

Dolomitenladinische Ähnlichkeitsprofile aus dem Gadertal

Ein Werkstattbericht zur dialektometrischen Analyse des *AD-I*

Roland Bauer

1. Vorstellung des Projekts *AD-DM*

In diesem Beitrag soll von einem Forschungsprojekt die Rede sein, das sich seit Beginn des Jahres 2001 der dialektometrischen Analyse aller 884 im *AD-I*¹ publizierten Sprachatlaskarten widmet. Es handelt sich um das dankenswerterweise vom Wissenschaftsfonds FWF in Wien geförderte Projekt *AD-DM* (Projektnummer P 14566-HIS). Da der dolomitenladinische Sprachatlas *AD-I* den Lesern dieser Zeitschrift u.a. durch die Publikation von 11 Arbeitsberichten² bestens bekannt ist, braucht er wohl hier nicht gesondert vorgestellt zu werden.

1.1. Problemstellung

Sprachatlantenn (wie der *AD-I*) sind matrixartig organisierte, geolinguistische Datensammlungen, in welchen die dialektale/basilektale Befindlichkeit einer untersuchten Landschaft im Sinn eines synchronen Schnittes dokumentiert wird. Dabei werden auf einer Anzahl p einzelner Sprachatlaskarten mit einer Anzahl N im Raum verteilter Messpunkte im Feld erhobene dialektale Reflexe zu vorgegebenen Konzepten abgebildet. In der Romanistik wird die Abbildung des innerlin-

¹ *AD-I* = GOEBL/BAUER/HAIMERL (eds.) 1998; alle in diesem Beitrag verwendeten Abkürzungen werden in Kapitel 4 aufgelöst.

² Cf. AAVV 1986–1999.

guistischen Datenmaterials traditionellerweise über Volltextkarten in Form integraler phonetischer Transkriptionen bewerkstelligt, so dass in der Regel allen an den Dialektdaten Interessierten sämtliche Originaldaten zur Verfügung stehen³. Einzelne Sprachatlaskarten können nach bestimmten innerlinguistischen (i.e. phonetischen, lexikalischen u.a.) Kriterien ausgewertet oder (wie wir in der Fachsprache sagen) “taxiert” werden, um allfällige, hinter den Einzeldaten liegende Raumstrukturen herauszuarbeiten. In diesem Zusammenhang werden aus der Vielfalt unterschiedlicher Antworten pro Atlaskarte einige wenige Antworttypen/Klassen abstrahiert, die jeweils eine, für die Einzeltaxierung gewählte Merkmalsausprägung gemeinsam haben. Diese Taxierung geschieht meist dadurch, dass der Datenbefund pro Messpunkt aus der Karte gelesen, in einer Taxatliste kodiert und dann zu Visualisierungszwecken auf eine “stumme Karte”⁴ übertragen wird. Durch Zusammenführen der typologisch gleichartigen Taxierungserträge mehrerer solcher Karten auf so genannten *Dichtekarten* kann die Verteilung einzelner linguistischer Merkmale (bzw. der entsprechenden Merkmalsausprägungen) im Raum sichtbar gemacht werden. Dichtekarten – darauf muss schon an dieser Stelle besonders hingewiesen werden – haben a priori nichts mit Dialektometrie zu tun, da in diesem Zusammenhang ein wesentliches Element der dialektometrischen (= taxometrischen) Verfahrenskette, nämlich jenes der Ähnlichkeits- bzw. Distanzmessung, gänzlich fehlt⁵.

Die in Sprachatlanten vorhandene zweidimensionale Datenmatrix (N Messpunkte mal p Karten) sollte aber – im Gegensatz zur oben angesprochenen Auswertung und Verdichtung weniger Einzelkarten – vor allem dazu herangezogen werden, höherrangige *Ordnungsstrukturen, Muster und Symmetrien* im von der jeweils

³ Anders verfährt etwa die Germanistik, wo wegen der oft exorbitant großen Anzahl von Messpunkten in den Sprachatlanten meist keine Rohdaten, sondern nur “verdünnte” Forschungsergebnisse in der Form mittels Symbolverkodung erstellter Karten publiziert werden. Dies bringt jedoch erhebliche Nachteile hinsichtlich Überprüfbarkeit und zusätzlicher Datenauswertung mit sich (cf. APPEL 2002, 13). Besonders hartnäckige germanistische Dialektologen scheinen aber trotz dieser evidenten methodischen Defizite die (aus romanistischer Evidenz jeglicher Sachgrundlage entbehrende und somit unhaltbare) Meinung vertreten zu müssen, dass eine Dokumentation “nach Manier der romanischen Sprachatlanten [...] eine optische Verdichtung zu sprachlichen Raumbildungen [...] unmöglich macht” (DINGELDEIN 2002, 103). Zu einer genaueren Auseinandersetzung mit den methodischen Unterschieden zwischen germanistischer und romanistischer Sprachatlasarbeit cf. GOEBL 2002b.

⁴ Cf. dazu bereits JABERG 1906.

⁵ Cf. in diesem Zusammenhang die Methoden verkennende Fehlinterpretation von KREFELD 1994 (die leider auch von anderen Forschern unreflektiert übernommen wurde und wird, cf. LIVER 1999, 25–26 und KRAMER 2000, 47) sowie darauf replizierend bzw. zurechtrückend BERSCHIN 1996. Beispiele für Dichtekarten zum *AD-I* finden sich u.a. in BAUER im Druck, GOEBL 1992, 254–257 oder GOEBL 1999, 199–201.

untersuchten Sprechergemeinschaft dialektal bewirtschafteten Raum zu entdecken und ansprechend zu visualisieren. Dabei muss die Frage, ob solche *Ordnungsmuster* auch tatsächlich existieren, oder ob sie durch die Datenmessung in die Objektwelt projiziert werden, offen bleiben⁶. Die im Rahmen eines solcherart auch auf die Erkennung von *Typen* angelegten Prozesses⁷ gewonnenen Erkenntnisse können schließlich zur Erstellung neuer bzw. zur Verifizierung/Falsifizierung altbekannter Hypothesen herangezogen, also interpretiert und traditionellen Forschungsergebnissen zur Seite bzw. gegenübergestellt werden. Aufgrund der in der Regel sehr großen zu verarbeitenden Datenmengen (idealiter aller in einem Sprachatlas zur Verfügung stehenden Karten) können die o.a. Arbeitsvorgänge jedoch nur mit massiver Computerunterstützung durchgeführt werden. Der im Zusammenhang mit einer solchen, exhaustiven Auswertemethodologie angesprochene Klassifikationsvorgang besteht im Wesentlichen aus drei grundlegenden Arbeitsschritten bzw. Arbeitszielen, nämlich der Datenreduktion (= Selektion, Abstraktion, Verdichtung, hier: durch Taxierung der Atlaskarten), der Ähnlichkeits- bzw. Distanzmessung und der die Mustererkennung ermöglichenden Klassenbildung. Seit den 70er-Jahren des letzten Jahrhunderts spricht man in diesem Kontext von *Dialektometrie* als einer Spielart von *Numerischer Klassifikation* bzw. *Pattern Recognition / Mustererkennung*.

Mit der Ende 1998 erfolgten Publikation des *AD-I*, eines sprachkartographischen Werks zum südost-schweizerischen und nordost-oberitalienischen bzw. zentralrätoromanischen Dialektraum, liegt nun ein aktuelles sprachgeographisches Corpus vor, dessen dialektometrische Auswertung im Rahmen des Projekts *AD-DM* umgesetzt wird. Das Untersuchungsnetz des *AD-I* umfasst 217 Messpunkte, an denen zu rund 1.500 Konzepten basilektale Reflexe erhoben und in phonetischer Umschrift in den Fragebüchern notiert wurden. Alle Explorationsgespräche wurden zudem auf Band aufgezeichnet (und, was die 21 dolomitenladinischen Messpunkte betrifft, digitalisiert), so dass neben dem transkriptorischen auch ein akustisches "Kontrollarchiv" zur Verfügung steht. Der in sieben Bänden (vier Kartenbände, drei Indexbände) auf rund 3.600 Druckseiten publizierte *AD-I* liegt auch als elektronische Version auf drei CD-ROM bzw. auf einer DVD vor⁸.

⁶ Stichwort: *Universalienstreit*; cf. GOEBL 1980, 33–34.

⁷ *Typologie* kann dabei wie folgt definiert werden: "A typology is a statement that a large number of variables have only a small number of combinations of values which actually occur, with all other combinations being rare or nonexistent" (STINCHCOMBE 1968, 47). Zu einer Einführung in die Prinzipien der Sprachtypologie cf. ALTMANN/LEHFELDT 1973.

⁸ Cf. BAUER/GOEBL/HAIMERL 1999–2000 und 2002. Die elektronische Edition kann gegen einen bescheidenen Beitrag zu den Herstellungs- und Versandkosten via Internet bestellt werden (http://www.sbg.ac.at/rom/people/proj/ald/ald_home.htm).

1.2. Projektziele

Das Projekt *AD-DM* verfolgt als generelles Ziel die Anwendung der bislang im Bereich dialektometrischer Forschung eingesetzten Methoden und Techniken auf das Datencorpus des *AD-I* sowie deren Verfeinerung und Ausbau. In diesem Zusammenhang ergeben sich innerhalb unseres Projekts zunächst genuin geolinguistisch-dialektologisch induzierte Teilziele, und zwar: Auffindung möglichst hochrangiger Ordnungsstrukturen durch umfassende Sondierung des oberitalienisch-rätoromanischen Diasystems auf Basis aller *AD-I*-Daten, i.e. unter Auswertung des größten, für diesen Raum bislang zur Verfügung stehenden Datencorpus; Einbau der daraus resultierenden dialektometrischen Forschungserträge in inner- wie außer- bzw. metalinguistisch relevante Erkenntnisse/Theorien der "traditionellen", i.e. (meist) rein qualitativ ausgerichteten Dialektologie und Soziolinguistik; Dokumentation der Kompatibilität dialektometrischer und "traditioneller" Forschungsmethoden; Überprüfung, Stützung und kritische Hinterfragung der traditionellen Ergebnisse/Theorien im Lichte der Dialektometrie. Für den Bereich der Computerlinguistik sind ferner folgende Teilziele bzw. daraus erwachsende Synergien zu nennen: Erweiterung des Vorrats an dialektometrischen Methoden und Verfahren; deren Testung und Implementierung in marktüblichen EDV-Strukturen (u.a. durch Portierung auf die gängigsten netzwerkfähigen Betriebssysteme); Vorschlag möglicher Standards im Bereich computerunterstützter Dialektometrie zur Anwendung auf weitere romanistische, germanistische, anglistische u.a. Datencorpora (interdisziplinärer Impuls)⁹.

1.3. Projektstand

Derzeit (Anfang 2003) verfügt das Projekt *AD-DM* über gut 3.300 taxierte Arbeitskarten. Unter einer Arbeitskarte (AK) verstehen wir in diesem Zusammenhang die durch Taxierung einer originalen Sprachatlaskarte (OK) nach einem bestimmten Merkmal generierte Einzelauswertung. Da jede Sprachatlaskarte nach verschiedenen Merkmalen untersucht werden kann¹⁰, übersteigt die Anzahl der AK jene der publizierten OK bei weitem. Im Projekt *AD-DM* gehen wir diesbezüglich von einer durchschnittlichen Relation von 1 OK : 4,5 AK aus. Im Endausbau wird unser Datencorpus daher rund 3.500–4.000 AK aufweisen.

⁹ Cf. dazu eine Übersicht der bisher dialektometrisch bearbeiteten (französischen und italienischen) Corpora unter <http://ald.sbg.ac.at/dm>.

¹⁰ Die *AD-I*-Karte 261 *l'estate* (< lat. AESTÁTE) konnte z.B. nach vier phonetischen Kriterien taxiert werden, woraus folgende vier Arbeitskarten entstanden: AK 1: anlautendes AE-]; AK 2: vortoniges S impurum; AK 3: Hauptton Á]; AK 4: nachtoniges T.

Die im Projekt für die Taxierung verantwortlichen Mitarbeiter¹¹ stützen sich bei ihrer Arbeit auf ein Kompendium so genannter *Taxierungsrichtlinien*, wobei folgende Hauptkriterien zur Anwendung kommen:

1. Im Falle von (in Sprachatlanten naturgemäß nicht ganz vermeidbaren) *Datenlücken* werden nur jene OK taxatorisch vermessen, die zu mindestens 210 von insgesamt 217 Messpunkten Auskunft geben; es werden also maximal sieben Datenlücken toleriert.
2. Bei Auftreten von (im *AD-I* eher seltenen) *Mehrfachantworten* wird ausschließlich die Erstantwort berücksichtigt.
3. Jede beobachtbare Merkmalsausprägung soll an mindestens drei Messpunkten auftreten, um eine gewisse *Minimalvariation* zu gewährleisten.
4. Es wird zwischen zwei Grundtypen der Taxierung unterschieden, nämlich einerseits der so genannten *Mikro-Taxierung*, bei der z.B. die Entwicklung eines bestimmten lateinischen Nexus *im Detail* vermessen wird¹², und andererseits der *Makro-Taxierung*, bei der meist *binär*, also eher *grob* vorgegangen wird. In solchen Fällen geht es etwa um Präsenz vs. Absenz verschiedener Lautwandel-Phänomene (z.B. Diphthongierung, Epenthese, Palatalisierung, Rhotazismus usw.), bestimmter lexikalischer Makro-Merkmale (z.B. Zugehörigkeit der Belege zu einer Wortfamilie, Auftreten von Germanismen) oder aber morphologischer Typen (z.B. vokalischer vs. sigmatischer vs. palataler vs. unmarkierter Plural, einfache vs. doppelte Stellung des Subjektspronomens)¹³.
5. Jedes für die Taxierung herangezogene Merkmal wird über einen eindeutigen *Merkmals-Code* in der Datenbank festgehalten¹⁴. Zudem werden innerlingu-

¹¹ Es sind dies: Irmgard DAUTERMANN, Liza KLINGER, Susanne OLEINEK, Brigitte RÜHLINGER und Pavel SMEČKA. Ihnen allen sowie den im EDV-Bereich agierenden Mitarbeitern Edgar HAIMERL und Slawomir SOBOTA sei an dieser Stelle herzlichst für das bisher Geleistete gedankt!

¹² Hier irrt Johannes KRAMER, wenn er in seiner Besprechung zum *AD-I* (unter völliger Verkennung dialektometrischen Arbeitens) u.a. schreibt: “[...] auf eine «narrow transcription» der Formen kann man auch verzichten, weil man ja sowieso mit Worttypen arbeitet, bei denen Feinheiten sowieso [sic] unberücksichtigt bleiben” (2001, 170). Im Projekt *AD-DM* nehmen rund 75% aller Taxierungen explizit auf eine «narrow transcription» Bezug, während nur rund 15% aller Arbeitskarten bestimmten “Worttypen” gewidmet sind. Zu einer weniger polemischen Rezeption des *AD-I* cf. die Besprechungen von BERNHARD 2000, D’AGOSTINO/PENNISI 2000, FRAU 1999, IANNACARO 2000, KELLE 2001, SKUBIC 2000, SOLER 2002 oder VICARIO 1999. Eine Gesamtaufstellung aller bislang zum *AD-I* erschienenen Rezensionen findet sich im Internet unter http://www.sbg.ac.at/rom/people/proj/ald/ald_home.htm.

¹³ Konkrete Anwendungen phonetischer Mikro- und Makro-Taxierungen (anhand der *AD-I*-Karte 601 *Pietro*) finden sich in BAUER 2002, 81–87.

¹⁴ Ein etwa aus bairischem bzw. alemannischem Adstrat stammender, somit relativ junger Germanismus trägt *immer* den Code 015; haupttoniges Á[(in offener Stellung) wird *immer* mit 330, der vortönige Konsonantennexus GL *immer* mit 527 kodiert usw.

istisch affine Merkmalstypen (bzw. die entsprechenden Codes) zu *Merkmalsgruppen* zusammengefasst, um im Rahmen der Datenverrechnung auch auf bestimmte Teilcorpora – etwa Phonetik vs. Lexikon – zugreifen zu können.

In quantitativer Hinsicht ist unser aktuelles Corpus wie folgt strukturiert: (gerundete) 33% aller Taxierungen betreffen den Vokalismus, 29% den Konsonantismus, 15% das Lexikon und 23% die oben erwähnten Makro-Merkmale. Der aus dieser Verteilung ersichtliche numerische Überhang phonetisch relevanter Information ergibt sich aus dem Grundkonzept des *AD*: in der bereits abgeschlossenen Projektphase *AD-I* stand die phonetische Fragestellung im Vordergrund, während in der derzeit in Ausarbeitung befindlichen Projektphase *AD-II* gezielt lexikalische, morphologische und syntaktische Dialektdaten gesammelt werden¹⁵. Zu jeder der vier oben genannten Gruppen sind ferner mindestens 500 Arbeitskarten vorhanden, eine Quantität, die – wie entsprechende Tests bzw. Erfahrungen mit anderen dialektometrischen Corpora zeigen¹⁶ – bereits sehr stabile Verrechnungsergebnisse gewährleistet. Dies kann auch anhand des Vergleichs zweier, zum selben Prüfbezugspunkt erstellter Ähnlichkeitskarten leicht nachvollzogen werden. In BAUER (2002, 98) wurde ein erstes Ähnlichkeitsprofil zum Kunstpunkt 999 “italiano standard” publiziert¹⁷, bei dem 527 Arbeitskarten verrechnet wurden. Der direkte Vergleich mit dem hier auf Karte 1 abgebildeten Profil, das sich auf die sechsfache (!) Anzahl an Arbeitskarten stützt (hier: 3.302 AK, dort: 527 AK), zeigt jedoch annähernd identische Raumkammerungen. Die Validität eines Datencorpus mit mindestens 500 AK ist also eindeutig gegeben. Freilich können in solchen Fällen kaum dialektometrische Auswertungen gemacht werden, die sich auf Teilcorpora stützen, da dabei die Zellenhäufigkeit rasch unter eine kritische Größe (von etwa 200 AK) absinken würde. Dies ist bei der Mächtigkeit unseres hier vorgestellten Corpus nicht der Fall, gibt es doch selbst innerhalb der o.a. Großgruppen (Vokalismus, Konsonantismus usw.) noch eine Reihe von Merkmalen, deren Zellenhäufigkeit eine gesonderte Auswertung durchaus rechtfertigen würde: allein der vokalische Nachton und Auslaut kann auf knapp 500 AK nach-

¹⁵ Natürlich wäre es wünschenswert, die dialektometrische Analyse des *AD* nach der Publikation von *AD-II* durch den Einbezug großer lexikalischer, morphologischer und syntaktischer Datenmengen zu vervollständigen. Da dies jedoch frühestens ab dem Jahr 2010 der Fall sein wird, können dazu aus heutiger Sicht lediglich unverbindliche (wenn auch durchaus ernst gemeinte) Absichtserklärungen abgegeben werden.

¹⁶ Zur Verwendung so genannter Staffelforpora mit je 500 AK cf. GOEBL 1984, I, 71–73.

¹⁷ Leider wurde die Karte auf Grund eines Fehlers im Druckprozess nicht in Farbe, sondern nur in einer kaum lesbaren Schwarz-Weiß-Reproduktion abgebildet, weshalb jedem ausgelieferten Band ein Separatum mit Farbkarten beiliegen sollte. Dieses kann überdies jederzeit kostenlos beim Autor bezogen werden (Roland.Bauer@sbg.ac.at).

vollzogen werden, der Konsonantismus im Vorton ist durch 280 Einzeltaxierungen abgedeckt, die Pluralbildung von Adjektiven und Substantiven ist mit über 300 AK vertreten¹⁸, usw.

2. Zur Dialektometrie

Der im Jahr 1973 durch den südfranzösischen Romanisten und Autor des gaskognischen Sprachatlases ALG, Jean SÉGUY, geschaffene Fachterminus *dialektométrie* bzw. die darunter zu verstehende, aus heutiger Sicht zur quantitativen Linguistik zählende Disziplin der *Dialektometrie* (DM) kann, ganz elementar gesprochen, als Kombination aus (per definitionem mit Sprachatlanten befassender) *Sprachgeographie* und *numerischer Klassifikation* (auch *Taxometrie* genannt) angesehen werden¹⁹. So gesehen findet sich die DM, neben rezenten linguistischen Anwendungen wie etwa der Onomatometrie²⁰, in guter Gesellschaft einer ganzen Reihe weiterer messender Disziplinen (Soziometrie, Psychometrie, Ökonometrie usw.), die allesamt mehrere Dinge gemeinsam haben: eine aus N Objekten (hier: Messpunkten bzw. Ortsdialekten) und p Attributen (hier: Arbeitskarten bzw. sprachlichen Merkmalen) zusammengesetzte *Datenbasis*, den Einsatz *Daten reduzierender*, somit abstrahierender, synthetisierender bzw. simplifizierender *Verfahren* (die, etwa im Vergleich mit dem vollen Informationsgehalt einer originalen Sprachatlaskarte, naturgemäß einen gewissen Informationsverlust mit sich bringen) und das Forschungsziel, durch Klassifikation der jeweiligen Daten *Ordnungsstrukturen, Muster und Typen* aufzudecken, die ohne den Einsatz numerisch-mathematischer Verrechnungsverfahren nicht sichtbar gemacht werden könnten. Die methodische Weiterentwicklung und Perfektionierung der Disziplin DM lag seit Ende der 70-er Jahre des vergangenen Jahrhunderts v.a. in den Händen von Hans GOEBL, der in diesem Zusammenhang besonders auf den Kontakt (heutiger) Dialektometer zur weltweit etablierten "Szene" der numerischen Klassifikation hinweist, der SÉGUY und seinen direkten Nachfolgern noch verwehrt geblieben war²¹. Nahezu alle hier in der Folge präsentierten Verrechnungs- und Visualisierungsverfahren beruhen im Übrigen auf den bereits in GOEBL (1984) ausführlich erläuterten Prinzipien.

¹⁸ Siehe dazu Karte 7.

¹⁹ Cf. den Titel des Beitrags von GOEBL 1980 sowie GOEBL 1981, 349. Eine Einführung in die numerische Klassifikation bieten SNEATH/SOKAL 1973.

²⁰ Cf. dazu den Beitrag von Paul VIDESOTT in diesem Band.

²¹ Cf. dazu zuletzt GOEBL 2001, 857.

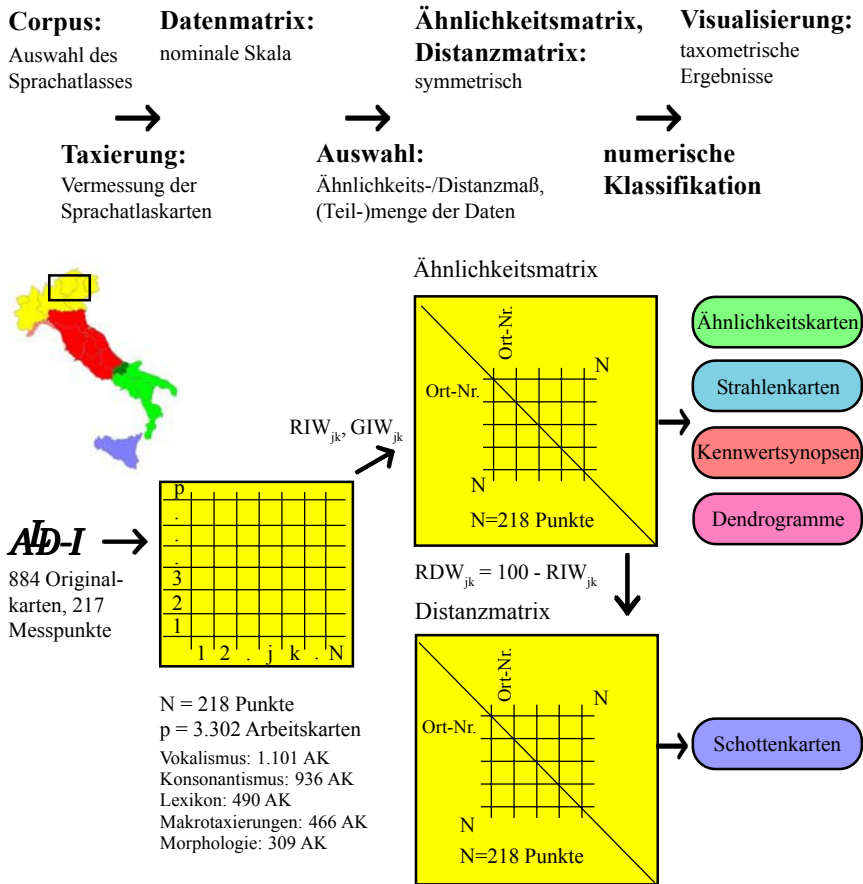


Abb. 1: Die dialektometrische Verfahrenskette

RIW_{jk} = „Relativer Identitätswert“
 RDW_{jk} = „Relativer Distanzwert“
 GIW_{jk} = „Gewichtender Identitätswert“

2.1. Die dialektometrische Verfahrenskette²²

Die in einem Sprachatlas abgebildete, mittels empirischer Feldforschung erhobene geolinguistische Realität stellt, wie bereits ausgeführt wurde, die Ausgangsbasis dialektometrischen Arbeitens dar. Durch die Vermessung (= Taxierung) der Sprachatlaskarten entsteht eine *Datenmatrix*, die aus N explorierten Ortschaften und p Arbeitskarten besteht. In unserem Fall handelt es sich dabei um 217 genuine *AD*-Messpunkte und um einen Zusatzpunkt (Nr. 999), der das Standarditalienische repräsentiert ($N = 218$). Die Einbeziehung dieses Kunstpunkts in unser Beobachtungsnetz erlaubt die Visualisierung dialektaler Ähnlichkeiten und/oder Distanzen aller Messpunkte mit dem Italienischen und umgekehrt. Der Wert p (Anzahl der Attribute) beträgt im Fall des hier vorgestellten Corpus 3.302 Arbeitskarten, woraus sich ein Produkt von 719.836 Messwerten (errechnet aus $N * p = 218 * 3.302$) ergibt, die derzeit in der Datenmatrix gespeichert sind.

Eigentlich beginnt (dialekto-)metrisches Arbeiten erst mit dem zweiten Schritt unserer Verfahrenskette, nämlich der Umwandlung der auf nominalem Skalenniveau befindlichen Datenmatrix in eine symmetrische *Ähnlichkeits-* oder *Distanzmatrix*. Zu diesem Zweck wird aus einem Fundus von bereits erprobten Maßen ein entsprechendes Ähnlichkeits- oder Distanzmaß gewählt. Dabei wird u.a. auf den so genannten *Relativen Identitätswert* (abgek. RIW_{jk}) zurückgegriffen, der auch zur Generierung der hier vorzustellenden Ähnlichkeitsprofile herangezogen wurde. RIW_{jk} bedeutet, dass – als Modellierung einer bereits vorab in den Köpfen der Fachwissenschaftler (und Dialektometer) vorhandenen Vision/Theorie/Vorstellung von “dialektaler Ähnlichkeit” – ausgehend von den in der Datenmatrix gespeicherten Taxaten die Vektoren zweier Messpunkte (j und k) miteinander verglichen werden, um die jeweilige Anzahl an übereinstimmenden Einträgen (Ko-Identitäten) und an divergierenden Messwerten (Ko-Differenzen) festzuhalten. Dabei haben alle Messwerte das absolut gleiche Gewicht. Nullstellen, also fehlende Taxateinträge in der Datenmatrix, werden nicht in den paarweisen Vergleich miteinbezogen²³. Dieses Procedere wird auf sämtliche Messpunktepaare angewendet, so dass schlussendlich (im Falle des *AD-I*) alle 218 Punkte mit den verbleibenden 217 Punkten auf ihre reziproke Ähnlichkeit hin überprüft werden. Die Hälfte aller (in 218 mal 217 Vergleichen) berechneten 47.306 Ähnlichkeitswerte kann aus Symmetriegründen ausgeklammert werden, da ein Vergleich des

²² Siehe dazu Abbildung 1.

²³ Nur so kann nämlich sichergestellt werden, dass zwei- oder mehrmaliges Auftreten von Nullstellen innerhalb eines Merkmalsvektors beim paarweisen Vergleich der Ortsvektoren nicht als gleichartige Merkmalsausprägung (d.h. als 100%ige Ähnlichkeit: $0 = 0$) interpretiert wird.

Punktes j mit dem Punkt k numerisch vollauf dem umgekehrten Vergleich k zu j entspricht. Die Ähnlichkeitsmatrix zum *AD-I* enthält somit 23.653 Einträge (= $47.306 / 2$); die entsprechende Formel dazu lautet: $N/2 * (N-1)$. Der gewählte Ähnlichkeitswert selbst, der immer zwischen 0 (= keine Ähnlichkeit) und 100 (= maximale Ähnlichkeit) liegt, berechnet sich nach einer einfachen Formel²⁴, indem die Gesamtsumme der festgestellten Ko-Identitäten zunächst durch die Summe aus Ko-Identitäten und Ko-Differenzen dividiert und das Ergebnis mit 100 multipliziert wird. Dazu ein Beispiel: hätten unsere Messpunkte j und k , unter Ansehung des Gesamtcorpus von 3.300 Arbeitskarten, 2.640 übereinstimmende Merkmalsausprägungen und demnach 660 divergierende Taxate, so ergäbe unsere Formel einen RIW_{jk} von 80, errechnet aus $(2.640 / 2.640 + 660) * 100$. In die Gemeinsprache übersetzt würde $RIW_{jk} = 80$ nun bedeuten, dass zwischen den Dialekten der Messpunkte j und k eine Ähnlichkeit von 80% besteht, da in beiden Dialekten 80% der beobachteten Merkmalsausprägungen (Taxate) übereinstimmen bzw. identisch sind.

Die graphische Umsetzung der in der Ähnlichkeitsmatrix gespeicherten Werte in der Form adäquater Visualisierungen stellt den letzten dialektometrischen Arbeitsschritt dar, von dem hier die Rede sein soll. Aus einer Vielzahl zur Verfügung stehender Optionen haben wir jene der *Ähnlichkeitskarten* oder *Ähnlichkeitsprofile* gewählt. Bei dieser Art der kartographischen Umsetzung wird ein einziger Messpunkt j als so genannter Prüfbezugspunkt gewählt und mit allen übrigen Punkten verglichen. Dadurch wird, was die Berücksichtigung der Ähnlichkeitswerte betrifft, nur auf eine einzige Zeile der Ähnlichkeitsmatrix (i.e. auf den Messpunktvektor j) zurückgegriffen. Alle in dieser Zeile gespeicherten Messwerte (RIW) werden mit Hilfe eines bestimmten *Segmentieralgorithmus* verschiedenen Klassen zugeordnet. Eine gängige Art der Segmentierung (MinMwMax) orientiert sich am jeweils kleinsten, am durchschnittlichen und am größten RIW. Dabei wird sowohl zwischen dem Minimalwert und dem arithmetischen Mittelwert, als auch zwischen dem Mittelwert und dem Maximalwert eine gleiche Anzahl jeweils gleich breiter Klassen oder Intervalle gebildet. Andere Segmentieralgorithmen (etwa MedMw oder Med) nehmen auf den Median Bezug, wodurch die einzelnen Klassen die gleiche Anzahl von Elementen (Messpunkten) aufweisen, aber unterschiedlich breit sind. Diese Typen der Segmentierung erzeugen einen geringeren Bildkontrast bzw. einen unruhigeren Flächeneindruck der Ähnlichkeitskarten als der hier verwendete MinMwMax, ihr Einsatz empfiehlt sich jedoch bei der Visualisierung höherrangiger Ordnungsstrukturen, bei denen die gesamte

²⁴ Cf. GOEBL 1984, I, 76.

Ähnlichkeitsmatrix zur Erstellung von synoptischen Darstellungen verschiedener Lageparameter der N Ähnlichkeitsverteilungen herangezogen wird²⁵.

Jeder im Rahmen der Segmentierung generierten Wertklasse, d.h. jedem Intervall, wird schließlich eine Farbstufe aus dem Sonnenspektrum zugeordnet, wobei überdurchschnittlich große Ähnlichkeitswerte mit warmen Farben (rot, orange, gelb) und unterdurchschnittliche Werte mit kalten Farben (grün, blau) signiert werden. Der für das jeweilige Ähnlichkeitsprofil gewählte Prüfbezugspunkt wird immer in weißer Farbe dargestellt.

Zur kartographischen Umsetzung (in unserem Fall: der in verschiedene Intervalle segmentierten Ähnlichkeitswerte) wird im Rahmen dialektometrischer Visualisierungen prinzipiell auf eine polygonisierte Grundkarte (Flächenmosaikkarte oder Choroplethenkarte) zurückgegriffen. Diese hat (etwa gegenüber der bei germanistischer Sprachatlasarbeit traditionellerweise eingesetzten Punktsymbolkarte) den Vorteil, dass kohärente Flächen signiert werden können, wodurch das räumliche Sehen bzw. das Erkennen räumlicher Strukturen maßgeblich erleichtert wird. Die Polygonisierung des Messpunktenetzes erfolgt übrigens mit Hilfe der so genannten Thiessen-Geometrie, die bereits im Jahr 1898 erstmals für sprachgeographische Belange verwendet wurde²⁶.

2.2. Das Dialektometrie-Programm *VDM*

Die Kurzbezeichnung *VDM* bedeutet *Visual DialectoMetry*. Damit ist ein modernes, auf allen heute üblichen *Microsoft-Windows*-Plattformen lauffähiges Programmpaket gemeint, das von unserem Kollegen Edgar HAIMERL (Ulm) in einem ebenfalls vom Wiener Wissenschaftsfonds FWF geförderten Forschungsprojekt in den Jahren 1998–2000 entwickelt wurde²⁷. Im Rahmen des hier vorgestellten Projekts *ALD-DM* wird *VDM* seit dem Jahr 2001 kontinuierlich weiterentwickelt. Zurzeit verwenden wir die Version 1.2.0.1 (November 2002).

²⁵ Aktuelle Beispiele für Kennwertsynopsen (z.B. der Standardabweichung und der Schiefe) zum französischen Sprachatlas ALF finden sich in GOEBL 2002a, 53–59.

²⁶ Es war der Germanist Karl HAAG (1898), der diese Technik im Rahmen der Sprachgeographie erstmals erprobte, bevor sie für längere Zeit in "sprachwissenschaftliche" Vergessenheit geriet. Die Bezeichnung "Thiessen-Geometrie" verweist hingegen auf den Meteorologen Alfred H. THIESSEN (1911). Das Prinzip der Erzeugung solcher Polygone wird ausführlich in GOEBL 1984, I, 90–92 erklärt.

²⁷ Laufend aktualisierte, detaillierte Angaben zu Aufbau und Einsatz des Programms *VDM* finden sich, in deutscher und englischer Sprache, im Internet unter <http://ald.sbg.ac.at/dm>.

VDM übernimmt zunächst das im Datenbankstandard von *Microsoft Access* erfasste Datencorpus. Diese Konstellation bringt den nicht zu unterschätzenden Vorteil mit sich, dass unsere Daten auch ohne expliziten Rückgriff auf *VDM* erfasst, gelesen, korrigiert und erweitert werden können, und berechtigt ferner zur Hoffnung, dass die Datenbank selbst dank des weltweit verbreiteten bzw. als Standard etablierten Formats auch in künftigen Benutzeroberflächen lesbar sein wird.

Dem Benutzer von *VDM* werden *grosso modo* drei Hauptfunktionen geboten: 1. die Verwaltung der Datenmatrix, 2. die Errechnung von Ähnlichkeits- bzw. Distanzmatrizen und 3. die taxometrische Behandlung letzterer samt Visualisierung.

2.2.1. Verwaltung der Datenmatrix

Die in der Datenmatrix gespeicherten, durch multiple Interpretationen aller Sprachatlaskarten gewonnenen Messdaten können innerhalb des Programms *VDM* in zweifacher Form ediert werden. Variante 1 betrifft die Darstellung der Taxate in Listenform. Dabei wird pro Arbeitskarte eine zweiseitige Liste generiert, die einerseits alle 218 Messpunktnummern und andererseits die Codes der an diesen Messpunkten jeweils beobachteten Merkmalsausprägungen trägt. Eine Legende erklärt dem Benutzer die linguistische Bedeutung der verwendeten Codes. In Ergänzung der einfachen Listenansicht bietet *VDM* die Möglichkeit, eine Kontrolldatei im Format RTF zu exportieren, die vom Layout her gesehen völlig deckungsgleich mit den im Rahmen der Taxierungsarbeit händisch ausgefüllten Taxatlisten ist. Dieser Export-Funktion kommt bei der in unserem Arbeitsablauf obligatorisch vorgesehenen Datenkontrolle bzw. "Fahnenkorrektur" größte Bedeutung zu.

Die zweite Möglichkeit der Verwaltung der Datenmatrix besteht in der kartographischen Umsetzung der Taxatlisten in polygonisierte Arbeitskarten. Dabei wird jeweils eine Zeile der Datenmatrix verwendet, wobei der pro Ort anfallende Messwert als farbiges Polygon in die Karte eingetragen wird. Eine Kartenlegende informiert auch hier über die linguistische Bedeutung der vergebenen Farben. Diese können übrigens durch den Benutzer interaktiv und beliebig verändert werden. Durch Anklicken eines der verfügbaren Messpunkte wird zusätzlich (und nach kurzer Wartezeit) ein Informationsfenster geöffnet, das Angaben zum aktivierten Messpunkt (Nr., Ortsname), zur jeweils bearbeiteten Originalkarte (Kartenummer, Kartentitel) und zur abgebildeten Arbeitskarte (Taxierungskriterium, Nummer, Name und Frequenz des jeweiligen Taxats) bereit stellt²⁸.

²⁸ Ein Beispiel für eine polygonisierte Arbeitskarte (zum Auslaut -U in PĚTRU, *AD-I*-Karte 601 *Pietro* "Peter") findet sich in BAUER 2002, 97.

Eine sehr hilfreiche, erst in der bislang letzten Programmversion realisierte Editions-Funktion von *VDM* betrifft die durch den Benutzer individuell steuerbare Ausgabereihenfolge der Arbeitskarten bzw. der Taxatlisten. Als Default-Einstellung wird dabei eine der Nummerierung der Originalkarten entsprechende Reihung vorgeschlagen, so dass zunächst alle Arbeitskarten zu *AD-I 4 l'aceto*, dann jene zu *AD-I 5 acido/acida* usw. und schließlich jene zur (alphabetisch) letzten Sprachatlaskarte *AD-I 884 lo zolfo* angezeigt werden. Alternativ dazu kann eine aus linguistischer Sicht überaus nützliche Serie aller typologisch gleichen Arbeitskarten (i.e. aller zu ein und demselben Merkmal durchgeführten Taxierungen) über ein entsprechendes Menü genauso einfach abgerufen werden wie etwa eine Ausgabenreihenfolge, die sich nur an bestimmten Merkmalen orientiert. Die Einstellung "Plural" z.B. bewirkt, dass nur Arbeitskarten angezeigt werden, bei denen als Taxierungsmerkmal die Pluralbildung gewählt wurde.

Insgesamt wird durch diese (hier bewusst sehr präzise beschriebenen) Editions-Optionen die in diesem Zusammenhang wohl größtmögliche *Transparenz* des gesamten Datencorpus sichergestellt, kann doch jeder Benutzer bis ins letzte Detail nachvollziehen, auf welche Art und Weise wie viele bzw. welche Sprachatlaskarten taxiert wurden. Dies wiederum gewährleistet eine einfache Verifikation/Falsifikation der Zusammensetzung und der Güte aller erfassten (und später dialektometrisch weiterverarbeiteten) Daten, ein Umstand, der – so bleibt zu hoffen – der v.a. unter italienischen Dialektologen noch immer weit verbreiteten Skepsis gegenüber dialektometrischer Forschung²⁹ einigen Wind aus den Segeln nehmen könnte.

2.2.2. Errechnung von Ähnlichkeits- und Distanzmatrizen

Auch der zweite Schritt der dialektometrischen Verfahrenskette kann in *VDM* auf sehr komfortable Art erledigt werden. Nach dem Laden des gewünschten Untersuchungscorpus (in Form einer *Access*-Datenbank), kann der Benutzer aus derzeit drei zur Verfügung stehenden Ähnlichkeitsmaßen wählen. Es sind dies: der in Kapitel 2.1 näher beschriebene RIW (*Relativer Identitätswert*), der auf die zeilenweise Frequenz der Taxate in der Datenmatrix Bezug nehmende GIW (*Gewichtender Identitätswert*³⁰) und eine für dialektometrische Zwecke adaptierte Version der DEM (*Durchschnittliche Euklidische Metrik*³¹). Ferner kann ange-

²⁹ PENNISI 1991, 162 spricht sogar von einer "demonizzazione" der DM.

³⁰ Cf. dazu Kapitel 2.3.3.3.

³¹ Cf. dazu BOCK 1974, 35–39.

geben werden, ob die zu errechnenden Ähnlichkeiten in Distanzen konvertiert werden sollen oder nicht. Die dritte Option betrifft schließlich die Auswahl des für die Erzeugung der Ähnlichkeitsmatrix relevanten Corpus, wobei hier – neben dem Gesamtkorpus – alle in der Datenbank vorab definierten Teilcorpora zur Verfügung stehen. Letztere können sowohl über die Angabe von regionalen Kriterien (z.B. gesamtes Untersuchungsnetz vs. Friaul vs. Trentino usw.) als auch über das Setzen innerlinguistischer Filter (z.B. alle Merkmale vs. Konsonantismus vs. Lexikon usw.) in beliebigen Kombinationen generiert werden. Alle errechneten Matrizen können direkt in der Projektdatenbank abgelegt oder aber als eigene Dateien extern gespeichert und bei Bedarf erneut geladen werden.

2.2.3. Numerische Klassifikation und Visualisierung der Ergebnisse

Die auch Insider immer wieder beeindruckende Leistungsfähigkeit von *VDM* kann am besten anhand der Vielfalt der im Programm gebotenen Klassifikations- und Visualisierungsoptionen nachvollzogen werden, die wir aus Platzgründen hier natürlich nicht im Einzelnen vorstellen können.

Neben den bereits erwähnten Ähnlichkeitsprofilen, auf die in dieser Präsentation in Form konkreter Beispielkarten noch näher einzugehen sein wird, verweisen wir hier v.a. auf einen reichen Vorrat an *synoptischen Darstellungen*, die (im Gegensatz zu den Ähnlichkeitskarten) die gesamte Ähnlichkeitsmatrix ($N * N$) und dabei u.a. folgende Kennwerte der jeweils N Ähnlichkeitsverteilungen berücksichtigen: Minimum, Mittelwert, Maximum, Standardabweichung, Variationskoeffizient, Schiefe, Kommunikationsgüte, Interaktionsprodukt und Intersextilabstand.

Zudem bietet *VDM* die Erstellung so genannter *Zwischenpunktkarten* an. Dazu gehören die mit Distanzwerten operierenden *Schottenkarten*, wobei in diesem Zusammenhang nicht ganze Polygonflächen, sondern nur die Polygonseiten eingefärbt und in ihrer Strichstärke verändert werden. Dadurch können dialektale Abschottungen zwischen benachbarten Orten in Form quantitativer Isoglossen-Bündelungen visualisiert werden. Komplementär zur Schottenkarte kann in der aktuellen Version von *VDM* auch die so genannte *Strahlenkarte* angezeigt werden, die mit ebenfalls eingefärbten und in der Strichstärke variablen Verbindungsgeraden zwischen jeweils zwei Polygon-Mittelpunkten (= Messpunkten) operiert, wodurch wechselseitige Ähnlichkeitsrelationen zwischen direkt benachbarten Orten veranschaulicht werden können.

Als die in diesem Zusammenhang elaborierteste Form numerischer Klassifikation bietet *VDM* schließlich verschiedene *hierarchisch-agglomerative* Verfahren

der *Clusteranalyse* an. Der englische Terminus *Cluster Analysis* stammt aus der Psychologie und aus der Anthropologie, wo schon in den 30er-Jahren des 20. Jahrhunderts erste Verfahren entwickelt wurden³². Dabei “geht es darum, die Fälle einer Stichprobe so in mehreren Clustern (Gruppen von Fällen) zusammenzufassen, daß sie innerhalb des Clusters möglichst homogen, die Cluster untereinander aber deutlich verschieden sind”³³. Im Zusammenhang mit unseren Daten sind mit den “Fällen einer Stichprobe” die einzelnen Ortsdialekte gemeint, “Cluster” kann demnach im Sinne einer Gruppe oder einer Klasse von Dialekten verstanden werden, wobei jeder Dialekt in ein Cluster fallen *muss* und kein Dialekt in mehr als ein Cluster fallen *darf*. Es handelt sich also um eine *exhaustive* und zugleich *disjunkte* Klassifikation, bei der sich die einzelnen Klassen nicht überschneiden dürfen³⁴. Als Output der *VDM*-Clusteranalysen werden so genannte *Dendrogramme* (hierarchische Stammbäume) sowie die räumliche Verteilung der die jeweiligen Bäume konstituierenden Äste in Form polygonisierter Farbkarten generiert. Anhand der Dendrogramme kann im Einzelnen nachvollzogen werden, auf welchem Fusionsniveau bzw. auf welcher Heterogenitätsstufe welche und wie viele Blätter (einzelne Ortsdialekte) zu Ästen (Dialektgruppen) verschmelzen oder agglomeriert werden. Neben einer rein synchronen (i.e. geolinguistischen) Betrachtungsweise kann ein Stammbaum im Übrigen auch diachron (i.e. lexikostatistisch) interpretiert werden³⁵.

2.3. Dolomitenladinische Ähnlichkeitsprofile

Alle in der Folge vorgestellten Ähnlichkeitsprofile sind nach den in Abschnitt 2.1 erklärten Prinzipien erstellt worden. Die entsprechenden Abbildungen bestehen jeweils aus vier Versatzstücken: 1. aus einer eingefärbten Polygonkarte, in der der jeweilige Prüfbezugspunkt als weißes, mit einem roten Pfeil gekennzeichnetes Polygon aufscheint, 2. aus einer Farblegende, in der die Ähnlichkeitswerte in sechs Intervalle aufgeteilt und zudem die Frequenz *n* der Messpunkte pro Intervall angegeben ist, 3. aus einem Histogramm, das die Frequenz und die numerische Verteilung der Messpunkte auf die Intervalle in zwölfacher Abstufung zeigt und 4. aus einer Textlegende, in der alle bei der Kartengenerierung berücksichtigten Parameter aufgelistet werden.

³² Cf. dazu BAILEY 1994, 34.

³³ BROSIUS 1989, 175.

³⁴ Cf. BOCK 1974, 22. Zum Problem der Überschneidung (“*overlapping clusters*”) cf. ferner BAILEY 1994, 43.

³⁵ Weiterführende Hinweise zum dialektologischen Wert von Dendrogrammen finden sich in GOEBL 1992.

Als Segmentieralgorithmus wurde bei sämtlichen Karten der oben beschriebene MinMwMax verwendet, allerdings in der Variante MMinMwMaxX. Dabei sind der/die jeweils höchste(n) und niedrigste(n) Messwert(e), M- und -X, bzw. die davon betroffenen Messpunkte auf der Karte dadurch gekennzeichnet, dass die entsprechenden roten bzw. blauen Polygone mit einem weißen Gitterraster hinterlegt sind. Dies ermöglicht das leichte Auffinden der dem jeweiligen Prüfbezugs punkt ähnlichsten bzw. unähnlichsten Orte.

Zur Ähnlichkeitsmessung wurde in den Karten 1–5 und 7 der RIW_{jk} herangezogen, der sich, wie bereits in Kapitel 2.1 erwähnt, an der Vorstellung orientiert, dass dialektale Ähnlichkeiten durch die Messung der relativen Anzahl übereinstimmender Sprachmerkmale abgebildet werden können.

Die Karten wurden übrigens mit Hilfe einer VDM-eigenen Exportfunktion erstellt, die es ermöglicht, alle zu einem dialektometrischen Visualisierungstyp gehörenden Bestandteile als einzelne Graphikdateien zu exportieren, um sie dann in einem Layout-Programm zu einer druckfertigen Vorlage zusammenzustellen³⁶.

2.3.1. Interpretation der Karte 1 (siehe Seite 244)

Das erste Ähnlichkeitsprofil (Karte 1) ist auf der Basis des Gesamtcopus von 3.302 Arbeitskarten erstellt worden und berücksichtigt somit hauptsächlich phonetische und lexikalische sowie auf Makro-Merkmale bezogene Variation. Als Prüfbezugs punkt wurde der am Südrand des Untersuchungsnetzes positionierte Kunstpunkt Nr. 999 “italiano standard” gewählt. Die Karte zeigt also dialektale Ähnlichkeiten aller 217 *AD-I*-Messpunkte mit dem Standarditalienischen auf.

Legende und Histogramm entnehmen wir zunächst, dass relativ mehr Ortsdialekte (knapp 60%) überdurchschnittliche Affinität zum Italienischen aufweisen als umgekehrt. Die drei höchsten, über dem arithmetischen Mittelwert liegenden Farbklassen (rot, orange, gelb) betreffen nämlich insgesamt 128 Messpunkte (36 + 53 + 39), während unter dem Mittelwert nur 89 Punkte (23 + 19 + 47) zu liegen kommen. Das Italienische ist also gut in das Gesamtnetz integriert. Anders formuliert könnte man sagen, dass bereits knapp 60% aller im *AD-I* erfassten Ortsdialekte in überdurchschnittlichem Maß vom Italienischen beeinflusst werden, während sich nur rund 40% mehr oder weniger gut gegen innerlinguistische “Angriffe” seitens des Italienischen verteidigen können.

³⁶ Dies gilt auch für alle über unsere Projekt-Homepage <http://ald.sbg.ac.at/ald/alddm/> abrufbaren Beispielvisualisierungen (Dichtekarten, Ähnlichkeitskarten, Schottenkarten und Clusteranalysen).

Dem Italienischen am absolut nächsten steht der (metaphorisch gesprochen auch als “bester Freund” zu bezeichnende) trentinische P. 119 Lèvico am Westeingang der Val Sugana mit einem RIW_{999, 119} von 76,46³⁷. Dies bedeutet, dass im Ortsdialekt von Lèvico mehr als drei Viertel aller von uns verrechneten 3.302 Merkmale bzw. aller (nach Abzug der Nullstellen) gültigen Taxate mit dem Italienischen identische Ausprägungen aufweisen. In einer ähnlichen Lage (zwischen 71 und 76% schwankende Ähnlichkeitswerte) befinden sich, innerhalb des Trentino, drei ebenfalls rot eingefärbte Messpunkte aus der Val Sugana (P. 115 Strigno, P. 116 Castello Tesino und P. 117 Tezze), zwei Ortschaften aus dem Fleimstal (P. 104 Tèsero, P. 105 Cavalese) sowie Trient (P. 122) und San Michele all’Adige (P. 66) in der Val Lagarina. Den Löwenanteil aller dem Italienischen besonders nahe stehenden Dialekte zieht jedoch das Veneto bzw. das Venedische auf sich: so fallen dort alle Messpunkte der Provinz Vicenza sowie ein Gutteil der in den Provinzen Treviso und Venedig explorierten Ortschaften in die rot eingefärbte Zone. Schließlich zählen dazu noch zwei auf (politisch gesehen) friaulischem Territorium liegende Punkte (P. 214 Pordenone, P. 216 Prata di Pordenone). Im *AD-I*-Ausschnitt tritt somit das Venedische als jener Dialekttyp in Erscheinung, der dem Italienischen in innerlinguistischer Hinsicht unter allen berücksichtigten Dialekten am ähnlichsten ist. Dazu zählen ferner die unter starkem venedischen Einfluss stehenden Dialekte des östlichen Trentino (Val Sugana) und der westfriaulischen-venedischen Übergangszone im Bereich von Pordenone.

Als zweithöchste Klasse (orange Signatur auf Karte 1), die sich in der Art eines konzentrischen (Halb-)Kreises um die roten Gebiete legt, erkennen wir das Zentral- bzw. Südtrentinische (Val Lagarina, Val di Cembra, Bozner Unterland), das venedisierte Osttrentinische des Primiero sowie wiederum das Venedische selbst (Nordteil der Provinz Treviso, Südteil der Provinz Belluno). Alle 53 in dieser Klasse subsumierten Messpunkte verfügen über Ähnlichkeits- bzw. RI-Werte zwischen 67 und 71%.

In direktem Anschluss an diese Gebiete kommen die gelb signierten Messpunkte der dritthöchsten Wertklasse zu liegen (Ähnlichkeitswerte zwischen 61 und 66%). Dabei treten an der Westflanke einerseits die im nördlichen Trentino gelegenen, ladino-anaunischen Dialekte der Täler Sulzberg und Nonsberg und andererseits die westtrentinisch-lombardischen Mundarten Judikariens in Erscheinung. Im Nordosten unseres Untersuchungsgebietes macht sich überdies eine markante, zwischen dem Dolomitenladinischen und dem Westfriaulischen gelegene Zone bemerkbar, der unschwer als agordinisch-cadorinisch-comelianischer Korridor

³⁷ Siehe dazu das mit weißem Gitterraster signierte, rote Polygon 119 im Zentrum von Karte 1. In der Klassifikation spricht man in diesem Zusammenhang von “polar types” (cf. BAILEY 1994, 23).

identifiziert werden kann. Zusammen mit den ladino-anaunischen Dialekten des Sulzbergs und des Nonsbergs gehören die letztgenannten, gemeinhin als ladino-venedisch bzw. ladino-cadorinisch klassifizierten Sprachräume bekanntlich zu den von ASCOLI (1873) als “anfizone” bezeichneten, mehr oder weniger ladinierten Übergangszonen.

Die Betrachtung der vierthöchsten (grün signierten) Klasse führt uns in den Bereich der Dialektgebiete mit bereits unterdurchschnittlicher Ähnlichkeit zum Italienischen. Nur mehr 53–61% aller Taxate decken sich dort mit den standarditalienischen Merkmalsausprägungen. In diese Klasse fällt zunächst ein kompakter (alpin-)lombardischer Block (Veltlin, Puschlav, Val Camonica) inkl. einiger westtrentinisch-lombardischer Messpunkte (Val Rendena, Judikarien). Die übrigen Messpunkte dieser Klasse fallen fast ausnahmslos auf rätoromanisches Territorium, sei es in den südlichen Dolomitentälern Fassa, Buchenstein (mit Colle Santa Lucia, Laste und Rocca Pietore) und Ampezzo, sei es im Westen Friauls.

Die verbleibende Rätoromania ist hell- bzw. dunkelblau signiert, gehört also zu den am wenigsten italianisierten bzw. mit dem Italienischen verwandten Gebieten unseres Untersuchungsnetzes. Zur vorletzten, nur mehr 45–53% Ähnlichkeit mit dem Italienischen aufweisenden Gruppe gehören Enneberg, das obere Fassatal und Westfriaul. Am absolut weitesten vom Italienischen entfernt erscheinen hingegen die Bündnerromania (Engadin und Münstertal) sowie die nördlichen Dolomitentäler (Gadertal und Grödnertal) mit 36–44% Ähnlichkeit.

Als absolute Antipode tritt dabei der oberengadinische Messpunkt 8 Brail mit einem $RIW_{999,8}$ von 36,09 in Erscheinung³⁸. Das dortige *Puter* teilt nur mehr gut ein Drittel aller untersuchten Sprachmerkmale mit dem Italienischen. Das Oberengadinische von Brail stellt übrigens, wie eine Analyse aller 218 Ähnlichkeitskarten gezeigt hat, aus der Sicht von 170 *AD-I*-Messpunkten bzw. Prüfbezugsunkten den absolut unähnlichsten Dialekt dar. Es wird also von einem Großteil unseres Untersuchungsnetzes zum “Prügelknaben” gemacht³⁹. Dies ist innerhalb des Gesamtcorpus insofern bemerkenswert, als lediglich weitere neun Messpunkte überhaupt als negative Antipoden in Frage kommen, die jedoch, im Gegensatz zu Brail, nur aus jeweils 1–10 Bezugspunkten “Prügel” beziehen⁴⁰.

³⁸ Siehe dazu das mit weißem Gitterraster signierte, blaue Polygon im Nordwesten von Karte 1.

³⁹ Zur metaphorischen Umschreibung von interdialektalen Ähnlichkeiten bzw. Differenzen anhand der sozialpsychologischen Konzepte “Freundschaft – bester Freund” bzw. “Feindschaft – Prügelknabe” cf. GOEBL 1994, passim.

⁴⁰ Ab der nächsten Version von *VDM* (2003) wird es auch möglich sein, eigene Antipoden-Karten abzurufen, die sowohl die als “beste Freunde” bzw. “Prügelknaben” in Frage kommenden Messpunkte, als auch das jeweilige Einzugsgebiet, aus dem die “Streicheleinheiten” bzw. “Prügel” stammen, abbilden.

2.3.2. Interpretation der Karte 2 (siehe Seite 245)

Das zweite Ähnlichkeitsprofil (Karte 2) basiert wiederum auf dem gesamten Corpus von 3.302 Arbeitskarten. Als Prüfbezugspunkt wurde dieses Mal ein Ort aus dem Gadertal, nämlich P. 83 San Martin de Tor / St. Martin in Thurn gewählt. Die Karte zeigt also die “Welt” der dialektalen Affinitäten (stets nach RIW_{jk}) aus der Sicht der nördlichen Ladinia.

Werfen wir zunächst wieder einen Blick auf Legende und Histogramm, um uns die Gesamtverteilung der Ähnlichkeitswerte vor Augen zu führen. Im Gegensatz zu dem auf Karte 1 abgebildeten Profil stellt sich die Ähnlichkeitsverteilung jetzt als “linkssteil” (d.h. mit einem starken Überhang bei den im Histogramm links vermerkten, niedrigsten Ähnlichkeitswerten) dar, was bedeutet, dass wesentlich weniger Ortsdialekte (nämlich nur gut ein Drittel) überdurchschnittliche Affinität zum Gadertalischen von St. Martin aufweisen als umgekehrt. Das Gadertalische ist also um vieles schlechter in das Gesamtnetz eingebunden bzw. integriert als das Italienische. So kommen nur 75 Messpunkte (= 7 + 3 + 65) in den warm (i.e. rot, orange und gelb) eingefärbten Farbklassen oberhalb des Mittelwerts zu liegen, während 142 Punkte (= 24 + 41 + 77) unter dem Mittelwert angesiedelt sind. Die Bandbreite an auftretenden RI-Werten unterscheidet sich insofern vom standarditalienischen Ähnlichkeitsprofil (Karte 1), als die Maxima hier deutlich höher liegen und bis knapp an die 90%-Grenze heranreichen.

Als absolut “bester Freund” figuriert dabei der (im Untersuchungsnetz des *AD-I*) direkte Nachbarort von St. Martin, nämlich P. 84 La Val / Wengen mit einem $RIW_{83,84}$ von 89,43⁴¹, den bei insgesamt 3.302 erstellten Arbeitskarten nur die Merkmalsausprägungen von rund 330 Karten von St. Martin trennen. Ähnliches gilt für alle übrigen Messpunkte des Gadertals, die hier eine kompakte (rote) Klasse bilden. Im oberen Talbereich (P. 85 San Linêrt / St. Leonhard, P. 89 Calfosch / Kolfuschg, P. 90 Corvara und P. 91 San Ciascian / St. Kassian) pendeln sich dabei die RI-Werte zwischen 85 und 88 ein, während am nördlichen Taleingang RI-Werte um 83% verzeichnet werden. Davon sind in erster Linie das *Maréo* von Enneberg / La Pli (P. 81) sowie das *Ladin* von Rina / Welschellen (P. 82) betroffen⁴².

Auch die zweithöchste Gruppe (orange Signatur) ist für Kenner der Dolomitenladinia auf den ersten Blick klar zu identifizieren. Es handelt sich um die drei

⁴¹ Siehe das mit weißem Gitterraster signierte, rote Polygon im Nordwesten von Karte 2.

⁴² Zu den hier in orthographisch vereinfachter Form wiedergegebenen Bezeichnungen der jeweiligen Ortsdialekte cf. die im Band 1 des *AD-I* abgedruckte Karte 3 “nome dialettale della parlata”.

grödnerischen Messpunkte P. 86 Bula / Pufels, P. 87 S. Cristina / St. Christina und P. 88 Sëlva / Wolkenstein. Das Grödnerische weist dabei durchwegs RI-Werte von 70 auf. Die Klassen 1 und 2 stellen zugleich die nördliche, vom Deutschen überdachte Ladinia dar, die sich, laut Auskunft unseres Ähnlichkeitsprofils, hier deutlich von den südlichen, italienisch überdachten Talschaften abhebt. Dies ist als erstes, (auch) aus dialektometrischer Sicht erkenntliches Indiz für die innere Raumgliederung der Dolomitenladinia zu werten.

Die dritte (gelb signierte) Gruppe weist tendenziell noch immer mehr Ko-Identitäten (nämlich rund 50–62%) als Ko-Differenzen mit St. Martin auf. Sie geht bezüglich ihrer räumlichen Struktur jedoch weit über die jeweils nur ein einzelnes Talsystem umfassenden Klassen 1 (Enneberg und Gadertal) und 2 (Grödnertal) hinaus und besteht aus mehreren Teilgebieten, die ihrerseits als eigene Dialektlandschaften (oder zumindest als Teile davon) gefasst werden können. Im Nordwesten finden wir zunächst alle bündnerromanischen Messpunkte (PP. 1–12) mit Ähnlichkeitswerten, die zwischen 50 und 55 liegen. Dazu gesellen sich, direkt im Südosten an das Münstertal anschließend, drei nordostlombardischen Ortsdialekte im Veltlin (PP. 15–17), ferner das Alpinlombardische von Poschiavo (P. 20) sowie Teile der oberen Val Camonica (PP. 25–28). Auf trentinischem Territorium kommen je drei Messpunkte aus Sulz- und Nonsberg hinzu, deren Ähnlichkeitswerte sich allesamt am untersten Ende des hier in Frage kommenden Intervalls bei etwa 50 einpendeln. Die gelben Polygone der Osthälfte unseres Netzes betreffen zwei Teilgebiete der Rätoromania, nämlich einerseits die gesamte südliche Dolomitenladinia (Fassatal, Buchenstein, Ampezzo) inkl. einiger mehr oder weniger ladinisierter, am Südrand vorgelagerter Ortschaften und andererseits Westfriaul, das durch einen cadorinischen Pfad (PP. 134–136) mit Cortina d’Ampezzo (P. 92) verbunden ist. Innerhalb der historischen Dolomitenladinia (inkl. P. 139 Laste und P. 140 Rocca Pietore) schwanken die Ähnlichkeitswerte zwischen 55 (P. 92 Cortina) und 62 (Buchenstein und oberes Fassatal), in den südlich (Fleimstal, P. 102 Forno; Agordino, P. 140 Falcade) bzw. südöstlich (Cadore) davon gelegenen Übergangsgebieten liegen sie hingegen nur mehr bei durchschnittlich 50⁴³. In Friaul finden wir schließlich RI-Werte zwischen 50 und 56, wobei eine deutliche Abnahme von Nord nach Süd zu bemerken ist.

Zur Vermeidung von Fehlinterpretationen sei mit Nachdruck folgender Hinweis gegeben: das in Ähnlichkeitsprofilen gemeinsam erfolgende Auftreten ver-

⁴³ Die hier angesprochenen Unterschiede können übrigens gut visualisiert werden, wenn man die Intervallzahl von 6 auf beispielsweise 10 erhöht. *VDM* erlaubt dabei Segmentierungen, die zwischen minimal 2 (ergibt ein sehr grobkörniges, um nicht zu sagen “grob-schlächtiges” Bildtypenprofil) und maximal 20 Intervallen (ergibt ein sehr detailliert strukturiertes Bildtypenprofil) liegen.

schiedener (hier: bündnerromanischer, ostlombardischer, ladino-anaunischer, dolomitenladinischer, ladino-venedischer und westfriaulischer) Dialekte kann keine zuverlässige Auskunft über allfällige Ähnlichkeitsrelationen zwischen den angeführten Dialekten selbst geben, sondern nur so interpretiert werden, dass alle Dialekte einen ungefähr gleich großen Abstand vom bzw. eine ungefähr gleich große Ähnlichkeit mit dem Prüfbezugspunkt (hier dem Gadertalischen von St. Martin) haben. Um Ähnlichkeiten zwischen z.B. Friaulisch und Bündnerromanisch darzustellen, müssten unbedingt Prüfbezugspunkte aus diesen beiden Sprachlandschaften gewählt werden.

Wie schon eingangs erwähnt, weisen gut zwei Drittel aller Messpunkte unterdurchschnittliche Affinitäten zum (hier von St. Martin repräsentierten) Gadertalischen auf. Die grün signierte Klasse grenzt dabei meist direkt an die soeben vorgestellte (gelb eingefärbte) Gruppe an. Ihr RIW liegt jeweils zwischen 45 und 49%. Zu dieser numerisch größten Klasse gehören 77 Messpunkte aus folgenden lombardischen, trentinischen und venedischen Land- bzw. Talschaften: Val Camonica, Judikarien, Val Rendena, Val Lagarina, Val di Cembra, Fleimstal, Primiero, südliches Agordino, Val di Zoldo, nördliches Cadore und Comelico.

Alle hellblau eingefärbten Polygone repräsentieren Messpunkte mit noch geringerer Ähnlichkeit (42–45%). Die 41 davon betroffenen Ortschaften liegen zum Großteil im prä-alpinen Veneto, und zwar im Bereich der Provinz Treviso, im Süden der Provinz Belluno und im friaulisch-venedischen Übergangsraum bei Pordenone. Hinzu kommen vereinzelte Messpunkte aus der Lombardei und aus dem Trentino.

Die dem Gadertalischen von allen am unähnlichsten Dialekte (24 dunkelblau signierte Orte) haben im Schnitt nur mehr rund 40% innerlinguistische Gemeinsamkeiten mit St. Martin. Die räumliche Ausbreitung dieser Gruppe ist in etwa deckungsgleich mit jenen venedischen und trentinischen Gebieten (Provinzen Vicenza und Treviso bzw. Val Sugana), die auf Karte 1 als besonders dem Standarditalienischen nahe stehende (und dort rot eingefärbte) Ortsdialekte identifiziert werden konnten. Auch das Standarditalienische selbst fällt in diese Gruppe. Der dort gemessene RIW_{83,999} beträgt 40,30. Aus der Sicht von St. Martin gibt schließlich das rund 25 km nordwestlich von Vicenza gelegene Schio (P. 178) den "Prügelknaben" ab, dessen RIW_{83,178} bei 38,55 liegt⁴⁴.

⁴⁴ Siehe das mit weißem Gitterraster signierte, blaue Polygon im Süden von Karte 2.

2.3.3. Interpretation der Karten 3–6 (siehe die Seiten 246–249)

Eine Frage, auf die bislang noch überhaupt nicht eingegangen werden konnte, betrifft die Eruierung jener Merkmalsgruppen, die für die in den Ähnlichkeitsprofilen zum Vorschein kommenden Raumstrukturen hauptverantwortlich sind/sein könnten. Bislang (Karten 1 und 2) wurde nämlich immer auf die gesamte Datenmatrix rekuriert, wobei bekanntlich alle Merkmale in *adansonistischer*⁴⁵ Weise, i.e. gleich bzw. nicht gewichtet wurden. Eine Auswertung, die sich an bestimmten Teilcorpora orientiert, kann nun Auskunft über mehr oder weniger zur Raumbildung beitragende Merkmalsgruppen liefern. Zu diesem Zweck haben wir das in Karte 2 abgebildete Ähnlichkeitsprofil zum Gadertalischen von St. Martin in Thurn insgesamt fünf Mal mit jeweils nach innerlinguistischen Kriterien gefilterten Teilcorpora nachverrechnet. Karte 3 zeigt das Ähnlichkeitsprofil anhand des Vokalismus (1.101 Arbeitskarten), Karte 4 betrifft ausschließlich den Konsonantismus (936 AK), Karte 5 das Lexikon (490 AK, ohne Berücksichtigung der Taxierung von Germanismen), Karte 6 noch einmal das Lexikon (490 AK; anstelle des RIW_{jk} wurde hier jedoch das gewichtende Ähnlichkeitsmaß GIW_{jk} verwendet) und Karte 7 die Morphologie (309 AK zur Pluralbildung). Aus Platzgründen können wir beim Vergleich der Karte 2 mit den Karten 3–7 lediglich auf die markantesten Unterschiede im Bereich der gesamten Ähnlichkeitsverteilung sowie auf die jeweiligen Minima und Maxima eingehen.

2.3.3.1. Vergleich der Karten 2 und 3 (Vokalismus)

Die Asymmetrie der Ähnlichkeitsverteilung von Karte 2 stellte sich im Histogramm als linkssteile Kurve dar, wodurch der anteilmäßige Überhang an dem Gadertalischen unähnlichen (bzw. nur unterdurchschnittlich ähnlichen) Messpunkten zum Ausdruck kam. Das Gadertalische erschien also, nach Auskunft der Ähnlichkeitsverteilung zu Karte 2, eher schlecht in das Gesamtnetz integriert. Im Histogramm zu Karte 3, in der 1.101 nur den Vokalismus betreffende Arbeitskarten berücksichtigt wurden⁴⁶, sehen wir hingegen eine um vieles ausgewogenere Ähnlichkeitsverteilung. 94 Messpunkte liegen unterhalb, die Mehrheit von 123 Messpunkten jedoch oberhalb des Mittelwerts. Fast drei Viertel aller Ortschaften befinden sich schließlich in den beiden um den Mittelwert angesiedelten (grün und gelb signierten) Klassen. Dazu gehören auch alle rätoromanischen Gebiete

⁴⁵ Cf. ADANSON 1763.

⁴⁶ Rund 33% beziehen sich dabei auf den vokalischen Vorton, 45% auf den Nachton und der Rest (42%) auf den Hauptton.

mit Ausnahme der nördlichen Dolomitenladinia. Der gadertalische Vokalismus scheint insgesamt bedeutend besser in das Untersuchungsnetz eingebettet zu sein als das Gesamtcorpus.

Die Breite des Gesamtintervalls aller beim Vokalismus gemessenen Ähnlichkeitswerte unterscheidet sich zwar nicht wesentlich von jener des Gesamtcorpus, allerdings kann eine deutliche Verschiebung nach unten festgestellt werden. So kommen etwa die Maxima nicht mehr an die in Karte 2 vermerkten 90% heran, während die beim Gesamtcorpus im Bereich von 38% liegenden Minima jetzt bis knapp an die 20%-Grenze hinunter reichen. Der "Prügelknaube" (P. 173 Recoaro) ist somit datenseitig deutlich stärker exponiert und liegt wiederum in der Provinz Vicenza, und zwar in direkter Nachbarschaft der auf Karte 2 eruierten Antipode (P. 178 Schio). Zur Klasse der bezüglich des Vokalismus am weitesten vom Gadertalischen entfernten Orte gehört übrigens auch das Standarditalienische mit einem RI-Wert von 22,67. Als "bester Freund" tritt wiederum P. 84 La Val / Wengen in Erscheinung, allerdings teilt der dortige Ortsdialekt beim Vokalismus nur mehr rund 78% aller Merkmalsausprägungen mit St. Martin.

Auch die auf Karte 3 im direkten Vergleich mit Karte 2 ersichtliche Raumkammerung stellt sich entsprechend ausgeglichen dar, dominieren doch die Gebiete mit durchschnittlicher Ähnlichkeit das Gesamtbild. Als dem Gadertalischen von St. Martin affinste Dialektgebiete treten wiederum die beiden Täler der nördlichen Dolomitenladinia sowie ein Messpunkt aus dem Unterengadin (P. 1 Tschlin) in Erscheinung.

2.3.3.2. Vergleich der Karten 2 und 4 (Konsonantismus)

Die Grundlage für den Vergleich der Karten 2 und 4 bilden 936 Arbeitskarten zum Konsonantismus⁴⁷. Die Asymmetrie der Ähnlichkeitsverteilung im Histogramm von Karte 4 gleicht jener von Karte 2, da auch hier die überwiegende Mehrzahl an Messpunkten (nämlich 143 von 217) unterdurchschnittliche Ähnlichkeitswerte aufweist. Der Konsonantismus entspricht also offensichtlich viel stärker der auf Karte 2 ersichtlichen Profilbildung als der Vokalismus.

Als "beste Freunde" des Konsonantismus von St. Martin sind Enneberg und das gesamte Gadertal (mit dem Spitzenwert von über 95% Ähnlichkeit in P. 89 Kol-

⁴⁷ Rund 42% beziehen sich dabei auf den Konsonantismus vor, 58% auf den Konsonantismus nach dem vokalischen Hauptton.

fuschg), ferner das Grödnertal (RIW ca. 72) und Buchenstein (RIW ca. 70) zu erkennen. Die noch mit überdurchschnittlichen Ähnlichkeitswerten (55–68%) ausgestattete dritte Klasse (gelbe Signatur) reicht einerseits von der Bündnerromania über das Veltlin in den Nonsberg und andererseits vom Fleimstal und Fassatal über Ampezzo exakt bis an die nordwestliche Außengrenze des Friaulischen heran und beinhaltet auch die im Süden und Osten der Ladinia vorgelagerten Übergangszonen (Agordino, Cadore, Comelico). Am untersten Ende der Werteskala stoßen wir zunächst auf die Besonderheit, dass nur ein einziger Ort in der niedrigsten Klasse zu liegen kommt, der somit notgedrungen auch den “Prügelknaben” abgibt. Es handelt sich dabei um die südbairische Sprachinsel Lusern (P. 118) mit einem RI-Wert von 33,37, an der der romanische Mesoлект (und nicht etwa der germanische Basilekt) erhoben wurde. Die vorletzte (hellblau signierte) Klasse weist Ähnlichkeitswerte zwischen 40 und 46% auf und kommt mit 35 Messpunkten wiederum v.a. im Bereich der Provinz Vicenza zu liegen.

2.3.3.3. Vergleich der Karten 2, 5 und 6 (Lexikon)

In Karte 5 tritt uns ein ausschließlich anhand lexikalischer Merkmale generiertes Ähnlichkeitsprofil gegenüber, wobei insgesamt 490 Arbeitskarten verdichtet wurden. Es handelt sich dabei vorwiegend um lexikalische Mikro-Analysen, bei denen sowohl Unterschiede im Wortstamm als auch allfällige Affigierungen (Prä- und Suffixe) gesondert taxiert wurden. Lexikalische Makro-Phänomene, wie etwa die Zugehörigkeit der Belege zu Wortfamilien, das Auftreten von Germanismen oder Abweichungen der Antworten vom etymologischen Typ des (italienischen) Kartentitels sind hier noch unberücksichtigt geblieben, da sich dieser Bereich derzeit noch in Ausarbeitung befindet. Das verrechnete lexikalische Corpus ist also noch nicht komplett, weshalb auch die Interpretation des Ähnlichkeitsprofils nur als provisorisch anzusehen ist. Zudem muss in Erinnerung gerufen werden, dass der *AD-I per definitionem* auf die Erfassung und Dokumentation phonetisch relevanter Information ausgelegt ist, das Lexikon also nur eine untergeordnete Rolle spielen kann.

Auf dem Histogramm zu Karte 5 ist zunächst eine linksversetzte oder linksschiefe Ähnlichkeitsverteilung erkennbar, die darauf verweist, dass nur relativ wenige Ortsdialekte überdurchschnittlich große lexikalische Gemeinsamkeiten mit dem Gadertalischen aufweisen, während die überwiegende Mehrzahl aller Messpunkte unterdurchschnittliche Ähnlichkeiten, also große lexikalische Differenzen zeigt. Die dabei auftretenden Minima sind jedoch höher als bei allen anderen hier vorge-

stellten Ähnlichkeitsprofilen. Jeder Messpunkt teilt nämlich zumindest 44% aller lexikalischen Merkmalsausprägungen mit dem Gadertalischen. Auch die höchsten RI-Werte überschreiten die in den bislang präsentierten Ähnlichkeitsprofilen gemessenen Maxima. Der nördlich von St. Martin gelegene Messpunkt Nr. 82 Rina / Welschellen verfügt etwa über einen $RIW_{83,82}$ von 95,69 (= Maximum) und auch das übrige Gadertal (rote Signatur) weist Ähnlichkeitswerte zwischen 91 (P. 90 Corvara) und 94% (P. 81 La Pli / Enneberg) auf, wobei innerhalb des Tals ein leichtes Absinken der lexikalischen Ähnlichkeit von Nord nach Süd zu bemerken ist. Mit schon deutlich geringerer Ähnlichkeit folgen die (orange signierten) Messpunkte im Grödnertal (RIW zwischen 75 und 77) und in Buchenstein (69–70). Die innerladinische Nord-Südgrenze wird hier (wie schon im Profil zum Konsonantismus) zum zweiten Mal durchbrochen, was auf eine Art lexikalische (und konsonantische) Vermittlerrolle des Buchensteinischen zwischen der nördlichen und der südlichen Ladinia hinweist. Am untersten Ende der Skala finden wir dieses Mal die Bündnerromania, die mit P. 8 Brail auch den (bereits einmal aus der Sicht des Standarditalienischen in Erscheinung getretenen) “Prügelknaben” ($RIW_{83,8} = 44,33$) stellt. In lexikalischer Hinsicht scheint also, nach Auskunft von Karte 5, das Bündnerromanische weniger Ähnlichkeit mit dem Gadertalischen zu haben als etwa das Standarditalienische, das dem Lexikon von St. Martin hier genauso nahe steht wie etwa das Friaulische. Dieser Befund darf freilich nicht vorschnell interpretiert werden, denn es handelt sich, wie bereits erwähnt, um ein relativ kleines Corpus, dessen Zusammenstellung überdies nur über Mikro-Taxierungen erfolgte. So würde schon die Einbeziehung der eingangs erwähnten Makro-Phänomene genauso Verschiebungen mit sich bringen, wie der Einsatz eines auf die Frequenz der Antworttypen Bezug nehmenden Ähnlichkeitsmaßes.

Letzteres kann auf Karte 6 verifiziert werden, die anhand des *Gewichtenden Identitätswerts* GIW_{jk} erstellt wurde. Im Gegensatz zum RIW_{jk} wird in diesem Zusammenhang die Häufigkeit, mit der eine bestimmte Merkmalsausprägung innerhalb eines Merkmalsvektors der Datenmatrix auftritt, gewichtet. “The philosophy here is that agreement among rare character states is a less probable event than agreement for frequent character states and should therefore be weighted more heavily”⁴⁸. Ohne Gewichtung gilt also: je höher die Frequenz eines Taxats in der Datenmatrix ist, desto höher ist auch sein Beitrag zum numerischen Gesamtertrag der Ähnlichkeitsmessung. Beim Einsatz gewichtender Verfahren werden daher häufig auftretende Ko-Identitäten gleichsam numerisch “benachteiligt”, seltene Übereinstimmungen numerisch “bevorzugt”⁴⁹. Das Histogramm von Karte 6 zeigt eine noch linkslasti-

⁴⁸ SNEATH/SOKAL 1973, 140.

⁴⁹ Zum messtheoretischen Hintergrund gewichtender Ähnlichkeitsmaße cf. GOEBL 1998, 981–984.

gere oder linksschieferer Ähnlichkeitsverteilung als jenes von Karte 5. Das Gadertalische (von St. Martin) scheint also bezüglich des Lexikons eine nur kleinräumig ausgreifende Sonderstellung einzunehmen, denn auch anhand des GIW treten uns relativ große Ähnlichkeiten nur im Bereich der nördlichen Ladinia gegenüber, wobei die deutliche Abstufung zwischen Gadertal (rot) und Grödnertal (orange) gut erkennbar wird. Die Zone der auf Karte 6 gelb signierten Messpunkte, deren jeweiliger GIW_{jk} größer als 19,16 ist und knapp oberhalb des Mittelwertes liegt, umfasst wiederum die gesamte südliche Dolomitenladinia inkl. der ihr vorgelagerten, mehr oder weniger ladinisierten Übergangsgebiete Agordino, Cadore und Comelico. Die wesentlichsten Unterschiede zu Karte 5 treten in den untersten Wertklassen auf, wo wir auf die mit Abstand niedrigsten bislang beobachteten Identitätswerte stoßen, was auf die Gewichtung zurückzuführen ist. Dabei kann man in den grün eingefärbten Gebieten gut die beiden rätoromanischen Kernzonen der Bündnerromania und (Nordwest-)Friauls erkennen, die durch lexikalisch in etwa gleich weit vom Gadertalischen entfernte Übergangszonen (v.a. oberes Veltlin, Sulzberg, Nonsberg, Fleimstal) miteinander verbunden sind. Das verbleibende Untersuchungsgebiet stellt sich als das lexikalisch am weitesten vom Gadertalischen entfernte (hell- und dunkelblau eingefärbte) und von Norden nach Süden durch eine progressive Zunahme der lexikalischen Differenz gestaffelte restliche Oberitalien dar, wobei der absolut niedrigste Identitätswert im zur Provinz Venedig zählenden P. 194 Portogruaro ($GIW_{83,194} = 10,07$) gemessen wird.

Wie einem visuellen Pauschalvergleich der Karten 5 und 6 entnommen werden kann, wirkt sich der Einsatz des gewichtenden Ähnlichkeitsmaßes GIW insofern auf das Bildprofil aus, als dadurch eine stärkere Polarisierung bzw. eine Erhöhung des Bildkontrastes erzeugt wird, die schon in der unterschiedlich akzentuierten Schiefe der beiden Histogramme zum Ausdruck kommt. So können zwar dialektale Ähnlichkeiten im direkten Umfeld des Prüfbezugspunktes (P. 83, St. Martin) durch die stärkere Gewichtung der eher niederfrequenten, typisch dolomitenladinischen Merkmalsausprägungen deutlicher sichtbar gemacht werden (= GI-Werte zwischen 19 und 90), andererseits werden die geolinguistischen Affinitäten mit steigender geographischer Distanz zum Referenzpunkt (= GI-Werte zwischen 10 und 19) durch die Gewichtung "benachteiligt" und fallen somit weniger deutlich strukturiert ins Auge.

2.3.3.4. Vergleich der Karten 2 und 7 (Morphologie)

Das letzte hier vorzustellende Vergleichsprofil betrifft ein kleines morphologisches Corpus. Karte 7 basiert auf der Auswertung von 309 Arbeitskarten zur Plu-

ralbildung bei Adjektiven und Substantiven. Dabei wurden vier Taxierungstypen unterschieden:

1. der dem Italienischen / Toskanischen entsprechende *vokalische Plural* vom Typ $\overline{[ácido - ácidí]}$,
2. der im Nominalbereich *nicht* (mehr) *markierte Plural* vom Typ $\overline{[áži - áži]}$,
3. der *sigmatische Plural* vom Typ $\overline{[garp - garps]}$ und
4. der *palatale Plural* vom Typ $\overline{[fort - foré]}$.⁵⁰

Ein fünfter Pluraltyp⁵¹, der v.a. bei im Singular auf R + Vokal auslautenden Nomina auftritt (Typ $\overline{[kor - korš]}$ ⁵²), wurde zum sigmatischen Plural gezählt. Bedingt durch die selektive Datenlage, gibt das auf Karte 7 abgebildete Ähnlichkeitsprofil (das wieder anhand des RIW_{jk} erstellt wurde) eine inhaltlich nur sehr eingeschränkte bildtypologische/geotypologische Auskunft, und zwar nur auf der Basis der oben zitierten vier Pluralarten. Die linkssteile Kurve der Ähnlichkeitsverteilung im Histogramm zu Karte 7 verrät zunächst, dass wir es hier mit einem überaus kleinen Kerngebiet zu tun haben (nur 49 von 217 Messpunkten liegen oberhalb des Mittelwerts), das sich numerisch sehr deutlich vom Gros der Dialekte (77% aller Messpunkte) mit typologisch nicht dem Gadertalischen entsprechender Pluralbildung abhebt. In Enneberg und im Gadertal (RIW zwischen 94 und 98), im Grödner Tal (RIW ca. 85) und im oberen Fassatal (RIW ca. 80) stoßen wir auf Ähnlichkeitswerte, die in Einzelfällen bis knapp an die 100%-Marke heranreichen, wobei wiederum P. 82 Rina / Welschellen als “bester Freund” von St. Martin figuriert, der es auf 98,7% Ähnlichkeit bringt. Die im *AD-I* erfassten Messpunkte der Bündnerromania gehören mit einem durchschnittlichen RIW von 61 zur zweithöchsten (orange eingefärbten) Klasse, genauso wie unser friaulischer Ausschnitt mit 66 (Provinz Pordenone) bis 71% (Provinz Udine) Ähnlichkeit. Am untersten Ende der zweiten Klasse pendeln sich die RI-Werte von Ampezzo (61%) und des bereits mehrfach auffällig gewordenen oberen Boite-Tales (PP. 134–136, RIW 54–61) ein. Jene zehn Messpunkte, die in die dritte (gelb signierte) Gruppe fallen, betreffen schließlich den Rest der (südlichen) Dolomitenladinia (Buchenstein mit rund 40%, Unterfassa mit 34–38% Ähnlichkeit) und Westfriauls (28–32%). Die Punkte 93 Colle S. Lucia, 138 Rocca Pietore und 139 Laste haben bereits nur mehr unterdurchschnittliche Ähnlichkeitswerte mit St. Martin (ca. 23%). Sie werden, gemeinsam mit Teilen Judikariens, der Val

⁵⁰ Alle Beispiele stammen aus dem *AD-I*, Band I, Karte 5 *acido / acidi*. Zum palatalen Plural cf. BENINCÀ / VANELLI 1995.

⁵¹ KRAMER (1976, 47) bezeichnet diesen Typ als “kontaminierten Plural”.

⁵² Beispiel aus *AD-I*, Band II, Karte 217 *il cuore / i cuori*.

Rendena und Westfriauls (z.B. PP. 203 Cimolais und 204 Erto) der grün signierten Gruppe zugeschlagen.

Das Ähnlichkeitsprofil von Karte 7 unterscheidet ganz deutlich zwischen rätoromanischen Kern- und Randgebieten einerseits (warme Farbsignaturen), und weit vom typisch rätoromanischen (sigmatischen) Plural entfernten lombardischen, trentinischen und venedischen Dialekträumen andererseits (blaue, i.e. kalte Farbsignaturen). Die größten Unterschiede finden sich im Puschlav und im Veltlin, wo der (in den Dolomiten viel weniger verbreitete) “nicht markierte” Pluraltyp vorherrscht und wo etwa im P. 17 Cepina nur mehr 9% Ähnlichkeit mit St. Martin gemessen werden können.

2.3.3.5. Vergleich einiger Kennwerte der Karten 2–7

Zum Abschluss möchten wir anhand einer synoptischen Tabelle einige markante Kennwerte miteinander vergleichen, um die Wirkkraft der jeweiligen Teilcorpora prägnanter herausarbeiten zu können. Die in Abbildung 2 präsentierte Tabelle enthält in den Spalten 1–3 Angaben zum berücksichtigten Corpus, in den Spalten 4–9 sind ausgewählte Kennwerte und davon betroffene Messpunkte eingetragen.

Die Spalten 4 und 5 unserer Tabelle beziehen sich auf die in den jeweiligen Ähnlichkeitsprofilen aufgetretenen Minimalwerte. Die größten Unterschiede zum Gesamtkorpus (Min = 39%) scheinen bei den Teilcorpora Vokalismus (22%) und Morphologie (9%) auf, woraus wir schließen können, dass diese Bereiche die Gesamtprofilierung eher verzerren als die übrigen Teilcorpora. Die von den jeweiligen Minima betroffenen “Prügelknaben” liegen zweimal im Westen des Veneto, einmal im östlichen Trentino und je einmal im Oberengadin bzw. im Veltlin. Es handelt sich dabei immer um nicht nur innerlinguistisch, sondern auch räumlich weit vom Gadertal entfernte Messpunkte. Wie ein Vergleich aller 218 im *AD-I*-Corpus generierbaren Ähnlichkeitskarten gezeigt hat, liegen die insgesamt 10 negativen Antipoden, die von den anderen Messpunkten Antipathien (“Prügel”) beziehen, so gut wie immer in den beiden hier erwähnten Gebieten, und zwar sechs Mal im Einzugsbereich der Provinz Vicenza, drei Mal in der Bündnerromania und einmal im Gadertal. Die damit angesprochenen Dialekte zeichnen sich u.a. dadurch aus, dass sie entweder zur vom Italienischen sehr weit entfernten, bündnerromanischen bzw. alpin-lombardischen Gruppe oder zum besonders “italophilen” Typ des Venedischen gehören.

Spalte 1	2	3	4	5	6	7	8	9
(Teil-) Corpus	Anzahl <i>n</i> AK	siehe dazu	RIW- Min	“Prügel- knabe”	RIW- Max	“besten Freund”	PP. < Mw	PP. > Mw
Gesamt	3.302	Karte 2	39%	P. 178	89%	P. 84	142	75
Vokalismus	1.101	Karte 3	22%	P. 173	78%	P. 84	94	123
Konsonantismus	936	Karte 4	33%	P. 118	95%	P. 89	143	74
Lexikon	490	Karte 5	44%	P. 8	96%	P. 82	163	54
Morphologie	309	Karte 7	9%	P. 17	99%	P. 82	168	49

Abbildung 2: Tabelle mit markanten Kennwerten aus fünf Ähnlichkeitsprofilen zum Prüfbezugs-
punkt 83 San Martin de Tor / St. Martin in Thurn

In den Spalten 6 und 7 sind die maximalen Ähnlichkeitswerte vermerkt. Der Vokalismus (78%) trägt, wie schon bei den Minima, auch hier weniger stark zur Gesamtprofilierung (89%) bei als die übrigen Teilcorpora (95–99%). Konsonantismus, Lexikon und Morphologie hingegen sind hauptverantwortlich für die Stellung des Gadertalischen als dialektales Kerngebiet, was auch dadurch unterstrichen wird, dass die jeweils “besten Freunde” allesamt innerhalb des Tals, in vier von fünf Fällen sogar in direkter Nachbarschaft zu St. Martin zu liegen kommen. Mit den in den Spalten 8 und 9 vermerkten Werten sprechen wir indirekt den auf die Asymmetrie der gesamten Ähnlichkeitsverteilung Bezug nehmenden Kennwert der *Schiefe* an, auf den hier nicht näher eingegangen werden kann⁵³. Es sei nur soviel verraten: positive Schiefewerte (i.e. linkslastige Verteilungen = fett gedruckte Zellen in Spalte 8) repräsentieren Dialekte bzw. dialektale Kerngebiete, die sich gut gegen allfälligen Sprachausgleich abschirmen können, negative Schiefewerte (i.e. rechtslastige Verteilungen = fett gedruckte Zelle in Spalte 9) hingegen stehen tendenziell für Gebiete mit weitgehend durchgeführtem Sprachausgleich, die gut in das Gesamtnetz integriert sind. Dies ist bei unseren Ähnlichkeitsprofilen nur beim Vokalismus der Fall, wo 94 Messpunkte unter und 123 Messpunkte über dem Mittelwert liegen (= negativer Schiefewert). Der gadertalische Vokalismus wird also vom Gesamtnetz um vieles besser akzeptiert als etwa der Konsonantismus, der übrigens eine Ähnlichkeitsverteilung zeigt, die

⁵³ Cf. dazu GOEBL 2001, 863–864.

(was das Messpunktaufkommen unter- bzw. oberhalb des Mittelwerts betrifft) ziemlich exakt jener der Gesamtcorpus entspricht. Lexikon und Morphologie zeigen schließlich die positivste oder linkslastigste Schiefe aller Teilcorpora mit jeweils gut drei Viertel aller Messpunkte unter dem Mittelwert, was (immer unter Berücksichtigung der dazu oben gemachten, corpusbedingten Einschränkungen) darauf hindeutet, dass die stärksten Bastionen der gadertalischen Ladinität in diesen beiden Bereichen zu finden sind.

3. Schlussbemerkung und Ausblick

Die in diesem Beitrag vorgestellten Ähnlichkeitsprofile stellen natürlich nur eine erste, bescheidene Kostprobe all jener “Bilder” bzw. “Muster” dar, die aus unserem Corpus dank *VDM* in sehr komfortabler Weise generiert werden können. So fördert (wenn wir hier nur auf den themakartographischen Typ der Ähnlichkeitskarte Bezug nehmen und alle anderen, oben erwähnten Visualisierungen vorderhand außer Acht lassen) schon die Berücksichtigung von Prüfbezugspunkten aus der südlichen Dolomitenladinia aufschlussreiche (auch innerladinisch relevante) Raumkammerungen zu Tage, genauso wie die Einbeziehung bündnerromanischer und friaulischer Referenzorte Einblicke in die interdialektalen “Freundschafts-” bzw. “Feindschafts-”Relationen zwischen den drei Teilgruppen der rätoromanischen Sprachfamilie erlaubt. Von großer geolinguistischer Aussagekraft sind ferner Ähnlichkeitsprofile aus den lombardisch-trentinischen, trentinisch-venedischen und venedisch-friaulischen Übergangsbereichen, die im *AD-I* dokumentiert sind⁵⁴. Sie alle können dazu beitragen, die Erkenntnisse der rein qualitativ ausgerichteten, “traditionellen” Dialektologie im Lichte der durch Datenverdichtung gewonnenen DM-Ergebnisse näher zu beleuchten bzw. gegebenenfalls kritisch zu hinterfragen und zurechtzurücken. Da es sich bei unserem Projekt aber um “Work in Progress” handelt, darf der interessierte Leser an dieser Stelle auf eine Fortsetzung der *AD-DM*-relevanten Berichterstattung in einer der nächsten Nummern der Zeitschrift *Ladinia* sowie auf die zwischenzeitlich im Internet⁵⁵ zugängliche Projektdokumentation verwiesen werden.

Um Missverständnissen und Fehlinterpretationen vorzubeugen seien als Schlusswort nochmals jene Grundprinzipien in Erinnerung gerufen, die beim Betrachten und “Lesen” dialektometrischer “Bilder” keinesfalls aus den Augen (bzw. aus dem Kopf) verloren werden dürfen:

⁵⁴ Cf. dazu BAUER 2003.

⁵⁵ Projekt-Homepage: <http://ald.sbg.ac.at/dm/alddm/>.

1. die DM bezieht ihre Daten ausschließlich aus *Sprachatlanten*⁵⁶: dies impliziert, dass die in den Sprachatlaskarten mehr oder weniger treffsicher abgebildete dialektale “Realität” für dialektometrische Zwecke *tel quel* als repräsentativ akzeptiert wird. Somit steht und fällt die Qualität der dialektometrischen Auswertung mit der Qualität des untersuchten Sprachatlasses. Diese Prämisse gilt im Übrigen für jede Art von Klassifikation⁵⁷. Variationslinguistische Fragestellungen, die über die reine diatopische Dimension hinausgehen, können demnach auch nicht (bzw. nur schwer) mit dialektometrischen Methoden beantwortet werden. Ferner ist durch die Auswahl eines bestimmten Sprachatlasses vorgegeben, wie viele und welche *Objekte* (hier: Messpunkte bzw. Ortsdialekte) bzw. wie viele und welche *Eigenschaften* (hier: sprachliche Merkmale) beobachtet werden (können).
2. die DM verdichtet durch *Datenreduktion*: dies bedeutet, dass die Rohdaten im Rahmen der Taxierung simplifiziert und transformiert werden, wobei es einerseits zu Informationsverlusten kommt, die aber andererseits im Interesse einer *höherrangigen* Erkenntnis bewusst in Kauf genommen werden.
3. die DM misst *Ähnlichkeiten/Distanzen*: verfahrenstechnisch wird dabei ein bestimmtes Ähnlichkeitsmaß gewählt, in dessen Hintergrund jedoch eine in den Köpfen der Fachwissenschaftler (und Dialektometer) *a priori* vorhandene Vision von “dialektaler Ähnlichkeit/Distanz” steht. Insofern modelliert das gewählte Ähnlichkeitsmaß einen vorab existenten Ähnlichkeitsbegriff nach. Der in diesem Beitrag hauptsächlich eingesetzte RIW_{jk} simuliert z.B. die quantitative (bzw. sogar lineare) Vorstellung, dass zwei Dialekte einander umso ähnlicher sind, je mehr Merkmale sie gemeinsam haben.
4. die DM arbeitet, wie jede mustererkennende Disziplin, *Relationen* zwischen den beobachteten *Objekten* (Ortsdialekten) heraus: dabei werden möglichst viele *Eigenschaften* (Merkmale bzw. Merkmalsausprägungen) quantitativ in Synopse gesetzt, um vom Besonderen (einzelne Merkmale) auf das Allgemeine (Dialekttyp) schließen zu können. Es handelt sich somit um eine *induktive* Vorgehensweise.

4. Abkürzungsverzeichnis

AK	Arbeitskarte, taxierte Sprachatlaskarte
AIS	Atlante Italo-Svizzero / Sprach- und Sachatlas Italiens und der Südschweiz
ALD	Atlant linguistisch dl ladin dolomitich y di dialec vejins / Atlante linguistico del ladino dolomitico e dei dialetti limitrofi / Sprachatlas des Dolomitenladinischen und angrenzender Dialekte

⁵⁶ “Extra atlantes linguisticos nulla salus dialectometrica” (GOEBL 1991, 343).

⁵⁷ “[...] a classification is no better than the dimensions or variables on which it is based” (BAILEY 1994, 2).

ALF	Atlas linguistique de la France
ALG	Atlas linguistique de la Gascogne
CARD	CARtografia Dialettologica / CARTographie Dialectologique
DEM	Durchschnittliche Euklidische Metrik
DM	Dialektometrie
DVD	Digital Versatile Disc
FWF	Fonds zur Förderung der wissenschaftlichen Forschung (Wien)
GIW	Gewichtender Identitätswert
IRS	Index Retrieval System
j	Vektor des Messpunkts j
k	Vektor des Messpunkts k
(M)Max	(absolute) Maximum
Med	Median
(M)Min	(absolute) Minimum
Mw	arithmetischer Mittelwert
N	Anzahl der Objekte / Messpunkte / Ortsdialekte
OK	Originalkarte, originale Sprachatlaskarte
p	Anzahl der Attribute / Merkmale / Arbeitskarten
P(P).	Messpunkt(e)
RDW	Relativer Distanzwert
RIW	Relativer Identitätswert
RTF	Rich Text Format
VDM	Visual DialectoMetry

5. Literaturverzeichnis

- AAVV: *Arbeitsberichte 1–11 zum AD-I / Relazioni di lavoro 1–11 per l'AD-I*, in: “Ladinia”, X, 1986, 5–32; XI, 1987, 183–218; XII, 1988, 17–56; XIII, 1989, 185–229; XIV, 1990, 259–304; XV, 1991, 203–254; XVI, 1992, 169–184; XVII, 1993, 125–152; XIX, 1995, 307–330; XX, 1996, 191–221; XXIII, 1999, 281–301.
- ADANSON, M.: *Famille des plantes I*, Paris 1763.
- AD-I: GOEBL, H./BAUER, R./HAIMERL, E. (Eds.): *Atlant linguistisch dl ladin dolomitich y di dialec vejins, 1a pert. / Atlante linguistico del ladino dolomitico e dei dialetti limitrofi, 1a parte. / Sprachatlas des Dolomitenladinischen und angrenzender Dialekte, 1. Teil*, Wiesbaden 1998, 7 Bände.
- ALTMANN, G./LEHFELDT, W.: *Allgemeine Sprachtypologie. Prinzipien und Meßverfahren*, München 1973.
- APPEL, H.-W.: *Die Datenbank zum Forschungsprojekt Phonologie Niedersächsischer Dialekte und das Niedersächsische Dialektarchiv*, in: “Dialectologia et Geolinguistica”, 10, 2002, 13–30.
- ASCOLI, G. I.: *Saggi ladini*, in: “Archivio glottologico italiano”, 1, 1873, 1–556.
- BAILEY, K. D.: *Typologies and Taxonomies. An Introduction to Classification Techniques*, London / New Delhi 1994.

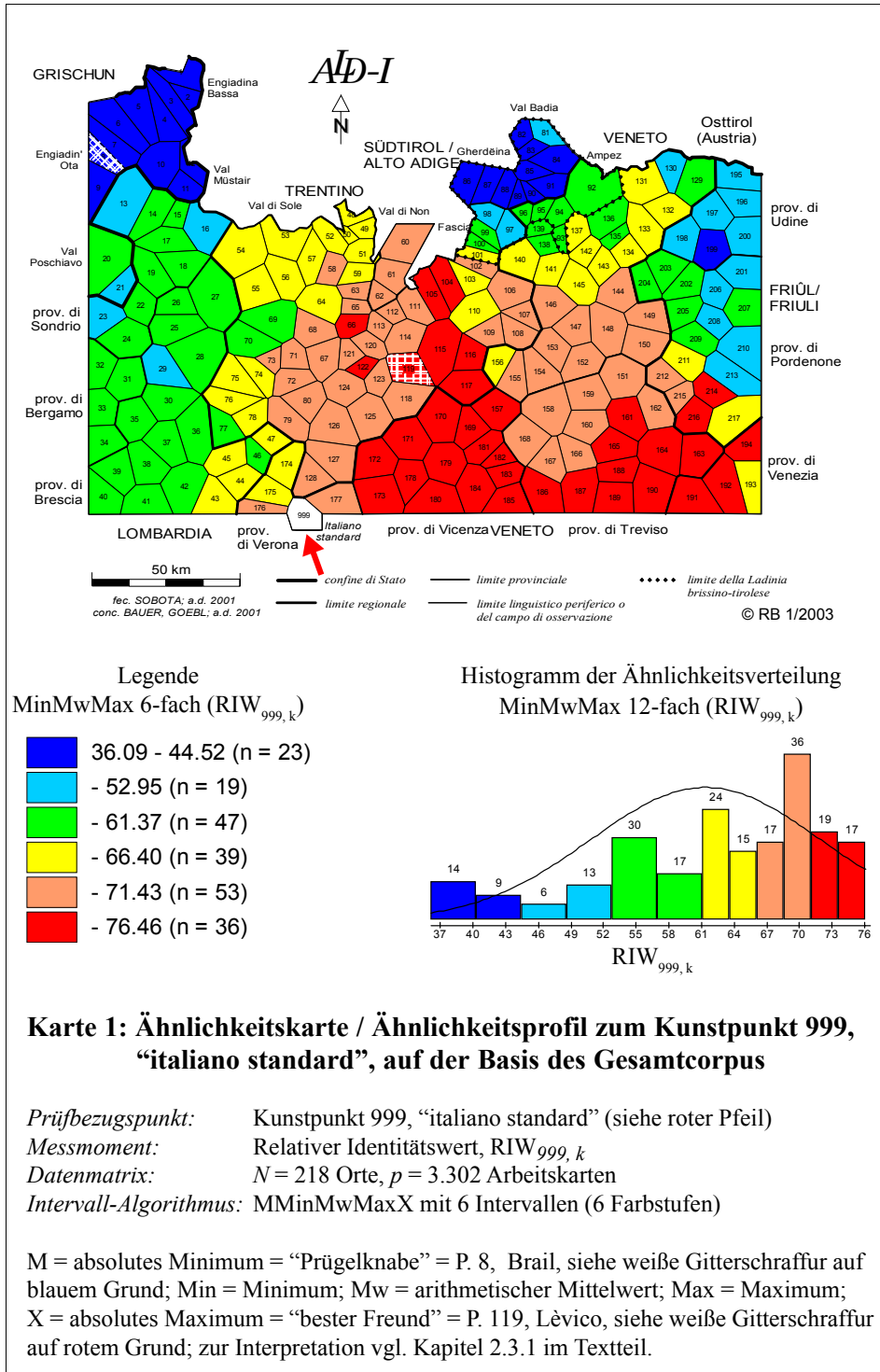
- BAUER, R.: *Peder, Pire, Piero: namenkundliche Splitter aus dem romanischen Alpenraum als Prolegomena zu einer Dialektometrisierung des AD-I*, in: ANREITER, P./ERNST, P./HAUSNER, I. (eds.), *Namen, Sprachen und Kulturen. Imena, Jeziki in Kulture. Festschrift für Heinz-Dieter POHL zum 60. Geburtstag*, Wien 2002, 79–99.
- BAUER, R.: *Sguardo dialettometrico su alcune zone di transizione dell'Italia nord-orientale (lombardo vs. trentino vs. veneto)*, in: BOMBI, R. / FUSCO, F. (eds.), *Parallela X, Sguardi reciproci*, Udine, 2003, 65–91.
- BAUER, R.: *Analyse- und Visualisierungsverfahren sprachgeographischer Daten am Ende des 20. Jahrhunderts*, in: DAHMEN, W. et al. (eds.), *Romanistik und Neue Medien. Romanistisches Kolloquium XVI*, Tübingen, im Druck.
- BAUER, R./GOEBL, H./HAIMERL, E.: *Der elektronische AD-I. CD-ROM 1: Kartographieprogramm CARD, Index Retrieval System IRS, Sprechender Sprachatlas: Karten 1–216; CD-ROM 2: Sprechender Sprachatlas: Karten 1–438; CD-ROM 3: Sprechender Sprachatlas: Karten 439–884*, Salzburg 1999–2000.
- BAUER, R./GOEBL, H./HAIMERL, E.: *AD-I-DVD. Sprechender Sprachatlas: Karten 1–884*, Salzburg 2002.
- BENINCA, P./VANELLI, L.: *Il plurale palatale in friulano: saggio di analisi autosegmentale*, in: *Scritti di Linguistica e Dialettologia in onore di Giuseppe FRANCESCATO*, Trieste 1995, 25–46.
- BERNHARD, G.: *Besprechung des AD-I*, in: "Dialectologia et Geolinguistica", 8, 2000, 93–95.
- BERSCHIN, H.: *Questione ladina. Grundrechnungsarten und Dialektometrie*, in: "Ladinia", XX, 1996, 187–189.
- BOCK, H. H.: *Automatische Klassifikation. Theoretische und praktische Methoden zur Gruppierung und Strukturierung von Daten (Cluster-Analyse)*, Göttingen 1974.
- BROSIUS, G.: *SPSS/PC+. Advanced Statistics und Tables. Einführung und praktische Beispiele*, Hamburg et al. 1989.
- D'AGOSTINO, M./PENNISI, A.: *Besprechung des AD-I*, in: "Rivista Italiana di Dialettologia", 24, 2000, 279–281.
- DINGELDEIN, H. J.: *Besprechung zu: KÖNIG, W./SCHRAMBKE, R., Die Sprachatlanten des schwäbisch-alemannischen Raumes, Baden 1999*, in: "Dialectologia et Geolinguistica", 10, 2002, 102–103.
- FRAU, G.: *Besprechung des AD-I*, in: "Incontri linguistici", 22, 1999, 215–221.
- GOEBL, H.: *Dialektgeographie + Numerische Taxonomie = Dialektometrie. Anhand rätoromanischer und oberitalienischer Dialektmaterialien (AIS)*, in: "Ladinia", IV, 1980, 31–95.
- GOEBL, H.: *Éléments d'analyse dialectométrique (avec application à l'AIS)*, in: "Revue de Linguistique Romane", 45, 1981, 349–420.
- GOEBL, H.: *Dialektometrische Studien. Anhand italo-romanischer, rätoromanischer und galloromanischer Sprachmaterialien aus AIS und ALF*, Tübingen 1984, 3 Bände.
- GOEBL, H.: *"Ma il distintivo necessario del determinato tipo sta appunto nella simultanea presenza e nella particolare combinazione di quei caratteri". Methodische und wissenschaftsgeschichtliche Bemerkungen zum Diskussionskomplex "Unità ladina"*, in: "Ladinia", XIV, 1990, 219–257.
- GOEBL, H.: *Una classificazione gerarchica di dati geolinguistici tratti dall'AIS. Saggio di dialettometrica dendrografica*, in: "Linguistica", XXXI/I, 1991, 341–352.

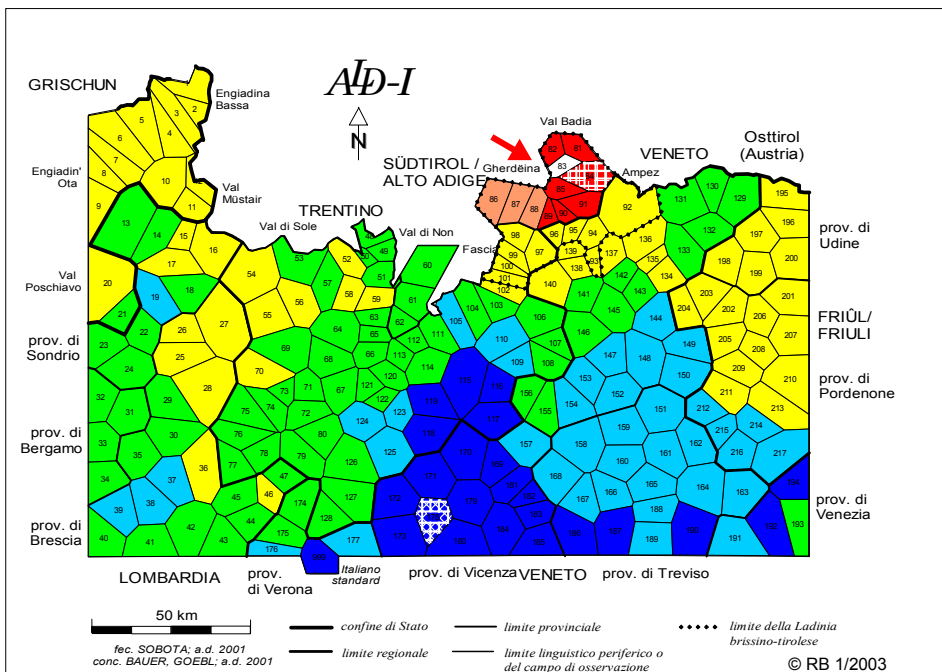
- GOEBL, H.: *Dendrogramme im Dienst der Dialektologie. Zwei hierarchisch-agglomerative Klassifikationen von Daten des Sprachatlases AIS*, in: KLENK, U. (ed.), *Computatio linguae*. Aufsätze zur algorithmischen und quantitativen Analyse der Sprache, Stuttgart 1992, 54–73.
- GOEBL, H.: *Spannungsverhältnisse in dialektalen Netzen. Ein Hinweis zu disziplinübergreifender Diskussion*, in: KLENK, U. (ed.), *Computatio Linguae II*, Stuttgart 1994, 63–83.
- GOEBL, H.: *Dialektometrische Beschreibung der Romania*, in: HOLTUS, G./METZELTIN, M./SCHMITT, C. (eds.), *Lexikon der Romanistischen Linguistik*, vol. 7: Kontakt, Migration und Kunstsprachen. Kontrastivität, Klassifikation und Typologie, Tübingen 1998, 977–1003.
- GOEBL, H.: *Die Germanismen im ladinischen Sprachatlas AD-I*, in: TATZREITER, H./HORNUNG, M./ERNST, P. (eds.), *Erträge der Dialektologie und Lexikographie. Festgabe für Werner BAUER zum 60. Geburtstag*, Wien 1999, 191–210.
- GOEBL, H.: *Dialektometrie*, in: HOLTUS, G./METZELTIN, M./SCHMITT, C. (eds.), *Lexikon der Romanistischen Linguistik. Band I/2: Methodologie (Sprache in der Gesellschaft / Sprache und Klassifikation / Datensammlung und -verarbeitung)*, Tübingen 2001, 856–874.
- GOEBL, H.: *Analyse dialectométrique des structures de profondeur de l'ALF*, in: "Revue de Linguistique Romane", 66, 2002a, 5–63.
- GOEBL, H.: *Sprachatlanten: woher? womit? wozu? Einige buntgemischte Reflexionen am Gartenzaun zwischen Romanistik und Germanistik*, in: ANREITER, P./ERNST, P./HAUSNER, I. (eds.), *Namen, Sprachen und Kulturen. Imena, Jeziki in Kulture. Festschrift für Heinz-Dieter POHL zum 60. Geburtstag*, Wien 2002b, 257–274.
- HAAG, K.: *Die Mundarten des oberen Neckar- und Donaulandes (schwäbisch-alemannisches Grenzgebiet: Baarmundarten)*, Reutlingen 1898.
- IANNÀCARO, G.: *Besprechung des AD-I*, in: "Mondo ladino", 24, 2000, 183–200.
- JABERG, K.: *Zum Atlas linguistique de la France*, in: "Zeitschrift für romanische Philologie", 30, 1906, 512.
- KELLE, B.: *Besprechung des AD-I*, in: "Zeitschrift für Dialektologie und Linguistik", 68, 2001, 247–252.
- KRAMER, J.: *Historische Grammatik des Dolomitenladinischen. Formenlehre*, Gerbrunn bei Würzburg 1976.
- KRAMER, J.: *Il problema storico-linguistico del ladino*, in: ZAMBONI, A./VIGOLO, M. T./CROATTO, E. (eds.), *Saggi dialettologici in area italo-romanza. Quinta raccolta*, Padova 2000, 35–50.
- KRAMER, J.: *Besprechung des AD-I*, in: "Zeitschrift für romanische Philologie", 117, 2001, 163–171.
- KREFELD, T.: *Der surselvische Wortschatz, die Questione ladina und die quantitative Arealtypologie*, in: "Ladinia", XVIII, 1994, 261–288.
- LIVER, R.: *Rätoromanisch. Eine Einführung in das Bündnerromanische*, Tübingen 1999.
- PENNISI, A.: *L'informatica per la dialettologia*, in: "Rivista Italiana di Dialettologia", 15, 1991, 137–164.
- SÉGUY, J.: *La dialectométrie dans l'Atlas linguistique de la Gascogne*, in: "Revue de Linguistique Romane", 37, 1973, 1–24.
- SKUBIC, M.: *Besprechung des AD-I*, in: "Linguistica", XL/1, 2000, 188–195.
- SNEATH, P. H. A./SOKAL, R. R.: *Numerical Taxonomy. The Principles and Practice of Numerical Classification*, San Francisco 1973.

- SOLÈR, C.: *Besprechung des A^hDⁱ*, in: “Annalas da la Societad Retorumantscha”, 115, 2002, 287–299.
- STINCHCOMBE, A. L.: *Constructing Social Theories*, New York 1968.
- THIESSEN, A. H.: *Precipitation Averages for large Areas*, in: “Monthly Weather Review”, 39, 1911, 1082–1084.
- VICARIO, F.: *Besprechung des A^hDⁱ-I*, in: “Ladinia”, XXIII, 1999, 261–265.
- VIDESOTT, P.: *Onomatometrie Ladiniens*, in diesem Band.

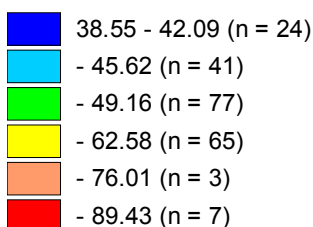
6. Kartenverzeichnis

- Karte 1: Ähnlichkeitskarte / Ähnlichkeitsprofil zum Kunstpunkt 999, “italiano standard”, auf der Basis des *Gesamtkorpus* (Seite 244).
- Karte 2: Ähnlichkeitskarte / Ähnlichkeitsprofil zum *A^hDⁱ-I*-Punkt 83, San Martin de Tor, auf der Basis des *Gesamtkorpus* (Seite 245).
- Karte 3: Ähnlichkeitskarte / Ähnlichkeitsprofil zum *A^hDⁱ-I*-Punkt 83, San Martin de Tor, auf der Basis des Teilcorpus *Vokalismus* (Seite 246).
- Karte 4: Ähnlichkeitskarte / Ähnlichkeitsprofil zum *A^hDⁱ-I*-Punkt 83, San Martin de Tor, auf der Basis des Teilcorpus *Konsonantismus* (Seite 247).
- Karte 5: Ähnlichkeitskarte / Ähnlichkeitsprofil zum *A^hDⁱ-I*-Punkt 83, San Martin de Tor, auf der Basis des Teilcorpus *Lexikon* (Seite 248).
- Karte 6: Ähnlichkeitskarte / Ähnlichkeitsprofil zum *A^hDⁱ-I*-Punkt 83, San Martin de Tor, auf der Basis des Teilcorpus *Lexikon* (Seite 249).
- Karte 7: Ähnlichkeitskarte / Ähnlichkeitsprofil zum *A^hDⁱ-I*-Punkt 83, San Martin de Tor, auf der Basis des Teilcorpus *Morphologie* (Seite 250).

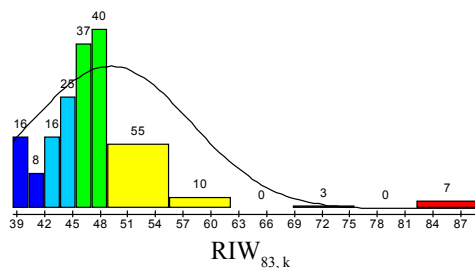




Legende
MinMwMax 6-fach (RIW_{83,k})



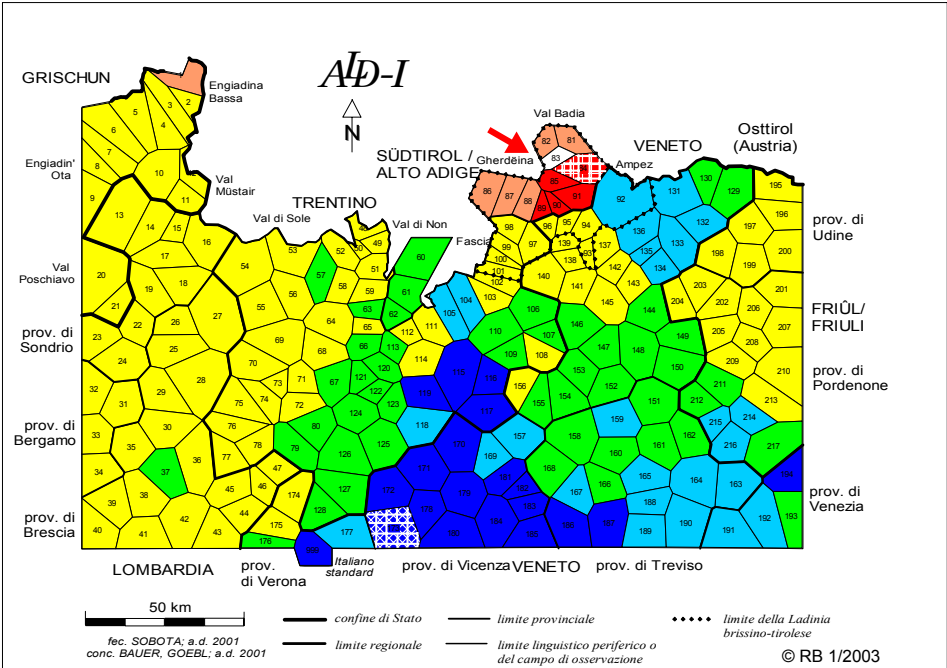
Histogramm der Ähnlichkeitsverteilung
MinMwMax 12-fach (RIW_{83,k})



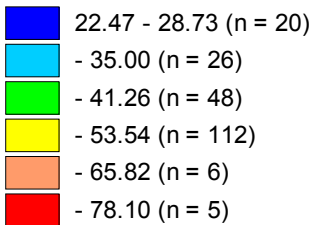
Karte 2: Ähnlichkeitskarte / Ähnlichkeitsprofil zum **ALD-I**-Punkt 83, San Martin de Tor, auf der Basis des Gesamtkorpus

Prüfbezugspunkt: P. 83, San Martin de Tor / St. Martin in Thurn (siehe roter Pfeil)
Messmoment: Relativer Identitätswert, RIW_{83,k}
Datenmatrix: N = 218 Orte, p = 3.302 Arbeitskarten
Intervall-Algorithmus: MMinMwMaxX mit 6 Intervallen (6 Farbstufen)

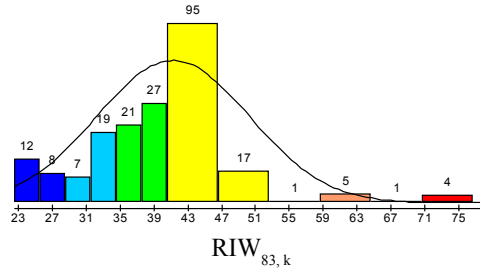
M = absolutes Minimum = “Prügelknabe” = P. 178, Schio, siehe weiße Gitterschraffur auf blauem Grund; Min = Minimum; Mw = arithmetischer Mittelwert; Max = Maximum; X = absolutes Maximum = “bester Freund” = P. 84, La Val, siehe weiße Gitterschraffur auf rotem Grund; zur Interpretation vgl. Kapitel 2.3.2 im Textteil.



Legende
MinMwMax 6-fach (RIW_{83,k})



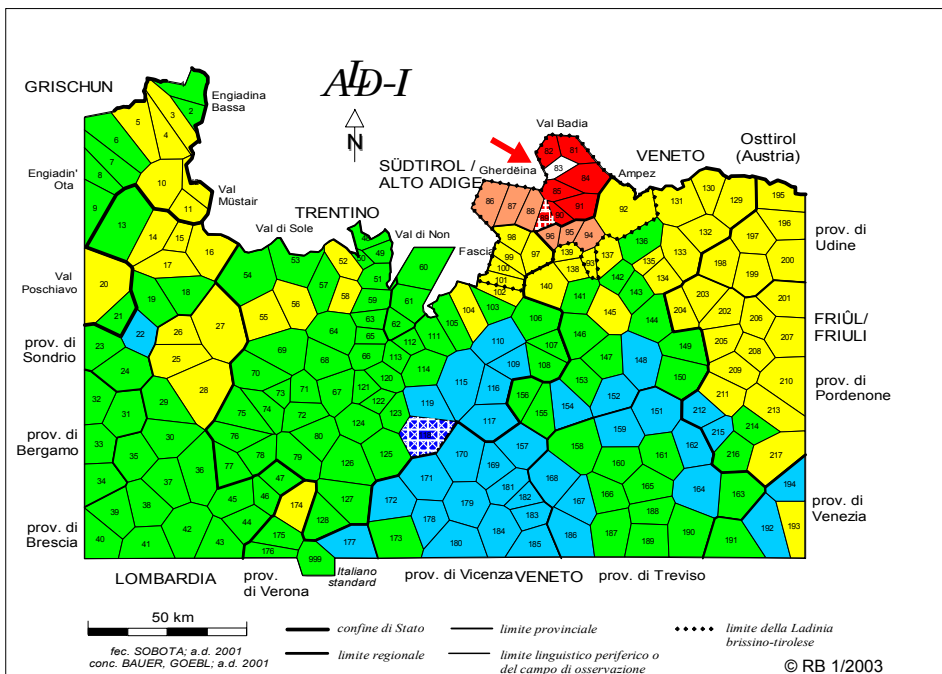
Histogramm der Ähnlichkeitsverteilung
MinMwMax 12-fach (RIW_{83,k})



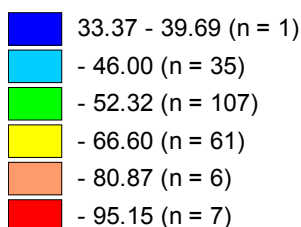
Karte 3: Ähnlichkeitskarte / Ähnlichkeitsprofil zum AD-I-Punkt 83, San Martin de Tor, auf der Basis des Teilcorpus Vokalismus

- Prüfbezugspunkt:* P. 83, San Martin de Tor / St. Martin in Thurn (siehe roter Pfeil)
Messmoment: Relativer Identitätswert, RIW_{83,k}
Datenmatrix: N = 218 Orte, p = 1.101 Arbeitskarten (Vokalismus)
Intervall-Algorithmus: MMinMwMaxX mit 6 Intervallen (6 Farbstufen)

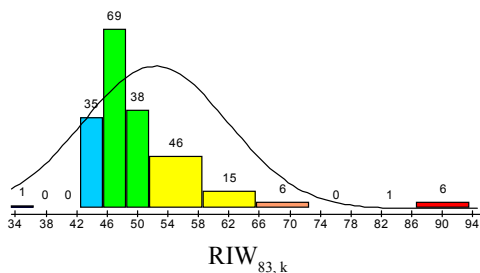
M = absolutes Minimum = “Prügelknabe” = P. 173, Recoaro, siehe weiße Gitterschraffur auf blauem Grund; Min = Minimum; Mw = arithmetischer Mittelwert; Max = Maximum; X = absolutes Maximum = “bester Freund” = P. 84, La Val, siehe weiße Gitterschraffur auf rotem Grund; zur Interpretation vgl. Kapitel 2.3.3.1 im Textteil.



Legende
MinMwMax 6-fach (RIW_{83, k})



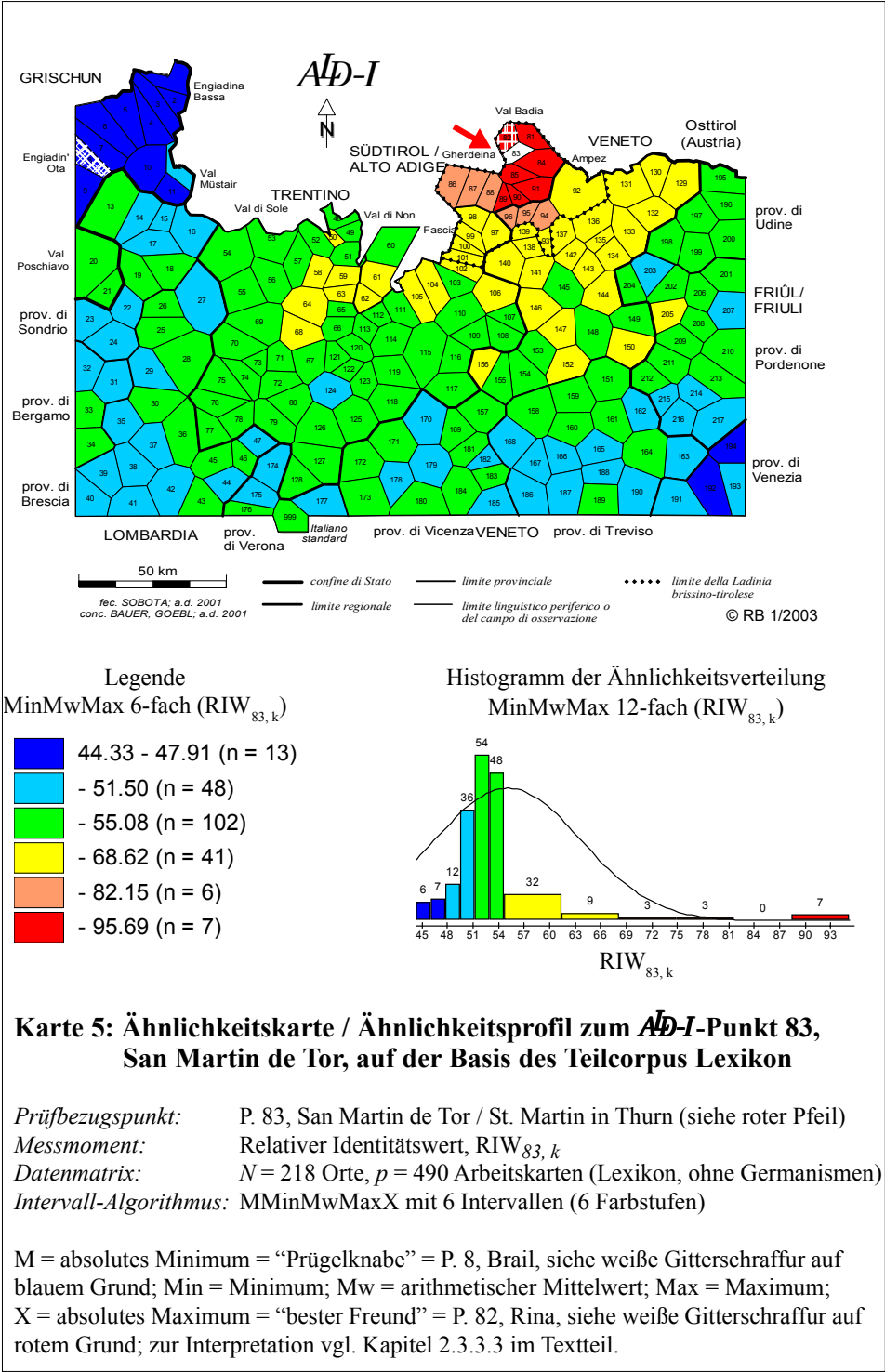
Histogramm der Ähnlichkeitsverteilung
MinMwMax 12-fach (RIW_{83, k})

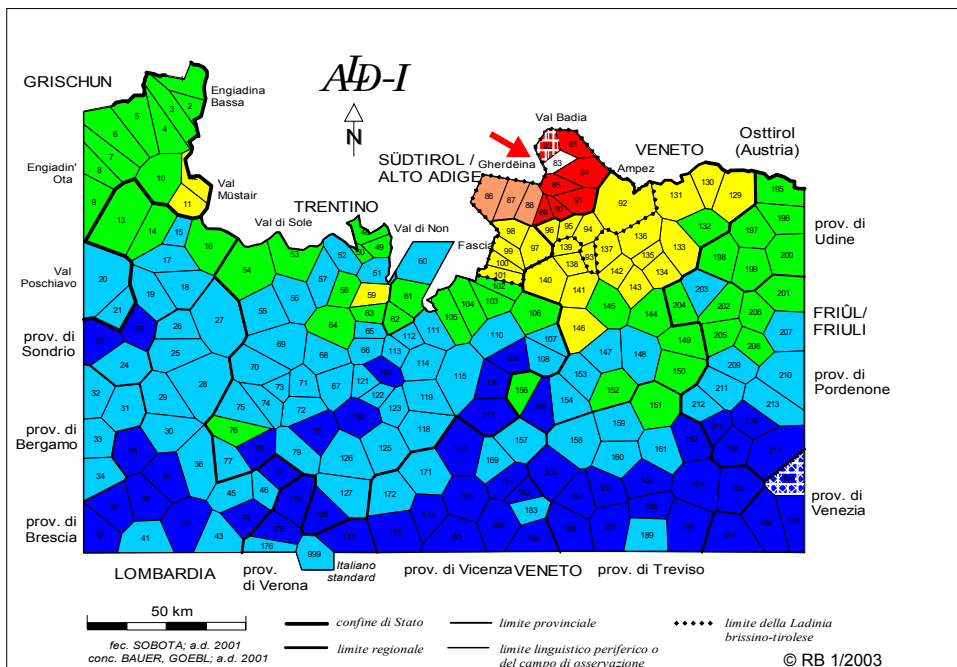


Karte 4: Ähnlichkeitskarte / Ähnlichkeitsprofil zum AD-I-Punkt 83, San Martin de Tor, auf der Basis des Teilcorpus Konsonantismus

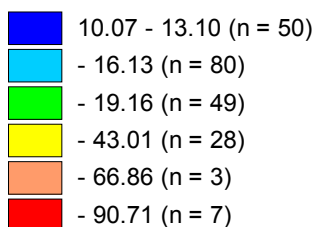
Prüfbezugspunkt: P. 83, San Martin de Tor / St. Martin in Thurn (siehe roter Pfeil)
Messmoment: Relativer Identitätswert, RIW_{83, k}
Datenmatrix: N = 218 Orte, p = 936 Arbeitskarten (Konsonantismus)
Intervall-Algorithmus: MMinMwMaxX mit 6 Intervallen (6 Farbstufen)

M = absolutes Minimum = “Prügelknabe” = P. 118, Lusern, siehe weiße Gitterschraffur auf blauem Grund; Min = Minimum; Mw = arithmetischer Mittelwert; Max = Maximum; X = absolutes Maximum = “bester Freund” = P. 89, Kolfuschg, siehe weiße Gitterschraffur auf rotem Grund; zur Interpretation vgl. Kapitel 2.3.3.2 im Textteil.

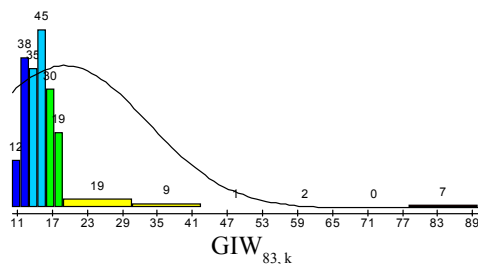




Legende

MinMwMax 6-fach (GIW_{83,k})

Histogramm der Ähnlichkeitsverteilung

MinMwMax 12-fach (GIW_{83,k})

Karte 6: Ähnlichkeitskarte / Ähnlichkeitsprofil zum AD-I-Punkt 83, San Martin de Tor, auf der Basis des Teilcorpus Lexikon

- Prüfbezugspunkt:** P. 83, San Martin de Tor / St. Martin in Thurn (siehe roter Pfeil)
- Messmoment:** Gewichtender Identitätswert, GIW_{83,k}
- Datenmatrix:** N = 218 Orte, p = 490 Arbeitskarten (Lexikon, ohne Germanismen)
- Intervall-Algorithmus:** MMinMwMaxX mit 6 Intervallen (6 Farbstufen)

M = absolutes Minimum = "Prügelknabe" = P. 194, Portogruaro, siehe weiße Gitterschraffur auf blauem Grund; Min = Minimum; Mw = arithmetischer Mittelwert; Max = Maximum; X = absolutes Maximum = "bester Freund" = P. 82, Rina, siehe weiße Gitterschraffur auf rotem Grund; zur Interpretation vgl. Kapitel 2.3.3.3 im Textteil.

