nunmehr möglich, in sehr komfortabler Weise diese Datenglättung (siehe dazu die Position C auf der Abbildung 1) mit Hilfe dreier quantitativer *Ähnlichkeits-* bzw. *Korrelations*-Maße (resp. Metriken)[15] vorzunehmen:

- mit der „Durchschnittlichen Euklidischen Metrik" (DEM)
- mit der „Durchschnittlichen Manhattan-Metrik" (DMM)
- mit dem „Produkt-Moment-Korrelations-Koeffizienten" von Bravais[16] und Pearson[17] [abgekürzt: *PMK* oder *r(BP)*].

14 Jeder dieser Karten-Typen beruht auf einer speziellen Auswertung des Inhalts der quadratischen Ähnlichkeitsmatrix (N^2) und steht damit im Dienst ganz spezieller (geo-) linguistischer Fragestellungen.

15 Die hier zitierten Metriken werden schon in meiner Habilitationsschrift (Goebl 1984) erwähnt. Siehe dazu aber auch die einschlägigen Literatur-Hinweise auf das taxometrische Handbuch von H. H. Bock (1974) in der abschließenden Bibliographie (hier in Kapitel 5.).

16 Auguste Bravais (1811–1863), französischer Physiker und Kristallograph.

17 Karl Pearson (1857–1936), britischer Mathematiker.

Die hier präsentierten *Beta*-Visualisierungen beruhen alle auf der Kombination einer basalen Ähnlichkeits-Berechnung mittels *RIW* und einer nachfolgenden Zusatz-Verrechnung mittels *PMK* [oder *r(BP)*.[18] Einsichtigerweise kann man die Leistungsfähigkeit und Tragweite der *Beta*-DM erst durch eine direkte und ausführliche Benützung der neuesten Version von VDM (2.24.2.3.) voll verstehen.[19]

1.2 Vorstellung der hier verwendeten Datenmatrix (*N* mal p)

Diese Datenmatrix wurde in den Jahren 2000-2003 in Salzburg anhand der Daten des ALD-I von einer drittmittelgeförderten Forschungsgruppe erarbeitet, die von Roland Bauer geleitet wurde. Er hat in seiner im Jahr 2009 veröffentlichten Habilitationsschrift diesen Vorgang genau beschrieben.[20]

Die im ALD-I enthaltene originäre Anzahl der Messpunkte (= 217) wurde dabei um zwei „Kunst-Punkte" erweitert; diese betreffen einerseits das Standard-Italienische (P. 999) und andererseits das Standard-Französische (P. 998). Es ergibt sich dadurch ein *N* von 219. Die Einsteuerung derartiger Kunstpunkte in ein Netz historisch gewachsener Basilekte ist von großem linguistischem Interesse, worauf wir aber in diesem Beitrag nicht näher eingehen können.

Die Gesamtzahl *p* der in dieser Datenmatrix enthaltenen Arbeitskarten beläuft sich auf 3.899; diese wiederum beruhen auf der Taxierung von rund 850 der 884 originalen Sprachkarten des ALD-I.[21] Ihre innere Variabilität erstreckt sich – entlang einer harmonisch abfallenden Kurve – zwischen 2 (mit hoher Frequenz) und 33 (mit sehr geringer Frequenz) Taxaten pro AK: siehe dazu die Abbildung 2. Diese von der *SaSDM* seit Langem und immer wieder beobachteten Regularitäten treten – so wie jene, die auf Abbildung 3 aufscheinen – bei allen Sprachatlanten auf. Sie gehören demnach zu den schon einmal erwähnten *Tiefenstrukturen*.

18 In der genuin mathematisch-statistischen Literatur wird statt PMK eher das Kürzel r(BP) verwendet.

19 Zur Erinnerung: das Programm VDM ist für jedermann frei verfügbar; siehe dazu auch den zusätzlichen Hinweis am Ende des 3. Kapitels.

20 Cf. Bauer (2009, S. 193–198). D Die in Bauer (2009) präsentierten M-Analysen beruhen auf einer etwas größeren Datenmatrix: diese verfügt über 4.020 AK und 25.604 Taxat-Areale (cf. Bauer 2009, S. 195).

21 Zu einer diesbezüglichen Auflistung siehe die Tabelle 38 in Bauer (2009, S. 173–179).

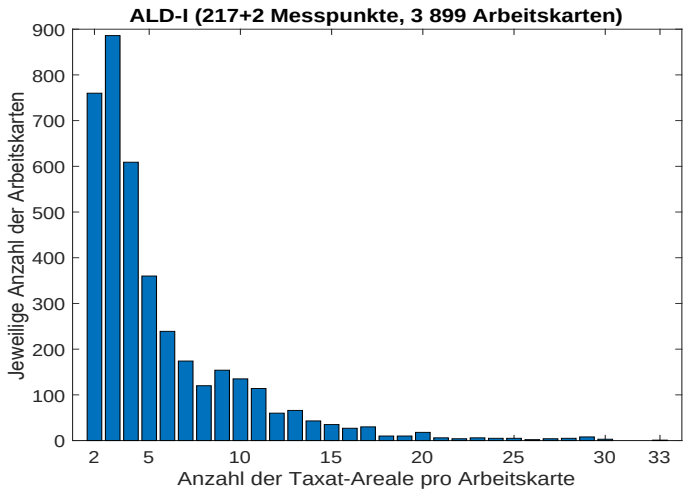


Abb. 2 – Verteilung von *einfach* und *komplex* strukturierten Arbeitskarten (AK).
x-Achse: innere Polynymie der AK (von 2 bis 33 Taxaten)
(links, erste Säule: 760 AK mit je 2 Taxaten [= binyme AK]; rechts, letzte „Säule“: 1 AK mit 33 Taxaten [= 33-nyme AK])
y-Achse: Anzahlen der verschieden strukturierten AK (von 886 [zweite Säule von links: trinyme AK] bis 1 [letzter Säulenstumpf, rechts]: 33-nyme AK).

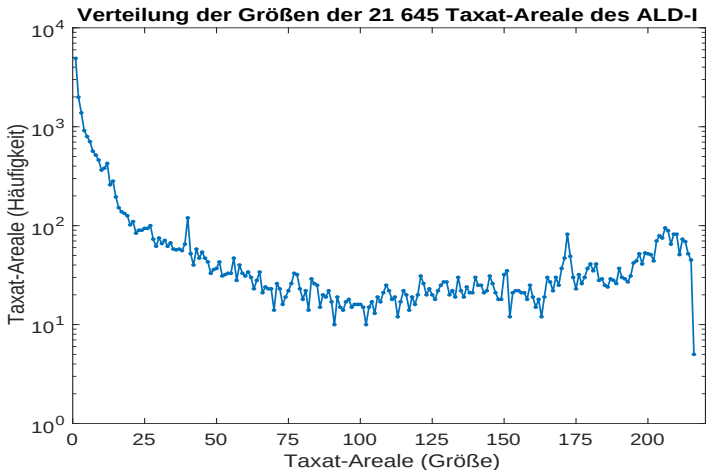


Abb. 3 – Verteilung von *kleinen* und *großen* Taxat-Arealen.
x-Achse: Größe der Taxat-Areale (von 2 bis 216 Messpunkten)
y-Achse: Anzahlen der verschieden großen Taxat-Areale (von 4.911 bis 5)
logarithmische Skalierung: $4.911 = 10^{3,691}$; $5 = 10^{0,699}$.

Die Gesamtmenge der in dieser 219 (= *N*) mal 3.899 (= *p*) messenden Datenmatrix enthaltenen Taxat-Areale (bzw. Taxate) beträgt exakt 21.645 Stück.[22] Diese betreffen vor allem die *Phonetik* (15.991, oder: 73,87%), aber auch die *Morpho-Syntax* (2.681, oder: 12,38%) und das *Lexikon* (2.973, oder: 13,73%).[23]

Man kann diese Zahlen auch zur Berechnung der Größe mittlerer Taxat-Areale heranziehen. Die Gesamtprodukt der 219 mal 3.899-Matrix beträgt 853.881 „Zellen". Diese werden von 21 645 Taxat-Arealen „bewirtschaftet". Aus der Division 853.881 : 21.645 ergibt sich der Wert 39,44 (Messpunkte). Das wäre die Größe eines durchschnittlichen Taxat-Areals unbestimmter linguistischer Zugehörigkeit.

Eine Applikation derselben Berechnungen auf die *Phonetik* und das *Lexikon* erbringt die folgenden Durchschnittswerte für die jeweils mittleren Taxat-Areale: *Phonetik*: 36,29 Messpunkte, *Lexikon*: 52,66 Messpunkte. Die *phonetische* Bewirtschaftung der ALD-Zone durch die dort ansässigen Dialektsprecher ist also kleinräumiger als die *lexikalische*.

Die räumliche Variabilität dieser 21.645 Taxat-Areale erstreckt sich in unserem Fall von 1 (umfassend also nur einen Messpunkt bzw. dessen Polygon) bis 216, wobei zwischen diesen beiden Polen erneut eine exponentielle Verteilung festzustellen ist: es gibt 4.911 (= $10^{3,691}$) Taxat-Areale mit der Minimal-Fläche 1, aber nur 5 (= $10^{0,699}$) Taxat-Areale mit Maximal-Fläche 216: siehe dazu die Abbildung 3. Auch diese Regularität gehört zu den in allen Sprachatlanten auftretenden *Tiefenstrukturen*.

Die auf unseren zehn Farbkarten sichtbaren Resultate beruhen also auf der taxometrischen Synopse von jeweils 3.899 *Arbeitskarten* (AK) bzw. von 21.645 *Taxat-Arealen*, die sich wiederum auf alle linguistischen Kategorien beziehen.

---

22 Siehe dazu auch die Graphiken und Tabellen bei Bauer (2009, S. 194–196). Für diverse Auszählungen am 3.899 AK umfassenden Korpus des ALD-I danke ich sehr herzlich Pavel Smečka (Salzburg), sowie auch Werner Goebl (Wien) für deren Visualisierung (siehe die Abbildungen 2 und 3) und für die zusätzliche Erstellung der Abbildung 1.

23 Siehe dazu auch Goebl (2018b, passim).

1.3 Zu den bei der Visualisierung benützten kartographischen Konventionen

Diese Konventionen wurden von mir in den 70er-Jahren sukzessive erarbeitet bzw. getestet und schlussendlich in kanonischer Form in meiner Habilitationsschrift von 1984 vorgestellt. Es handelt sich um zwei „Intervall-Algorithmen" (Kürzel: MINMWMAX und MEDMW) zur Aufteilung der zur Visualisierung anfallenden 218[24] oder 219[25] Messwerte bzw. Polygone auf sechs Intervalle bzw. Farbstufen. Die Letzteren sind nach der Chromatik des Regenbogens angeordnet, wobei die *kalten* Farbtöne (Dunkelblau, Blau, Grün) *unter*halb des arithmetischen Mittelwerts der 219 (bzw. 218) Einzelmessungen, und die *warmen* Farben (Gelb, Orange und Rot) *ober*halb desselben Mittelwerts liegen.

Bei MINMWMAX werden die numerischen Grenzen zwischen den jeweils drei zu beiden Seiten des Mittelwerts liegenden Intervallen (bzw. Klassen) durch Drittelung der Spanne zwischen *Minimum* und *Mittelwert* (für die Intervalle 1–3) bzw. der Spanne zwischen *Mittelwert* und *Maximum* (für die Intervalle 4–6) errechnet.

Bei MEDMW stellt ebenfalls das arithmetische Mittel die Grenze zwischen den kalten und warmen Farbtönen dar. Jedoch wird bei der Festlegung der Intervallgrenzen zu beiden Seiten des Mittelwerts darauf geachtet, dass einerseits die *unter*-durchschnittlichen (i. e. *kalt*-farben markierten) und andererseits die *über*-durchschnittlichen (i. e. *warm*-farben markierten) Klassen (Intervalle) hinsichtlich der Menge ihrer Klassenmitglieder (bzw. Messpunkte) jeweils möglichst gleich groß sind.

Ich verweise dazu auf die Zitierung der beiden Kürzel (MINMWMAX und MEDMW) ungefähr in der Mitte der oberen Statusleiste der ALD-I-Teilkarten sowie auf die numerischen Legenden im rechten oberen Eck der Karten 1–8. Man erkennt dort – jeweils zwischen runden Klammern und in der Zitierform (n = *x*) – die Menge der zu den sechs Farbklassen gehörenden Messpunkte.

Die Karten 1–3 und 5 wurden mittels MINMWMAX visualisiert, während die Karten 4 und 6–8 durch MEDMW erstellt worden sind.

---

24 Der Wert 218 (= N-1) gilt nur für die Ähnlichkeitskarten, also für die Karten 1–3. Die Reduktion von 219 auf 218 erklärt sich durch die Tatsache, dass bei allen Ähnlichkeitskarten das Polygon des Prüfbezugspunkts weiß verbleibt.

25 Der Wert 219 entspricht immer der maximalen Anzahl der Messpunkte (= N); er betrifft hier also alle Parameter- und Korrelationskarten (i. e. die Karten 4, 5 und 6).

In der rechten unteren Ecke aller Choroplethenkarten findet man außerdem Stabgraphiken (Histogramme) zu den jeweils visualisierten Häufigkeitsverteilungen. Diese veranschaulichen die in der darüber liegenden numerischen Legende aufscheinenden Zahlenverhältnisse unter Heranziehung der bei der flächigen Visualisierung eingesetzten Farben. Klarerweise gehorchen auch diese Histogramme den Direktiven des jeweils verwendeten Intervall-Algorithmus (MINMWMAX oder MEDMW).

2. Vorstellung und Interpretation des Inhalts der Karten 1–10

Allgemein gilt: Die Karten 1–8 enthalten jeweils zwei Sub-Karten: die oberen wurden jeweils nach den Direktiven der *Alpha*-DM berechnet, die unteren nach jenen der *Beta*-DM. Die Karte 9 zählt zur *Alpha*-DM, die Karte 10 zur *Beta*-DM.

2.1 Zu den Karten 1–3

Es handelt sich um sechs Ähnlichkeitskarten zu den Prüfbezugspunkten 1 (Tschlin, Unterengadin, Graubünden), 81 (Enneberg, Ladinien, Südtirol) und 195 (Forni Avoltri, Karnien, Friaul). Es sind das drei Messpunkte, die für die Ascoli-Ladinia repräsentativ sind. Wie schon erwähnt, wurden die jeweils oben befindlichen Teilkarten mittels der traditionellen *Alpha*-DM erstellt, die jeweils unten befindlichen Teilkarten dagegen durch die innovatorische *Beta*-DM. Man erkennt durch den optischen Vergleich der drei Karten-Paare sofort, dass die karten-interne Kohärenz im Sinne Ascolis in den jeweils unteren Teilkarten (alle erzeugt via *Beta*-DM) durch die Massierung von rot-orange-farbenen Polygonen deutlich besser ausgeprägt ist. Das gilt auch für die „Alterität" – gegenüber den drei rätoromanischen Blöcken im Norden – der am unteren Bildrand liegenden venedischen Gebiete, die großflächig von Dunkel- und Hellblau abgedeckt sind.

2.2 Zu den Karten 4 und 5

Hier geht es um zwei „Parameter"-Karten. Unter „Parameter" versteht man in der *deskriptiven* Statistik nicht nur einen der zahlreichen Kennwerte einer empirischen Häufigkeitsverteilung wie z. B. das *arithmetische Mittel* oder die Ex-

tremwerte *Minimum* und *Maximum*, sondern auch komplexere Kennwerte wie die *Standardabweichung* oder die *Schiefe*. Es hat sich schon in den 70er-Jahren gezeigt, dass die systematische Beachtung und Kartierung solcher Parameter für die Dialektometrie in linguistischer Hinsicht überaus bedeutsam und ertragreich sind. Die auf Karte 4 dargestellte *Schiefe* (*coefficient d'asymétrie de Fisher, coefficiente di asimmetria di Fisher, skewness*), deren Werte immer zu beiden Seiten von Null liegen, zeigt die Asymmetrie einer Ähnlichkeitsverteilung an. In aller Regel weisen die blauen (und negativen) Schiefen-Werte auf geolinguistische Übergangszonen hin, während die roten (und positiven) Schiefen-Werte die Existenz resilienter Dialektkerne signalisieren. In diesem Sinn darf man auf den beiden Hälften der Karte 4 rot-orange-farbene Verdichtungen im Unterengadin, in Ladinien und in Friaul erwarten, sowie auch – südlich davon – dunkel- und hellblau eingefärbte Übergangszonen. Tatsächlich trifft beides zu. Freilich fällt dabei auf, dass auf der unten gelegenen *Beta*-Karte das südliche Fassa-Tal (Mundart des Brach) in Dunkelblau erscheint und damit taxometrisch als „intermediär" rubriziert wird.

Die Karte 5 zeigt die geographische Variabilität von 219 *Differenz*-Werten, und zwar zwischen den *Maximal*- und *Mittel*-Werten der 219 Ähnlichkeitsverteilungen. Die dabei entstehenden Choroplethen-Profile ähneln weitgehend jenen der Karte 4 bzw. bestätigen diese.

2.3 Zur Karte 6

Die beiden Teile der Karte 6 zeigen die räumliche Schichtung jener 219 [= N] Korrelationen, die mittels des „Produkt-Moment-Korrelations-Koeffizienten" (*PMK*) zwischen den 219 *Messpunkt*-Vektoren (berechnet durch den „Relativen Identitätswert", *RIW*) der *Ähnlichkeits*-Matrix und den 219 *Messpunkt*-Vektoren der *Nachbarschafts*-Matrix (ermittelt anhand der „Geographischen Proximitäten", *GeoProx*) eruiert werden können.

Dabei werden die *GeoProx*-Werte unter Zuhilfenahme der geographischen *x*-/*y*-Koordinaten der ALD-Messpunkte, des altbekannten Satzes des Pythagoras ($a^2 = b^2 + c^2$) und einer einfachen Konversionsformel ($1 = s + d$) zwischen *Ähnlichkeiten* (*s*) und *Distanzen* (*d*) berechnet. Alle diese Berechnungen erfolgen VDM-intern und damit blitzschnell.

Die theoretische Grundannahme besteht hier in der Frage nach der Ausprägung der Konvergenz zwischen der *basilektalen* und der *euklidischen* Bewirt-

schaftung desselben Natur-Raumes: existiert hier eine Art von „Gleichschritt" oder eine mehr oder weniger große „Dissonanz"?

Man erkennt aus dem Choroplethenprofil der beiden Teilkarten, dass die diesbezüglich größten Konvergenzen (siehe dazu die Farben *Rot* und *Orange*) im Bereich des Venedischen und des östlichen Trentinischen anzutreffen sind, und auch, dass in Graubünden und Ladinien deutlich geringere Konvergenzen zwischen „Raum" und „Sprache" festzustellen sind. Die dunkel- und hellblau markierten Zonen signalisieren hingegen eine deutliche Divergenz zwischen diesen beiden Bewirtschaftungs-Modalitäten und verweisen damit auf Friktionen zwischen den Domänen von „Sprache" und „Raum", die *historische*, *politische*, *religiöse*, *soziale* oder sonstige ,,anthropische" Gründe haben können.

2.4 Zu den Karten 7 und 8

In beiden Fällen handelt es sich um „Zwischenpunkt"-Karten (frz. *cartes à interpoints*,[26] ital. *carte interpuntuali*) deren Bildaussagen entweder *differentiell* (Karte 7) oder *kommunikativ* (Karte 8) relevant sind. Die Erzeugung der hier zu Grunde liegenden 578 Zwischenpunkte erfolgte in mehreren Etappen:

- Triangulierung des Messpunkte-Netzes: es ergeben sich dabei 578 Dreiecksseiten.
- Errichtung der Streckensymmetralen über diesen 578 Dreiecksseiten.
- Verlängerung aller dieser Streckensymmetralen bis zu deren sternförmiger Verschmelzung. Die solcherart entstehenden 578 Polygone heißen *Thiessen-*, *Haag-* oder *Voronoi*-Polygone.[27]

Die Karte 7 benützt die 578 eruierten *Polygon-Seiten* zur nach Farbe und Strichdicke differenzierten Visualisierung und signalisiert damit – ganz wie bei den altbekannten Isoglossen-Bündeln – interpunktuelle Dialekt-Verschiedenheiten.

---

26 Die hier verwendete Metaphorik des Zwischen-Punkts (frz. *interpoint*) stammt aus der französischen Dialektologie, wo das Prinzip der Triangulierung eines Messpunkte-Netzes erstmals (1949–1950) systematisch vom Abbé Théobald Lalanne (1880–1952) angewendet worden ist. Dabei hat Lalanne in sehr hellsichtiger Weise die zwischen zwei benachbarten Messpunkten bestehenden Potential-Differenzen erfasst, gezählt und visualisiert.

27 Zur Bezeichnung dieser in vielen Disziplinen benützten Polygone sind die Namen mehrerer „Erfinder" bzw. „Erstverwender" im Umlauf:
- Alfred H. Thiessen, amerikanischer Meteorologe, 1872–1956.
- Karl Haag, schwäbischer Philologe, 1860–1908.
- Georgy Feodorovich Voronoi, russischer Mathematiker, 1868–1908.

Die dafür seit vielen Jahren eingebürgerten Namen lauten: dt. *Schottenkarte*, frz. *carte à cloisons*, ital. *carta isoglottica.*

Dagegen benützt die Karte 8 nur die 578 durch Triangulierung ermittelten *Dreiecks-Seiten*, und visualisiert diese nach den schon auf Karte 7 verwendeten kartographischen Normen (d. h. durch variable Strichdicke und Einfärbung). Dabei entsteht der Eindruck eines auf interpunktueller Kommunikation basierenden Netzwerks, das aus mehreren Interaktions-„Kernen" und einer größeren Anzahl von dazwischenliegenden Übergangszonen besteht. Die dafür seit vielen Jahren üblichen Bezeichnungen sind: dt. *Strahlenkarte*, frz. *carte à rayons*, ital. *carta a raggi*.

Wie bei der Karte 4 zeigt sich auch hier beim Vergleich der Visualisierungsresultate mittels MINMWMAX und MEDMW, dass durch die Applikation von MEDMW heuristisch bzw. sehpsychologisch besseren Visualisierungs-Resultate erzielt werden können.[28] Diese Erfahrung hat sich bei den Karten 5–8 wiederholt.

Auf der unteren Teilkarte (*Beta*-DM) der Karte 7 sieht man deutlich die markante Abschottung des Engadins, der (v. a. nördlichen) Ladinia und Friauls. Diese Prägnanz ist auf der oberen Teilkarte (*Alpha*-DM) geringer.

Auf der Karte 8 wird durch die optische Konvergenz zahlreicher dicker und roter Dreiecksseiten die Existenz mehrerer Interaktions-Kerne suggeriert, so z. B. im Trentino entlang der Etsch, in Gröden und im Gadertal sowie in den dialektal venedischen Provinzen Vicenza, Treviso und Belluno, die alle zur italienischen Region Veneto gehören.

2.5 Zu den Karten 9 und 10

In beiden Fällen handelt es sich um *hierarchisch-agglomerative* Klassifikationen, die auch *Baum-* bzw. *Cluster-Analysen*, *Baum-Graphen* oder *Arboreszenzen* genannt werden. Es ist das ein taxometrisches Verfahren, das direkt auf eine quadratische Ähnlichkeits- oder Korrelationsmatrix zugreift und sich dabei einer Vielzahl[29] verschiedener *Algorithmen* bedient. Das finale Produkt dieser Klassifikationen sind binär strukturierte Bäume, deren Ramifikationen (bzw. Astwerke)

28 Das Programm VDM erlaubt einen raschen Wechsel zwischen drei verschiedenen Intervallalgorithmen bei paralleler Benützung von jeweils bis zu 20 nach der Logik des Regenbogens sortierten Farbstufen.

29 In VDM sind sechs derartige Baum-Algorithmen implementiert.

vom Linguisten bzw. Dialektometer entsprechend ausgewertet werden müssen. Das bedeutsamste Resultat einer solchen Auswertung sind ein hinsichtlich seiner wichtigsten Verzweigungen (*Dendreme*)[30] eingefärbter Baum und eine auf denselben Farben beruhende Karte, die die Projektion (oder: Spatialisierung) dieser Dendreme in den Raum (*Choreme*) zeigt. Die obgenannten Algorithmen legen ex ante jene Distanzen fest, die vom Taxometer sowohl *innerhalb* (= *intra-*) als auch *zwischen* (= *inter-*) den Dendremen dieser stets binär strukturierten Bäume präferiert werden.

Die Karten 9 und l0 beruhen auf einem Algorithmus, der 1963 vom amerikanischen Statistiker Joe H. Ward Jr. (1926–2011) vorgeschlagen worden war. Der für (geo)linguistische Zwecke sehr gut brauchbare Ward-Algorithmus produziert in aller Regel Gruppierungen (zunächst als *Dendreme*), die im Falle ihrer Projektion in den Raum (als *Choreme*) sehr kohärente und vor allem linguistisch überaus plausible Areale ergeben.

Die Erzeugung dieser Klassen beruht im Prinzip auf der Fusion zwischen einander sehr nahen Ähnlichkeits- bzw. Distanz-Werten. Diese Fusionen beginnen bei den größten der $N/2\ (N-1)$[31] errechneten Ähnlichkeitswerte und enden bei deren Minima. Dabei entstehen $N-1$ hierarchisch angeordnete Bifurkationen (auf den Karten 9 und 10 erkennbar als nach links gekippte T-Formationen, etwa in dieser Form: ⊦), die alle zusammen die Gestalt eines Baumes ergeben. Dieser verfügt immer über *eine* „Wurzel" (bzw. *einen* „Stamm") und über N unverbundene „Blätter", zwischen denen $N-1$ Ramifikationen liegen, die linguistisch nützliche Informationen liefern. Bei den auf den Karten 9 und 10 quergedreht präsentierten Bäumen liegen die *Blätter links* sowie der *Stamm* bzw. die *Wurzel rechts*.

Bei der linguistischen Interpretation beginnt man immer bei den „obersten" bzw. wurzel-nahen Bifurkationen (hier ganz außen *rechts*). Dabei ist zu beachten, dass jene Klassen (Gruppen oder Dendreme), die sich nahe beim *Stamm* (oder: bei der *Wurzel*) befinden, eine deutlich größere innere Variabilität (*intra-*

30 Die *Dendreme* sind „bedeutsame" Teile eines Klassifikations-Baumes; die *Choreme* sind deren Entsprechungen auf der zweidimensionalen Fläche einer Karte bzw. Teile der aus der Baumgliederung abgeleiteten „Spatialisierung".

31 Diese Formel definiert den dialektometrisch verwertbaren Anteil einer quadratischen Ähnlichkeitsmatrix (N^2). Diese besteht immer aus zwei inhaltsgleichen Hälften und einer dialektometrisch nutzlosen Diagonale (N) mit dem Standardwert 1 (oder 100%); daher: $N^2/2 - N = N/2(N - 1)$.

group variability) aufweisen als jene, die sich in der Nähe der *Blätter* (hier also ganz außen *links*) befinden. Zudem können aus der Lage der stets T-förmigen Bifurkationen auch *diachrone* Folgerungen abgeleitet werden: je *näher* bei Stamm oder Wurzel → desto *älter*.

Auf den Karten 9 (zur *Alpha*-DM) und 10 (zur *Beta*-DM) verfügen die Dendrem-Chorem-Strukturen der beiden Bäume über dieselbe Farblogik und können daher sowohl einzeln als auch vergleichend betrachtet werden. So erkennt man auf Karte 9, dass die drei rätoromanischen Blöcke einerseits an der ersten (bzw. obersten) Verzweigung des Baumes hängen und andererseits die jeweiligen Dendreme in sich kohärente Choreme erzeugen.

Auf Karte 10 (zur *Beta*-DM) ist dieser Effekt noch viel deutlicher ausgeprägt: an der obersten und zugleich ersten Bifurkation „hängen" das Unter- und Oberengadinische (Vallader und Puter), das Grödnerisch-Gadertalische und das Friaulische, alle durch verschiedene Rot- und Violett-Töne markiert.

Dagegen bilden das Fassanische (inkl. Moena, P. 101) und das Buchensteinische (beide in Gelb) ein eigenes Dendrem, das aber – anders als auf der Karte 9 – an der unteren Erst-Verzweigung hängt.

3. Zusammenfassung

Die heuristische Nützlichkeit der *Beta*-DM wächst mit der Größe des jeweils untersuchten Grundnetzes. Die bislang anhand der Daten von ALF (638 Messpunkte) und AIS (407 Messpunkte) erzielten Glättungs-Effekte sind daher um vieles eindrücklicher, als das hier (Netzgröße: 217+2 Messpunkte) der Fall ist. Auch ergeben sich bei der *Beta*-Verwendung der metrischen Maße *DEM* und *DMM* an der Stelle von *PMK* sehr interessante Zusatz-Effekte, die hier nur angedeutet werden können.

Da klarerweise zur vollen Kenntnisnahme aller „Geheimnisse" der *Beta*-DM der direkte Zugang zur Version 2.24.2.3. von VDM erforderlich ist, wiederhole ich hier mein schon öfters mitgeteiltes Angebot: VDM ist für jedermann frei verfügbar!

Allerdings wäre es für seriös an VDM Interessierte angesichts von dessen Komplexität ratsam, ein paar Tage für die (unentgeltliche) Einschulung in Salzburg zu reservieren.

Bibliographie

AIS = Jaberg, Karl & Jud, Jakob (Hrsg.) (1928–1940). *Sprach- und Sachatlas Italiens und der Südschweiz*. Zofingen. Ringier, 8 vols. Neudruck: Nendeln: Krauss 1973. Im Netz: *Padua*: https://navigais-web.pd.istc.cnr.it/; *Zürich*: https://www.ais-reloaded.uzh.ch [letzter Abruf: 30. Juni 2025].

ALD-I = Goebl, Hans, Bauer, Roland, Haimerl, Edgar et al. (Hrsg.) (1998). *Atlant linguistich dl ladin dolomitich y di dialec vejins, 1ª pert/Atlante linguistico del ladino dolomitico e dei dialetti limitrofi, 1ª parte/Sprachatlas des Dolomitenladinischen und angrenzendet Dialekte, 1. Teil*. Wiesbaden: Dr. Ludwig Reichert Verlag, 7 vols. Im Netz [*digitALD*]: https://www.ald.gwi.uni-muenchen.de/?db=ald1

ALD-II = Goebl, Hans et al. (Hrsg.) (2012). *Atlant linguistich dl ladin dolomitich y di dialec vejins, 2ª pert/Atlante linguistico del ladino dolomitico e dei dialetti limitrofi, 2ª parte/Sprachatlas des Dolomitenladinischen und angrenzender Dialekte, 2. Teil.* Wiesbaden: Dr. Ludwig Reichert Verlag, 7 vols. Im Netz [*digitALD*]: https://www.ald.gwi.uni-muenchen.de/?db=ald2

ALF = Gilliéron, Jules & Edmont, Edmond (eds.) (1902–1910). *Atlas linguistique de la France*. Paris: Champion, 9 vols. Neudruck: Bologna: Forni 1968. (Im Netz: *Innsbruck*: https://diglib.uibk.ac.at/urn:nbn:at:at-ubi:2-4568; *Toulouse*: http://symila.univ-tlse2.fr/alf; *Grenoble:* https://lig-tdcge.imag.fr/cartodialect5/#/recherche

Bauer, Roland (2009). *Dialektometrische Einsichten: Sprachklassifikatorische Oberflächenmuster und Tiefenstrukturen im lombardo-venedischen Dialektraum und in der Rätoromania*. S. Martin de Tor: Istitut Ladin "Micurà de Rü".

Bock, Hans Hermann (1974). *Automatische Klassifikation. Theoretische und praktische Methoden zur Gruppierung und Strukturierung von Daten (Cluster-Analyse)*. Göttingen: Vandenhoeck & Ruprecht.

FEW (1922–2002) = Wartburg, Walther von (2001). *Französisches etymologisches Wörterbuch*. Basel: Zbinden et al., 25 vols.

Goebl, Hans (1977). *Rätoromanisch* versus *Hochitalienisch* versus *Oberitalienisch. Ladinia 1*, 39–71.

Goebl, Hans (1980). Dialektgeographie + Numerische Taxonomie: Dialektometrie. Anhand rätoromanischer und oberitalienischer Dialektmaterialien (AIS). *Ladinia 4*, 31–95.

Goebl, Hans (1981). Isoglossen, Distanzen und Zwischenpunkte. Die dialektale Kammerung der Rätoromania und Oberitaliens aus dialektometrischer Sicht. *Ladinia 5*, 23–55.

Goebl, Hans (1984). *Dialektometrische Studien. Anhand italoromanischer, rätoromanischer und galloromanischer Sprachmaterialien aus AIS und ALF. Unter Mitarbeit von Siegfried Selberherr, Wolf-Dieter Rase und Hilmar Pudlatz*, 3 vols. Tübingen: Niemeyer [Beihefte zur Zeitschrift für romanische Philologie, Bände 191–193].

Goebl, Hans (2005). La dialectométrie corrélative. Un nouvel outil pour l'étude de l'aménagement dialectal de l'espace par l'homme. *Revue de Linguistique Romane 69*, 321–367 [mit 24 Farbkarten].

Goebl, Hans (2018a). La face cachée de la géographie linguistique. Bref aperçu sur les «cartes muettes» produites pour l'ALF, l'AIS et le FEW. *Revue de Linguistique Romane 82*, 5–63 [mit 16 Farbkarten].

Goebl, Hans (2018b). Zu den zwei Grundbausteinen der Salzburger Dialektometrie: Arbeitskarten und Taxat(areal)e. In Guido Mensching, Jean-Yves Lalande, Jürgen Hermes & Claes Neuefeind (Hrsg.), *Sprache – Mensch – Maschine. Beiträge zu Sprache und Sprachwissenschaft, Computerlinguistik und Informationstechnologie für Jürgen Rolshoven aus Anlass seines sechsundsechzigsten Geburtstages* (S. 287–309). Köln: Kölner UniversitätsPublikationsServer [KUPS].

Goebl, Hans (2024). Du nouveau de l«'École Dialectométrique de Salzbourg»: brève présentation du mode «Béta» de la dialectométrie de Salzbourg. *Dialectologia* (*Special Issue) 14*, l–39.

Goebl, Hans & Smečka, Pavel (2016). The Quantitative Nature of Working Maps (WM) and Taxatorial Areas (TA): A Brief Look at two Basic Units of Salzburg Dialectometry (S-DM). In Emmerich Kelih, Róisín Knight, Ján Mačutek & Andrew Wilson (eds.), *Issues in Quantitative Linguistics 4. Dedicated to Reinhard Köhler on the occasion of his 65th birthday* (S. 113–127). Lüdenscheid: RAM-Verlag [mit vier Farbkarten].

Goebl, Hans & Smečka, Pavel (2022). Quarante-quatre regards «ensemblistes» sur les données de l'ALF. Une étude dialectométrique «à la salzbourgeoise». *Revue de Linguistique Romane 86,* 357–391 [mit 14 Farbkarten].

Hajek, Sandra (2018). *Struktur und Entwicklung der dialektalen Variation in der Campania. Eine dialektometrische Analyse*. Göttingen: Phil. Diss. [im Netz

verfügbar: https://ediss.uni-goettingen.de/handle/11858/00-1735-0000-002E-E364-B].

Jaberg, Karl (1908). *Sprachgeographie. Ein Beitrag zum Verständnis des* Atlas linguistique de la France. Aarau: Sauerländer. (Spanische Übersetzung: *Geografía lingüística, Ensayo de interpretación del* Atlas lingüístico de Francia. Granada: Universidad de Granada, 1959).

Köhler, Reinhard; Altmann, Gabriel & Piotrowski, Rajmund G. (Hrsg.) (2005). *Quantitative Linguistik/Quantitative Linguistics. Ein internationales Handbuch/ An International Handbook.* Handbücher zur Sprach- und Kommunikationswissenschaft, Band 27. Berlin, New York: De Gruyter.

Lalanne, abbé Théodore (1949–1950). *L'indépendance des aires linguistiques en Gascogne maritime.* Saint-Vincent-de-Paul: Chez l'auteur, 2 Faszikel. (Neudruck: Labatut: Atelier Histoire trait d'union, 2018).

SDS = *Sprachatlas der deutschen Schweiz* (1962–1997). Begründet von Heinrich Baumgartner und Rudolf Hotzenköcherle in Zusammenarbeit mit Konrad Lobeck, Robert Schläpfer, Rudolf Trüb und unter Mitwirkung von Paul Zinsli, herausgegeben von Rudolf Hotzenköcherle, fortgeführt und abgeschlossen von Robert Schläpfer, Rudolf Trüb und Paul Zinsli. Bern, Basel: Francke, 8 vols.

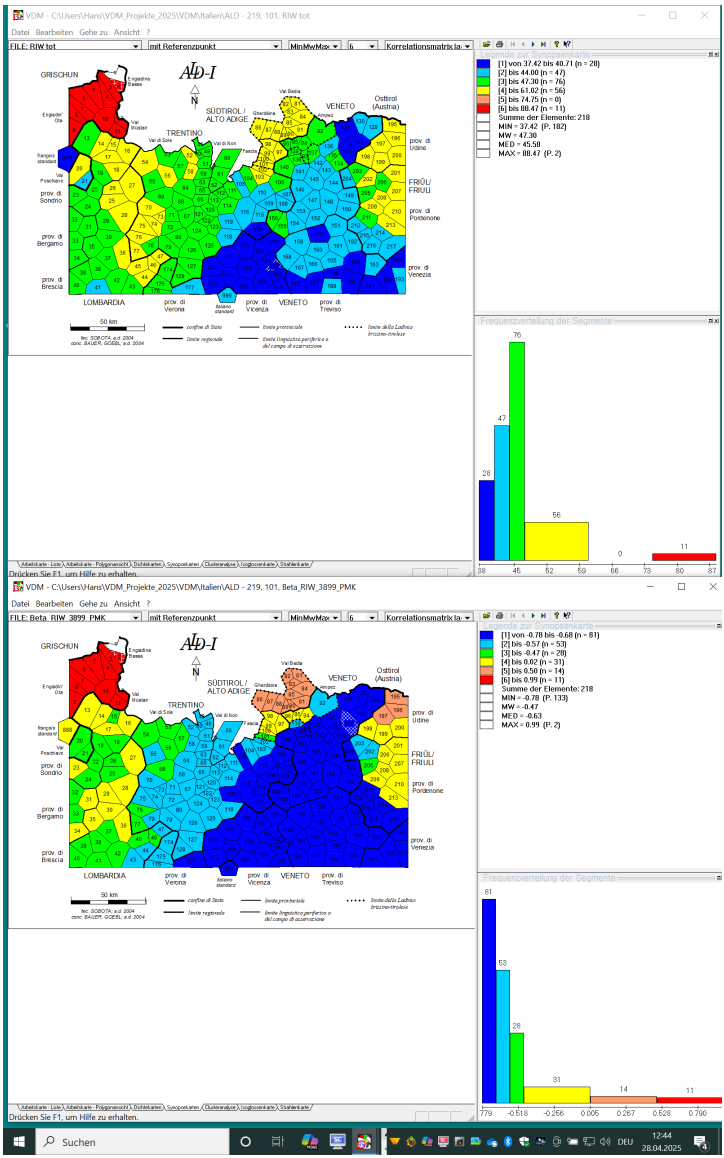
SED = Orton, Harold; Halliday, Wilfrid Joseph; Dieth, Eugen & Wakelin, Martyn F. (eds.) (1962–1971). *Survey of English dialects. The basic material*. Leeds: E. J. Arnold, 12 vols. [Neudruck: London: Routledge, 1998].

Séguy, Jean (1973). La dialectométrie dans l'«Atlas linguistique de la Gascogne». *Revue de Linguistique Romane 37*, l–24.

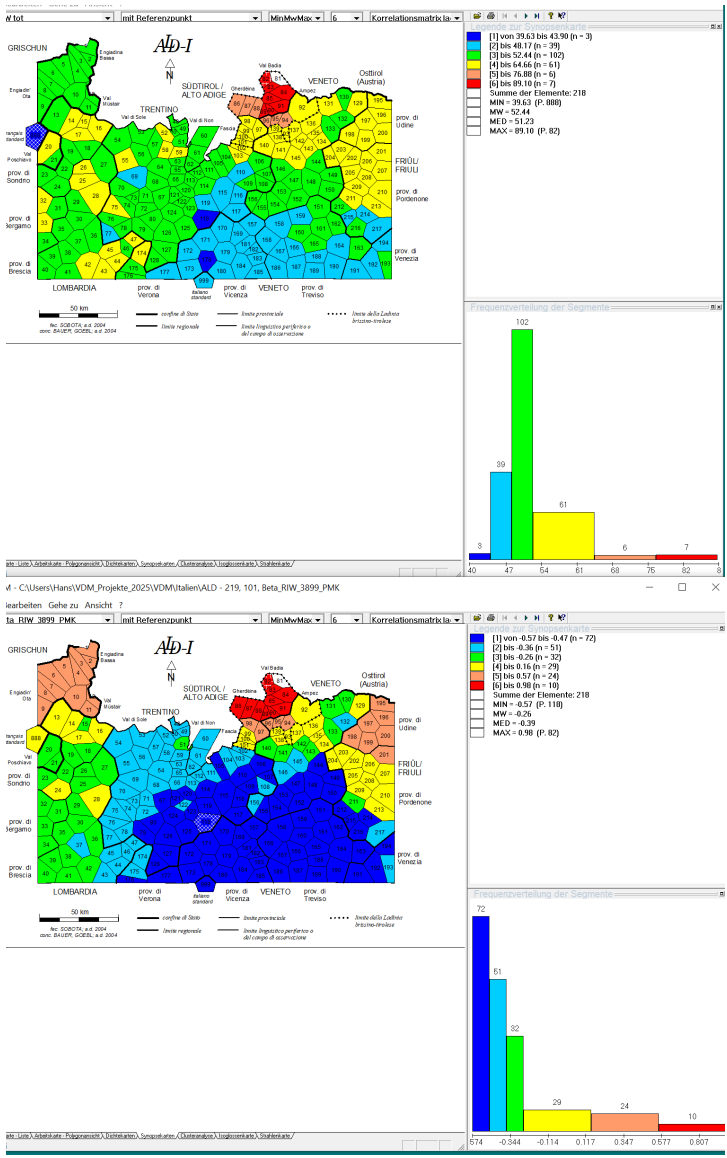
Videsott, Paul (2009). *Padania scrittologica. Analisi scrittologiche e scrittometriche di testi in italiano settenrionale antico dalle origini al 1525*. Tübingen: Niemeyer [Beihefte zur Zeitschrift für romanische Philologie, Band 343].

Abkürzungen

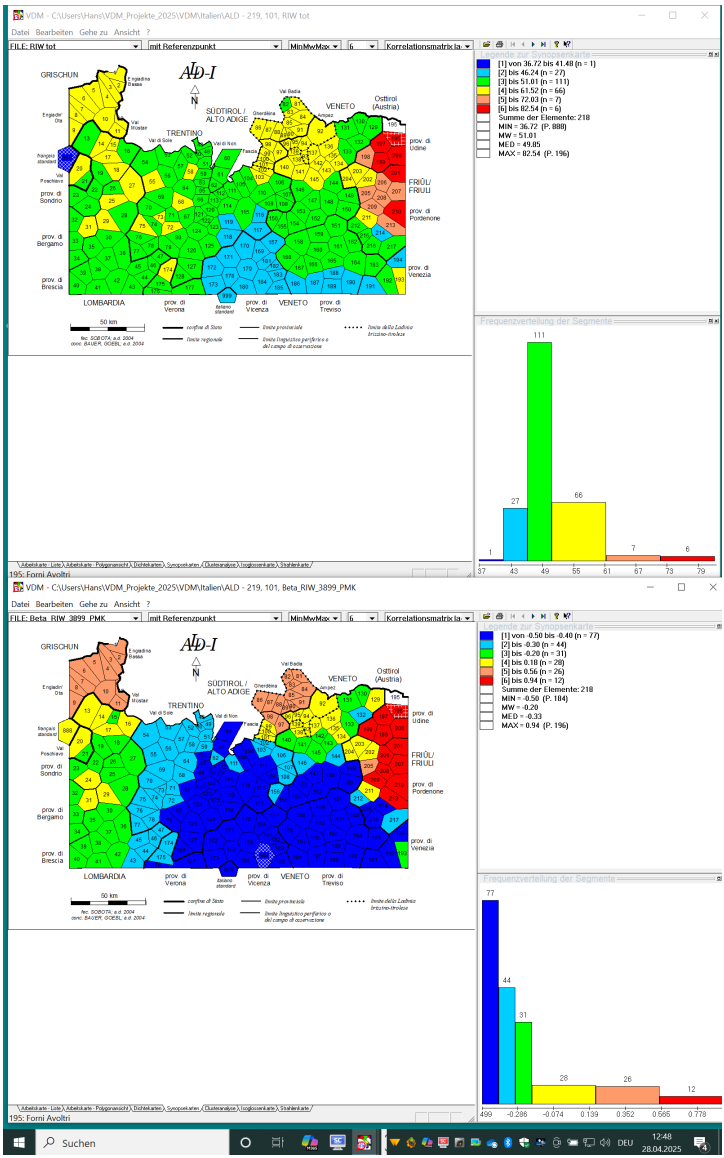
AK: Arbeitskarte(n); Definition in: Goebl 1984, I, 57 f.
DEM: „Durchschnittliche Euklidische Metrik"; Definition in: Bock 1974, 37.
DM: Dialektometrie, dialektometrisch.
DMM: „Durchschnittliche Manhattan Metrik"; Definition in: Bock 1974, 39.
GIW: „Gewichteter Identitätswert"; Definition in: Goebl 1984, I, 83 f.
p: in einer Datenmatrix mit den Dimensionen *N x p*: Anzahl der Arbeitskarten (AK).
PMK: „Produkt-Moment-Korrelations-Koeffizient"; Definition in: Bock 1974, 44 und 60.
r(BP): andere Abkürzung für *PMK.*
RIW: „Relativer Identitätswert": Standard-Ähnlichkeitsmaß der *SaSDM*; Definition in: Goebl 1984, I, 75.
SaSDM: Salzburger Schule der Dialektometrie
VDM: Visual DialectoMetry: Computerprogramm (© Edgar Haimerl).



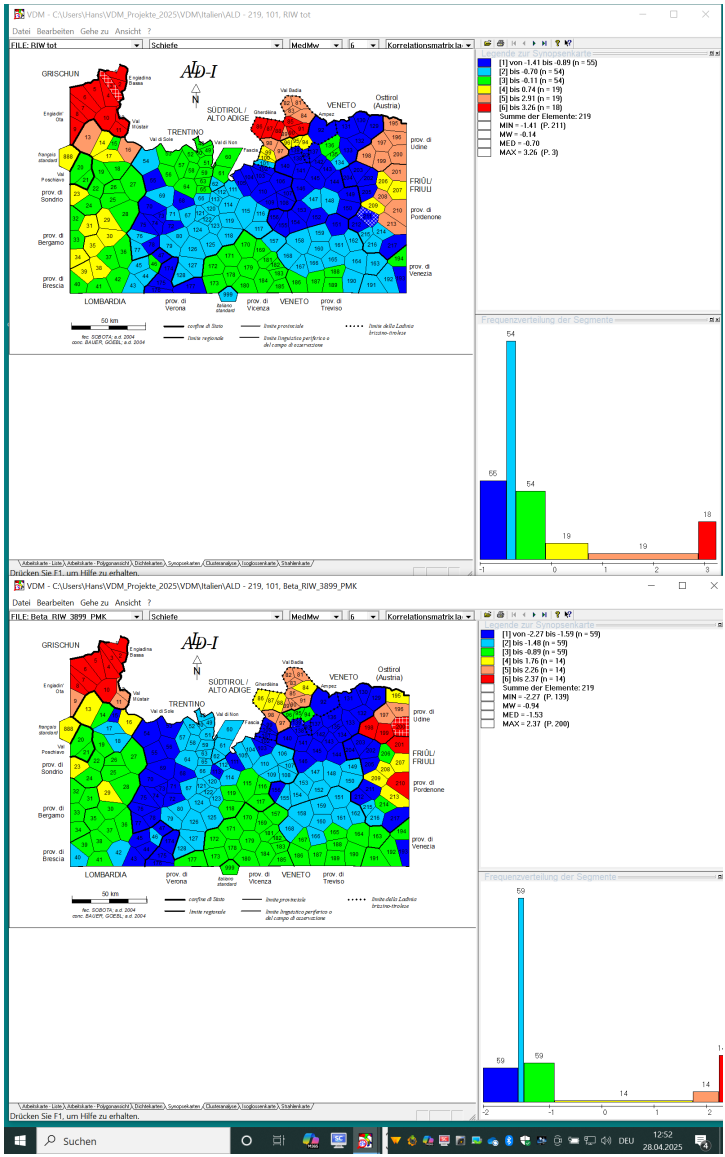
Karte 1 – Zwei Ähnlichkeitskarten zum ALD-I-Punkt 1 (Tschlin, Unterengadin)
Oben (*Alpha*-Dialektometrie): Ähnlichkeitsmessung: *RIW* (Relativer Identitätswert)
Unten (*Beta*-Dialektometrie): Ähnlichkeitsmessungen: *RIW* + *PMK*
(Produkt-Moment-Korrelations-Koeffizient [nach Bravais-Pearson])
Visualisierung: Intervallalgorithmus MINMWMAX 6-fach
Korpus: 3.899 Arbeitskarten aus ALD-I (alle linguistischen Kategorien).



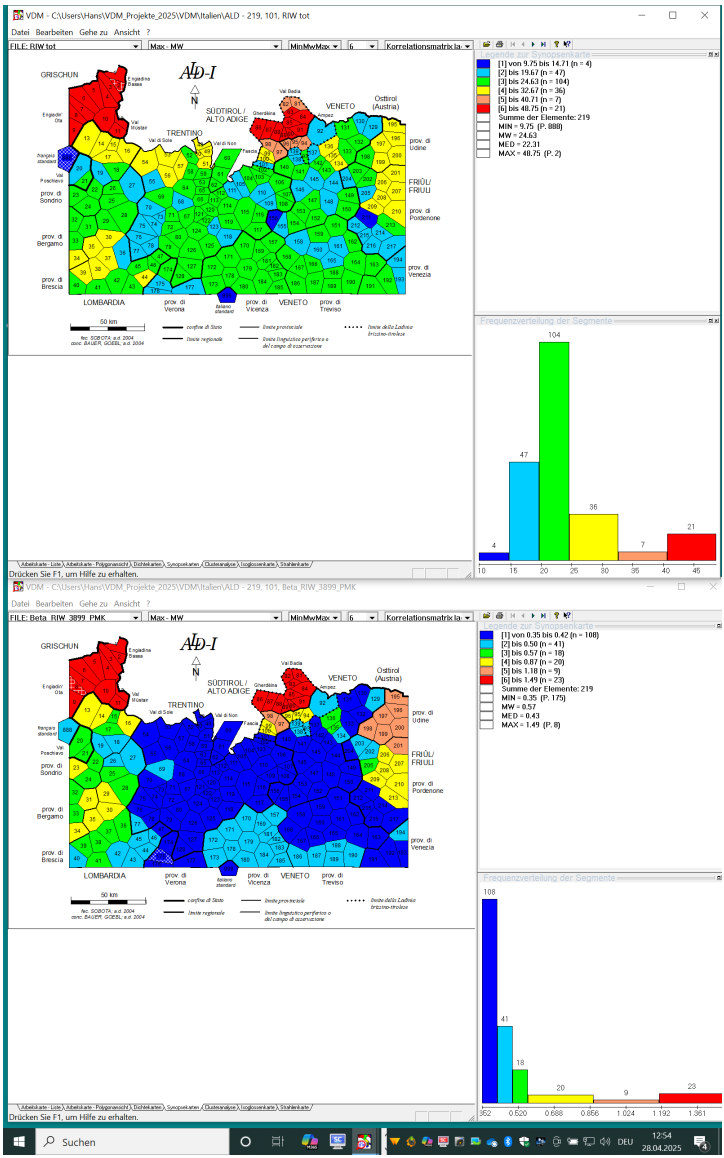
Karte 2 – Zwei Ähnlichkeitskarten zum ALD-I-Punkt 81 (La Pli/Enneberg, Südtirol)
Oben (*Alpha*-Dialektometrie): Ähnlichkeitsmessung: *RIW* (Relativer Identitätswert)
Unten (*Beta*-Dialektometrie): Ähnlichkeitsmessungen: *RIW* + *PMK*
(Produkt-Moment-Korrelations-Koeffizient [nach Bravais-Pearson])
Visualisierung: Intervallalgorithmus MINMWMAX 6-fach
Korpus: 3.899 Arbeitskarten aus ALD-I (alle linguistischen Kategorien).



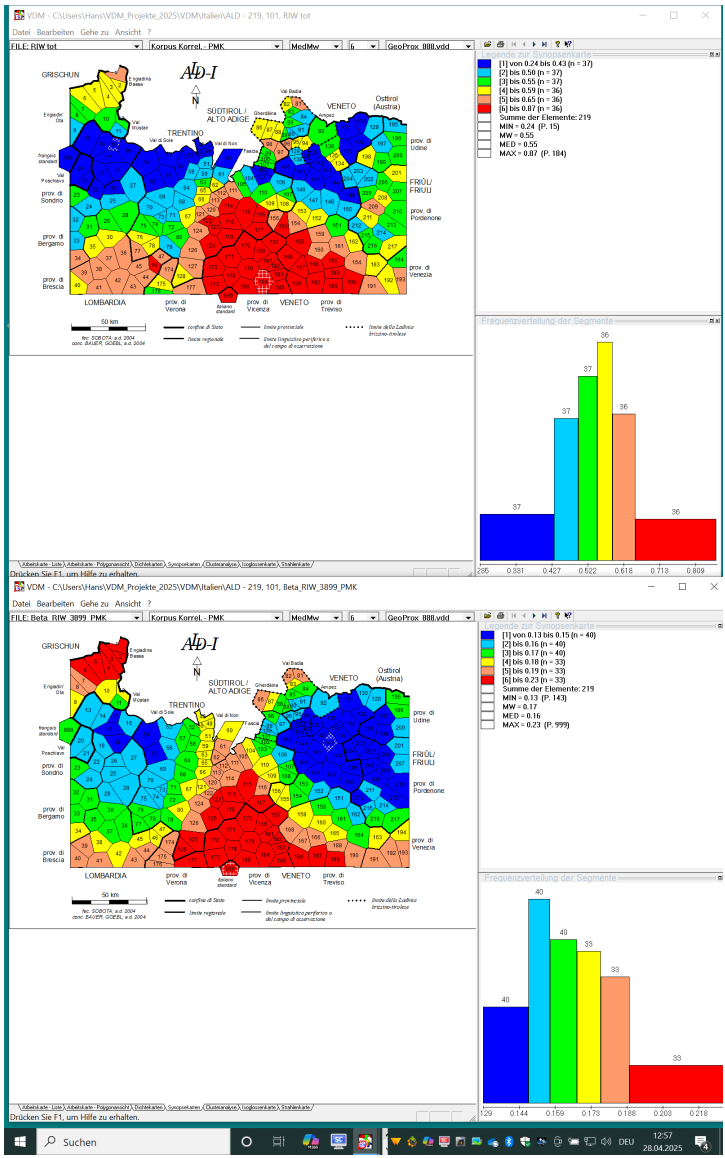
Karte 3 – Zwei Ähnlichkeitskarten zum ALD-I-Punkt 195 (Forni Avoltri, Friaul)
Oben (*Alpha*-Dialektometrie): Ähnlichkeitsmessung: *RIW* (Relativer Identitätswert)
Unten (*Beta*-Dialektometrie): Ähnlichkeitsmessungen: *RIW* + *PMK*
(Produkt-Moment-Korrelations-Koeffizient [nach Bravais-Pearson])
Visualisierung: Intervallalgorithmus MINMWMAX 6-fach
Korpus: 3.899 Arbeitskarten aus ALD-I (alle linguistischen Kategorien).



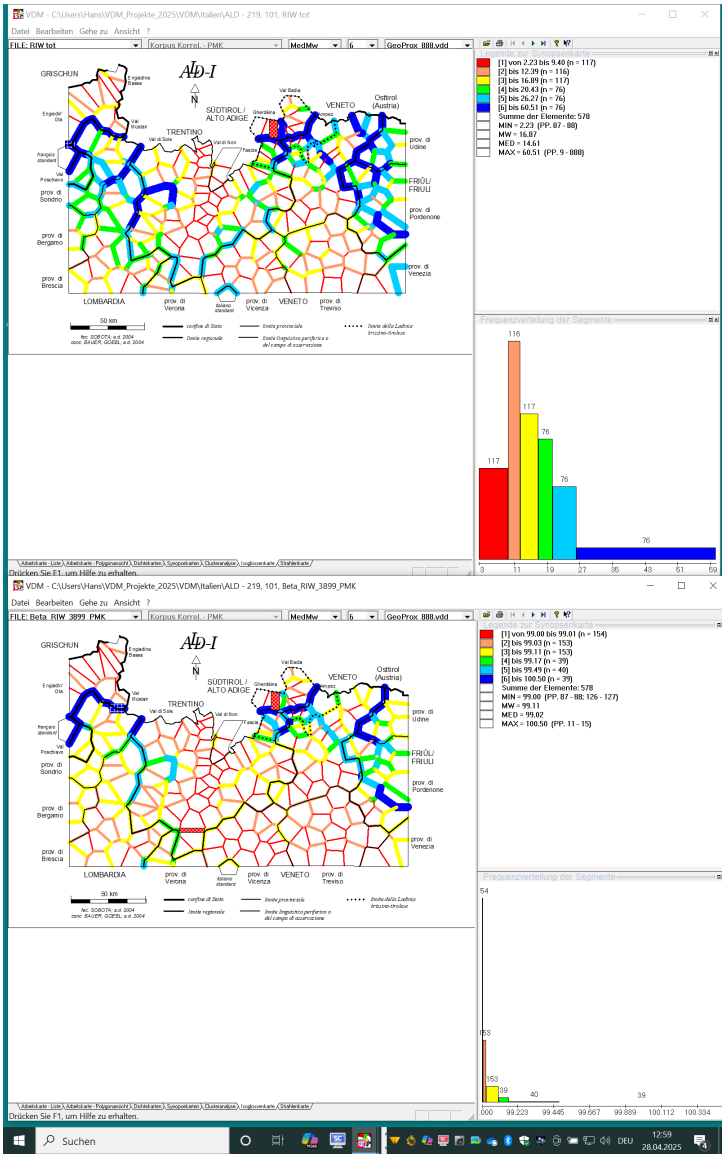
Karte 4 – Zwei Parameterkarten zum ALD-I: Synopse der Schiefen der 219 Ähnlichkeitsverteilungen
Oben (*Alpha*-Dialektometrie): Ähnlichkeitsmessung: *RIW* (Relativer Identitätswert)
Unten (*Beta*-Dialektometrie): Ähnlichkeitsmessungen: *RIW* + *PMK*
(Produkt-Moment-Korrelations-Koeffizient [nach Bravais-Pearson])
Visualisierung: Intervallalgorithmus MEDMW 6-fach
Korpus: 3.899 Arbeitskarten aus ALD-I (alle linguistischen Kategorien).



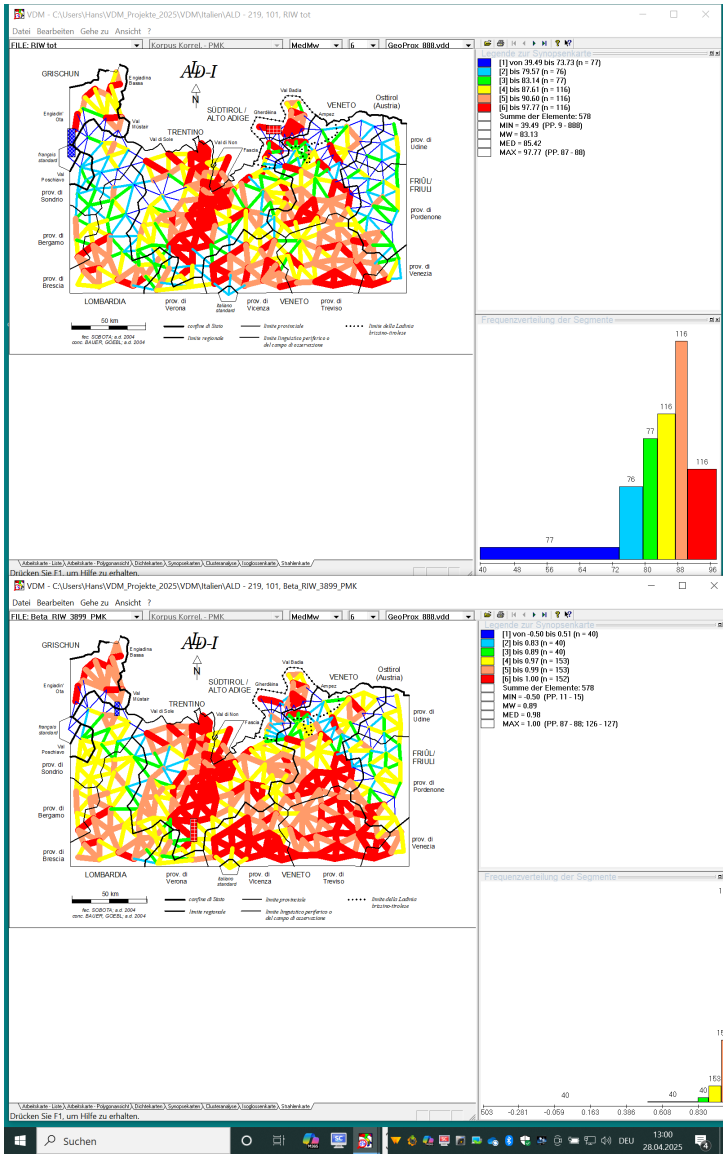
Karte 5 – Zwei Parameterkarten zum ALD-I: Synopse der Differenzen zwischen den *Maxima* (MAX) und den *Mittelwerten* (MW) der 219 Ähnlichkeitsverteilungen
Oben (*Alpha*-Dialektometrie): Ähnlichkeitsmessung: *RIW* (Relativer Identitätswert)
Unten (*Beta*-Dialektometrie): Ähnlichkeitsmessungen: *RIW* + *PMK* (Produkt-Moment-Korrelations-Koeffizient [nach Bravais-Pearson])
Visualisierung: Intervallalgorithmus MINMWMAX 6-fach
Korpus: 3.899 Arbeitskarten aus ALD-I (alle linguistischen Kategorien).



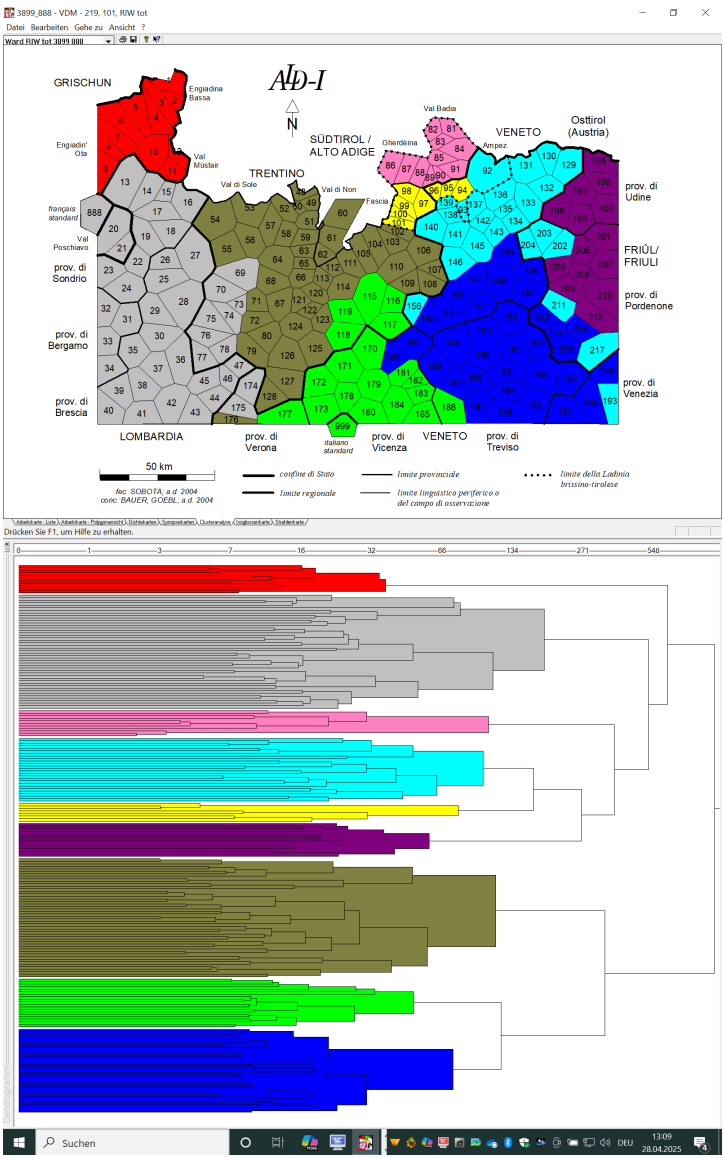
Karte 6 – Zwei Korrelationskarten zum ALD-I: *Ähnlichkeiten* (*RIW* bzw. *RIW* + *PMK*) und *Geographische Proximitäten* (GeoProx)
Oben (*Alpha*-Dialektometrie): Ähnlichkeitsmessung: *RIW* (Relativer Identitätswert)
Unten (*Beta*-Dialektometrie): Ähnlichkeitsmessungen: *RIW* + *PMK* (Produkt-Moment-Korrelations-Koeffizient [nach Bravais-Pearson])
Visualisierung: Intervallalgorithmus MEDMW 6-fach
Korpus: 3.899 Arbeitskarten aus ALD-I (alle linguistischen Kategorien).



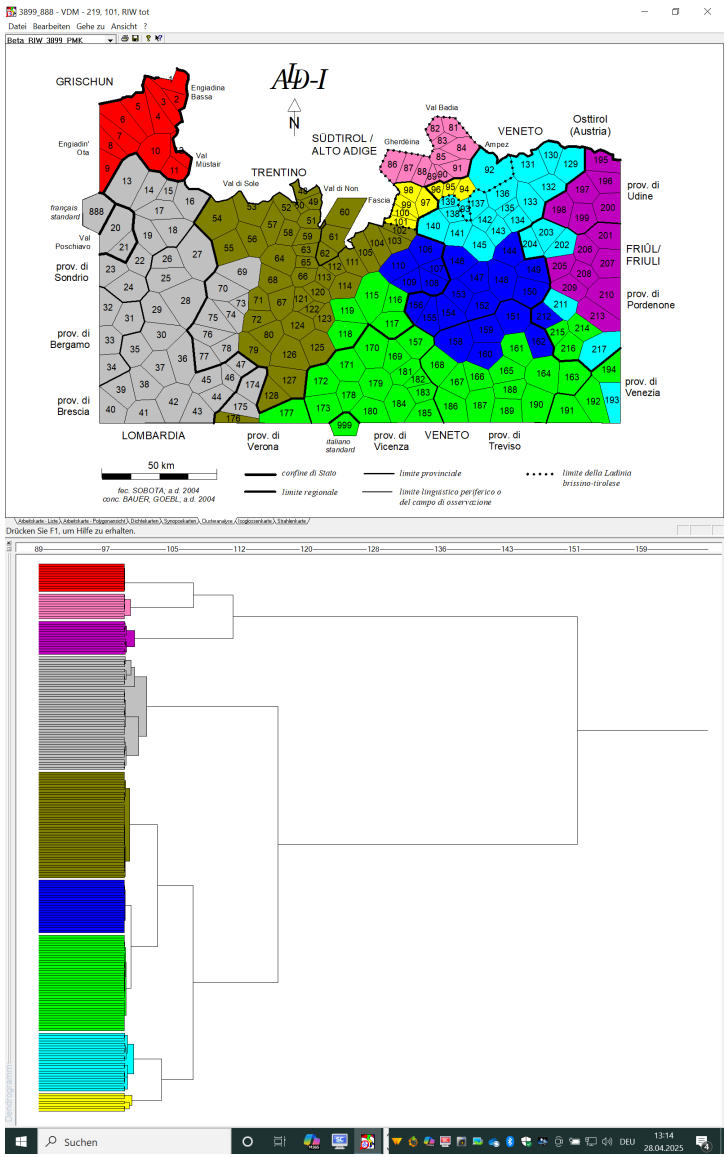
Karte 7 – Zwei Isoglossen-Synthesen zum ALD-I: Messung der interpunktuellen *Distanzen*
Oben (*Alpha*-Dialektometrie): Distanzmessung: *1 - RIW* (Relativer Identitätswert)
Unten (*Beta*-Dialektometrie): Ähnlichkeitsmessungen: *1 - RIW + PMK*
(Produkt-Moment-Korrelations-Koeffizient [nach Bravais-Pearson])
Visualisierung: Intervallalgorithmus MEDMW 6-fach
Korpus: 3.899 Arbeitskarten aus ALD-I (alle linguistischen Kategorien).



Karte 8 – Zwei Strahlen-Karten zum ALD-I: Messung der interpunktuellen Ähnlichkeiten
Oben (*Alpha*-Dialektometrie): Ähnlichkeitsmessung: RIW (Relativer Identitätswert)
Unten (*Beta*-Dialektometrie): Ähnlichkeitsmessungen: *RIW* + *PMK*
(Produkt-Moment-Korrelations-Koeffizient [nach Bravais-Pearson])
Visualisierung: Intervallalgorithmus MEDMW 6-fach
Korpus: 3.899 Arbeitskarten aus ALD-I (alle linguistischen Kategorien).



Karte 9 – Dendrographische Analyse nach der Methode von Joe Ward Jr.
Alpha-Dialektometrie: Ähnlichkeitsmessung: *RIW* (Relativer Identitätswert)
Visualisierung: durch neun in Baum und Karte (analog zu Karte10) ident eingefärbte Dendreme bzw. Choreme
Korpus: 3.899 Arbeitskarten aus ALD-I (alle linguistischen Kategorien).



Karte 10 – Dendrographische Analyse nach der Methode von Joe Ward Jr.
Beta-Dialektometrie: Ähnlichkeitsmessung: *RIW + PMK*
(Produkt-Moment-Korrelations-Koeffizient [nach Bravais-Pearson])
Visualisierung: durch neun in Baum und Karte (analog zu Karte 9) ident eingefärbte Dendreme bzw. Choreme
Korpus: 3.899 Arbeitskarten aus ALD-I (alle linguistischen Kategorien).