

Gewählte Methodik

Um die skizzierten Projektziele zu erreichen, sind folgende *Arbeitsschritte* nötig (in chronologischer Reihenfolge):

- a) Bibliographische Aufarbeitung des aktuellen Forschungsstandes zur Dialektometrie (innerhalb und außerhalb der Romanistik). Abfassung eines **Forschungsberichts** in Form einer kommentierten Bibliographie.
- b) Begleitende, externe **Programmiertätigkeit** zur Verbesserung bestehender Software-Lösungen (u.a. Implementierung einer Schnittstelle zwischen *IRS-Index Retrieval System*, dem Datenanalyse-Tool des ALD I, und dem Dialektometriepaket VDM) sowie zur Erweiterung des Vorrats an dialektometrischen Methoden und Verfahren (Upgrades bestehender Systemkomponenten).
- c) Methodische Prämisse: Erstellung eines standardisierten **Kodierungsschemas** zur Taxierung aller Karten des ALD I. In diesem Zusammenhang müssen zunächst alle in Frage kommenden, meist lateinischen Nexus (i.e. phonetische, morphologische oder lexikalische Merkmale) isoliert und zusammen mit den möglichen Merkmalsausprägungen in den untersuchten romanischen Dialekten (Bündnerromanisch, Lombardisch, Trentinisch, Dolomitenladinisch, Venedisch, Friulanisch) in Form einer Datenbank erfasst werden. In weiterer Folge wird jedem Nexus/Merkmal und den dazugehörigen Ausprägungen ein eindeutiger Code zugewiesen. Die solcherart generierte, *a priori* offene Tabelle dient sowohl als Grundlage für die Taxierung der Sprachatlaskarten/Kodierung der Taxierungserträge als auch als wesentliches Hilfsmittel für eine differenzierte dialektometrische Analyse, da die vergebene Kodierung den Zugriff auf Subcorpora, also die Auswahl von Untermengen des Gesamtkorpus je nach linguistischem Erkenntnisinteresse erlaubt (z.B. rein lexikalisch basiertes Teilcorpus; konsonantisch vs. vokalisch relevante Daten; Eliminierung bestimmter Merkmalsgruppen aus dem Gesamtkorpus).

Darüber hinaus kann das Kodierungsschema als Kontrolltabelle fungieren, anhand derer die Codes aller Taxierungen (und somit die Taxierungen selbst) auch noch *a posteriori* transparent nachvollzogen/verifiziert werden können, was bei bislang publizierten Auswertungen von Sprachatlanten in der Regel kaum möglich ist.

- d) **Computergestützte Taxierung** idealiter aller Merkmalsträger (i.e. aller Sprachatlaskarten) des ALD I nach möglichst vielen einzelnen Merkmalen. In EDV-technischer Hinsicht soll hier hauptsächlich das im Rahmen des FWF-Projekts ALD I entwickelte *Index Retrieval System* (IRS) zum Einsatz kommen (und erweitert werden), dessen Tauglichkeit bereits im Vorfeld getestet werden konnte.
- e) parallel zu d): **Manuelle Taxierung** von unter d) nicht abarbeitbaren Sprachatlaskarten. In weiterer Folge können die Erträge bereits (computergestützt und manuell) taxierter Teilcorpora in Form von Dichtekarten visualisiert werden.
- f) Standardisierte **Erfassung**, i.e. Eingabe aller in den Arbeitsschritten d) und e) gewonnenen Taxierungserträge/Taxate in Datenmatrizen zur Generierung einer (möglichst plattformunabhängigen) Datenbank sowie deren **Korrektur**.
- g) **Verrechnung** der solcherart erstellten Datenmatrix: Erstellung von Ähnlichkeits- bzw. Distanzmatrizen unter Einsatz verschiedener Maße; numerische Klassifikation am PC mittels Tools aus dem im Rahmen der FWF-Projekte "Dialektometrie" erstellten Programmpaket VDM (VisualDialectoMetry).
- h) "Zeitgemäße" (i.e. benutzerfreundliche, heuristisch hochwertige) **Visualisierung** der unter g) erzielten Verrechnungsergebnisse (Ähnlichkeitsprofile, Synopsekarten, Isoglossenkarten, Dendrogramme usw. unter Einbindung gängiger GIS-Systeme).
- i) **Interpretation** der Ergebnisse; Überprüfung von Hypothesen; Abfassung des Textteils.