

VerbaAlpina – Der alpine Kulturraum im Spiegel seiner Mehrsprachigkeit

Thomas Krefeld, Stephan Lücke

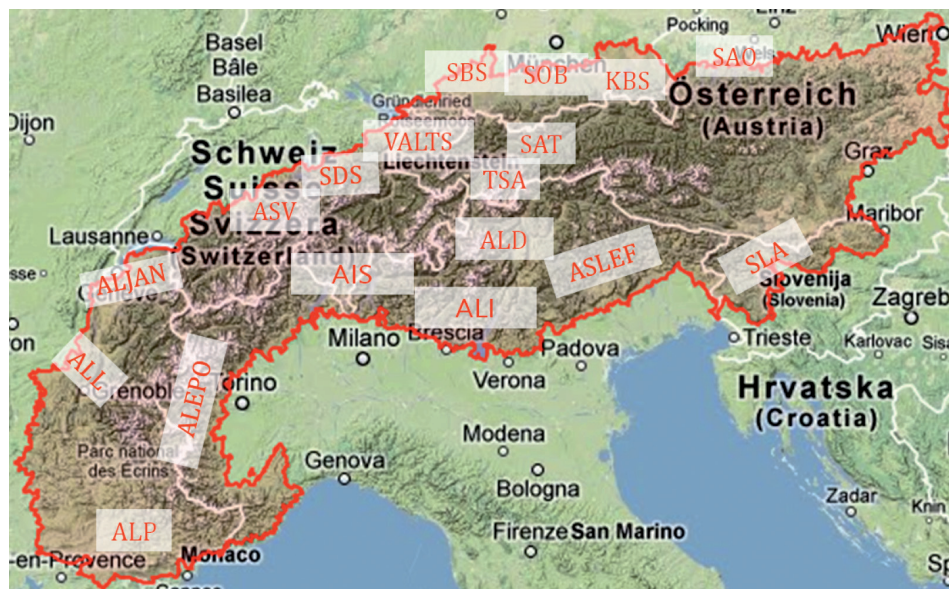
MEDIA MUTANTUR ET GEOGRAPHIA LINGUISTICA MUTATUR IN ILLIS

1. Sprachareale, Wortareale, Sachareale – und ihre Inkongruenz

Der Alpenraum zeichnet sich durch die – typisch europäische – Verbindung einer weit gehenden topo- und ethnographischen Homogenität (Geländeformationen, Fauna, Flora, Almwirtschaft, traditionelle Architektur, moderner Wintersport) mit einer ausgeprägten sprachlichen Heterogenität aus; denn er beheimatet, jeweils in Gestalt sehr unterschiedlicher Dialekte, romanische Sprachen (Französisch und Frankoprovenzalisch, Okzitanisch, Bündnerromanisch, Ladinisch, Friaulisch, Italienisch), das Deutsche und das Slowenische. Dieses komplexe Sprachgebiet ist zu einem guten Teil sprachgeographisch erschlossen¹

Wie Karte 1 aber ebenfalls zeigt, ist die ethnolinguistische Dokumentation ebenso parzelliert wie die Sprachen und Varietäten selbst; großräumig sind nur der AIS und der ALI, die wiederum das spezifische Alpenvokabular nur ansatzweise erfassen können. Daher ist es nicht leicht, sich einen ethnolinguistischen Überblick zu verschaffen. Gerade das wäre jedoch wünschenswert, denn bemerkens-

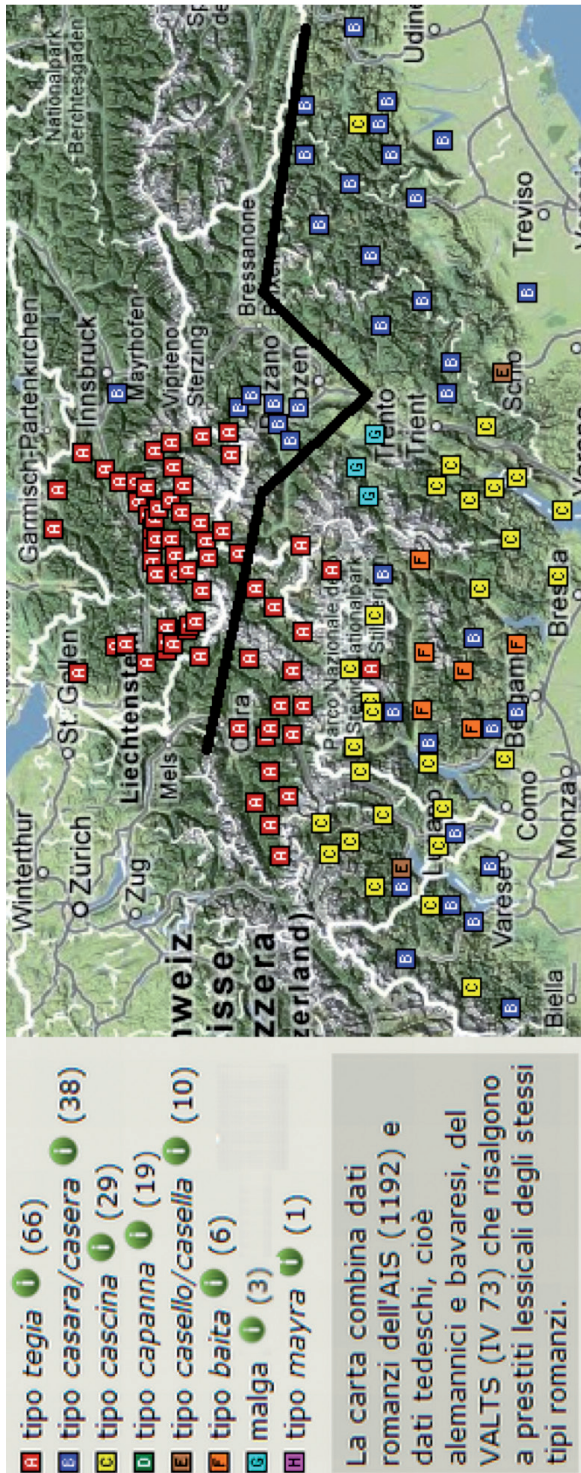
¹ Siehe Karte 1. Der *Atlas linguistique de la France* (ALF) ist hier bewusst nicht berücksichtigt. Die dort abgefragten Konzepte sind aus Sicht von *VerbaAlpina* wenig ergiebig. Gleiches gilt für die digitalen Atlanten ALAVAL und VIVALDI.



Karte 1: Erschließung des alpinen Kulturrums durch Sprachatlanten. Die rote Grenzlinie stellt den Perimeter der Alpenkonvention (<www.alpconv.org>) dar.

wert ist ja die Tatsache, dass Sprachareale, Wortareale und Sachareale häufig gerade nicht deckungsgleich sind: Wortareale, insbesondere im Blick auf spezifisch alpine Realien, überschreiten oft die Sprach- und Sprachfamiliengrenzen; gleichzeitig bezeichnen etymologisch identische Worttypen in ihren Verbreitungsarealen jedoch keineswegs immer dieselben Sachen, vor allem dann nicht, wenn sie auf Artefakte oder kulturell relevante Naturalia referieren. Diese Inkongruenzen und andere Fragen, die uns unten beschäftigen werden, lassen sich ausgehend vom folgenden Ausschnitt einer digitalen Karte CASCINA DI MONTAGNA / SENNHÜTTE² des AdIS (*Atlante linguistico digitale dell'Italia e della Svizzera meridionale*, <www.adis.gwi.uni-muenchen.de>) aufzeigen, in der Daten des AIS und des VALTS berücksichtigt wurden.

² In den Kartentiteln vorgegebene Konzepte werden in dieser Arbeit in Majuskeln gesetzt; Bedeutungsangaben stehen zwischen doppelten Anführungszeichen.



Karte 2: Sprachgrenzüberschreitende Wortareale am Beispiel SENNHÜTTE.³

³ Ausschnitt aus <http://www.adis.gwi.uni-muenchen.de/AIS.php?karte=true> → cartografia interattiva → lessico → LA CASCINA DI MONTAGNA (AIS 1192); die- se und die folgenden Karten werden im Zusammenspiel von servergespeicherten Daten und Geoinformationsdiensten (z.B. Google Earth) jeweils auf den Computern (*Clients*) der Internetbenutzer erzeugt; sie existieren jedoch nicht als eigenständige, nutzerunabhängige Graphiken.

Die (noch unvollständige) Karte zeigt eine Reihe von Heteronymen, die sich komplementär im Raum verteilen. In wenigen Fällen sind auch zwei Typen am selben Ort belegt, als ob es sich um Synonyme handelte.⁴ Die Heteronyme führen hier auf unterschiedliche sprachliche Strata zurück: Vorrömisch sind  *malga*⁵ und vielleicht  *baita*;⁶  *tegia* ist keltisch;⁷ lateinisch sind  *casara/casera*,⁸  *cascina* (zu lat. CAPSA “Gefäß” gekreuzt mit lat./rom. CASA “Hütte, Haus”)⁹ und wahrscheinlich  *mayra* (zu lat. MIGRARE “wandern”).¹⁰ Ihre Verbreitungsareale überschreiten teils innerromanische Sprachgrenzen (zwischen Bündnerromanisch, Ladinisch, Friaulisch und Italienisch), teils aber auch die romanisch-germanische Sprachfamiliengrenze. Im alemannischen und bairischen Sprachraum handelt es sich dann entweder um Relikte aus Substraten oder um Entlehnungen aus dem Romanischen, die nach der Germanisierung erfolgt sind.

Einen interessanten Spezialfall bilden die Wortareale, die auf inneralpine Migration zurückgehen, wie im Fall der Walser. Ein Beispiel zeigt die Verbreitung des vorrömischen Typs *balma(a)* “geschützte Stelle für Mensch und Vieh unter einem überhängenden Fels” (cf. z.B. Abb. 1). Die zugehörige Karte zeigt die Verbreitung, wie sie im AIS und im VALTS dokumentiert ist. In Wahrheit erstreckt sie sich nach Westen (Katalonien)¹¹ und Norden weit über das im AIS erfasste Gebiet hinaus. Allerdings zeigen die AIS-Daten auch, dass die Bezeichnung den Ostalpen fremd zu sein scheint; die Walser in Graubünden (Klosters), Liechtenstein und Vorarlberg (Lech) haben sie wohl in ihrer Ursprungsgegend, dem Wallis, aus dem Romanischen

⁴ In diesen Fällen stehen die beiden Kartensymbole direkt senkrecht übereinander.

⁵ Ein prototypisches Alpenwort; “si presta alle più varie congetture” (DELI 705; dort bibliografische Angaben).

⁶ Der DELI 105 verweist auf friaulisch *uàite, vùàite* “casotto di guardia” und schlägt ein germanisches Etymon, analog zu althochdeutsch *wabta* “Wache” vor; allerdings sind auch alemannisch *Beitz* und bairisch *Boaz(n)* “Wirtshaus” zu berücksichtigen, die nicht auf *wabta* zurückgehen können. Die areale Verteilung lässt eher an ein Substratwort denken.

⁷ Cf. REW 8616a.

⁸ Laut DELI 213 < latino CAESEARIA(M) “luogo ove si fa il caccio”, zu lat. CASEUS “Käse”.

⁹ Laut DELI 213 < “lat. parl. **capsia(m)*, per il lat. class. *capsa(m)* ‘cassa, recipiente’.” – Da die Formen des Typs *cascina* gelegentlich in derselben Gegend mit Formen des Typs *casina* (mit [z]) konkurrieren, liegt es nahe, eine Kreuzung mit dem Diminutiv von CASA (CASINA) im Sinne einer volksetymologischen Reinterpretation anzunehmen.

¹⁰ Diese Herkunft schlägt der *Vocabolario Treccani* vor (DURO 1997², <<http://www.treccani.it/vocabolario/maira/>>, letzte Abfrage 07.10.2014).

¹¹ Eine Google-Bildersuche ergibt zahlreiche Photos von *balma* genannten überhängenden Felsen, unter die oft Häuser, Kirchen oder Klöster gebaut wurden; z.B. *Santuario de la Balma* (<http://es.wikipedia.org/wiki/Santuario_de_la_Virgen_de_la_Balma>, letzte Abfrage 07.10.2014).

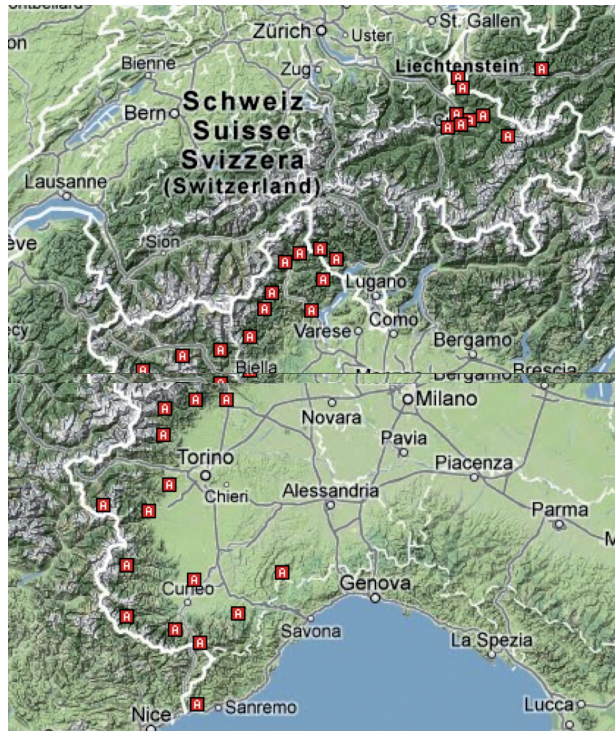
Abb. 1: *Die Balma della vecchia* im Val di Nibbio bei Verbania, Piemont (Quelle: <<http://in-valgrande.it/vegia/vegia.html>>, letzte Abfrage 18.02.2013).

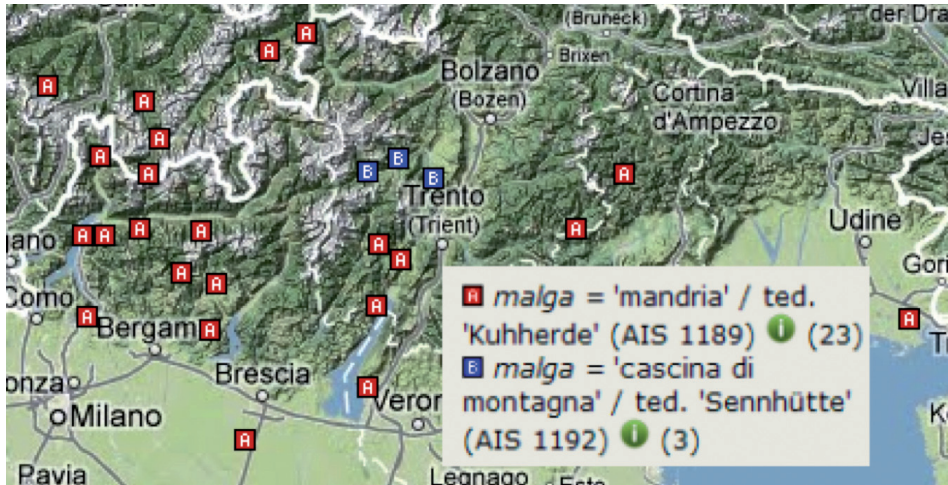


übernommen, so dass sie mit den Walsern und dem Walserischen nach Liechtenstein und Vorarlberg gekommen ist und dort keineswegs aus dem regionalen romanischen Substrat stammt. Das Wort

fehlt im Bündnerromanischen vollkommen; jedenfalls gibt es keinen Eintrag im umfassenden neuen Wörterbuch von DECURTINS 2012, und bei SCHORTA (1964, 31) heißt es s.v. *balma*: “Fehlt in R[omanisch]B[ünden] als Appell.; als Fl[ur]N[ame] weist es auf vorübergehende Anwesenheit von Walsern, die das Wort im Wallis aus dem Frankoprov. entlehnt haben.” Bemerkenswert sind allerdings die nicht-walserischen Belege im Churer Rheintal (Untervaz, Pfäfers, Vätis) und im Prättigau (Schiers), denn Walservörter verbreiten sich in der Regel nicht. In diesem Fall zeigt sich in ihrer Verbreitung allerdings die Entlehnungsmotivation, die wohl auch schon die Übernahme des Substratworts durch die Römer und später die Walser erklärt: Es ist kognitiv sinnvoll, für die charakteristische und nutzbare Felsformation eine spezifische Bezeichnung zu haben.

Karte 3: Verbreitung des Typs vorrömisch *balma(a)* in den romanischen Westalpen (nach AIS 424a) und bei den Walsern in Liechtenstein und Vorarlberg (nach VALTS IV, 71).





Karte 4: Bedeutungsunterschiede des Wortes *malga*.

Die erwähnte semantische Variation der von Atlanten erfassten Bezeichnungen lässt sich ebenfalls ausgehend von Karte 2 gut zeigen, denn der dort belegte Typ *malga* “Sennhütte” bezeichnet in metonymischer Relation zu SENNHÜTTE sehr oft die KUHHERDE (cf. AIS 1189).

Die genannten Beispiele wurden gewählt, weil sie eine doppelte sprachgeographische Herausforderung markieren, nämlich einerseits ein möglichst engmaschiges und großräumiges Netz von Aufnahmepunkten zu knüpfen und andererseits die erfassten Bezeichnungen auf eine möglichst präzise und übersichtliche Weise semantisch zu beschreiben. Beide Herausforderungen sind alles andere als neu, aber die Neuen Medien liefern durchaus neue Antworten, die im Folgenden unter den plakativen Schlagworten “Georeferenzierung” und “Crowdsourcing” skizziert werden sollen.¹² Vor diesem Hintergrund wurde von den Verfassern des vorliegenden Artikels das Projekt *Verba Alpina*¹³ entworfen, dessen Konzept hier zu Grunde liegt.

¹² Cf. zur digitalen Sprachgeographie die Beiträge in ELSPASS/KÖNIG 2008 sowie KREFELD 2009.

¹³ Der Hauptausschuss der *Deutschen Forschungsgemeinschaft* (DFG) hat am 01.07.2014 die Bewilligung einer ersten dreijährigen Förderperiode im Rahmen einer als Langzeitprojekt auf insgesamt neun Jahre angelegten Projektlaufzeit ausgesprochen. Die Projektarbeit wird spätestens am 01.01.2015 beginnen.

2. Georeferenzierung – bringt zusammen, was zusammengehört

Grundgedanke dieses Vorhabens ist es, die Abgeschlossenheit traditioneller gedruckter Publikationen zu überwinden und ein Portal zu schaffen, in das kontinuierlich möglichst viele georeferenzierte sprachliche Daten aus dem Alpenraum¹⁴ eingebracht werden. Darin sollen insbesondere drei Typen von Quellen zusammenfließen:

- [1] Retrodigitalisierung einschlägiger Arbeiten, d.h. der bereits publizierten Atlanten und Wörterbücher;¹⁵
- [2] Transfer aus parallelen Projekten;¹⁶
- [3] Neuerhebung von Daten durch *social media*.

¹⁴ Das Untersuchungsgebiet von *VerbaAlpina* deckt sich mit dem Gebiet der Alpenkonvention (siehe Karte 1). Die Erfassung von Atlas- und Wörterbuchdaten beschränkt sich auf die jeweiligen – und jeweils sämtlichen – Erhebungspunkte innerhalb des Perimeters der Alpenkonvention. Eine entsprechende Selektion ist mit Hilfe der raumbezogenen Datenformate und Funktionen der eingesetzten *MySQL*-Datenbank problemlos möglich und für eine Anzahl von Sprachatlanten (AIS, ALD, ALEPO, ASLEF, SDS, TPPSR und VALTS) bereits erfolgt. Die Ausrichtung des Projektes ist primär sprachwissenschaftlich und in der Sprachwissenschaft lexikalisch orientiert. Aufgrund der im Rahmen sprachwissenschaftlicher Analyse erforderlichen systematischen Bedeutungerschließungen und Klärungen von Etymologien erfährt das Projekt gleichsam von selbst und ganz organisch eine Erweiterung um ethnographische und historio-graphische Daten, die ebenfalls in strukturierter Form Eingang in die zentrale Datenbank finden. Auf diese Weise wird der Alpenraum in einer Vielzahl von Aspekten erschlossen, so dass eine Art Panoptikum dieses Kulturraumes entsteht. Die fächerübergreifende Qualität des gesammelten Datenmaterials verlangt bereits während der Entwicklungsphase interdisziplinären Austausch. Mittel- und langfristig bilden das entwickelte System und die gesammelten Daten eine gute Grundlage für interdisziplinär ausgerichtete Nachfolgeprojekte.

¹⁵ Die Daten werden entweder mittels OCR oder durch Abtippen erfasst und eindeutig kodiert (ASCII bzw. UTF-8/Unicode) sowie in strukturierter Form plattformunabhängig in einer *MySQL*-Datenbank gespeichert. Durch den Einsatz von *Client-/Server*-Prinzip und Browsertechnologie sind sämtliche Daten und Inhalte weltweit und betriebssystemunabhängig (*Windows*, *Mac/Unix*) verfügbar. Bei der Erfassung der Wörterbuch- und Atlasdaten werden lautschriftliche Besonderheiten (Basiszeichen, Diakritika) grundsätzlich zunächst in serialisierte ASCII-Zeichenfolgen transkodiert und in dieser Form in der Datenbank gespeichert (z.B. tsó => tso-/ [TPPSR I, Col. 1-3, Chevroix: “chaud”]). Für jede erfasste Publikation dokumentieren Zuordnungstabellen die jeweils verwendeten Transkodierungen. Darüber hinaus werden für die verschiedenen Transkodierungen Ersetzungsroutinen entwickelt, die die uniforme Ausgabe in IPA (Internationales Phonetisches Alphabet) leisten. IPA fungiert dabei als Referenztranskription, die publikationsübergreifende Vergleichbarkeit ermöglicht. Durch Speicherung der Daten in strukturierter Form sowie in ASCII- bzw. UTF-8/Unicode-Kodierung bei gleichzeitiger Dokumentation der Transkodierungen ist die Ausgabe des kompletten Datenbestands jederzeit in nahezu beliebigen Datenformaten möglich, darunter auch in XML. Auf diese Weise sind die Kriterien der Langzeitarchivierung erfüllt.

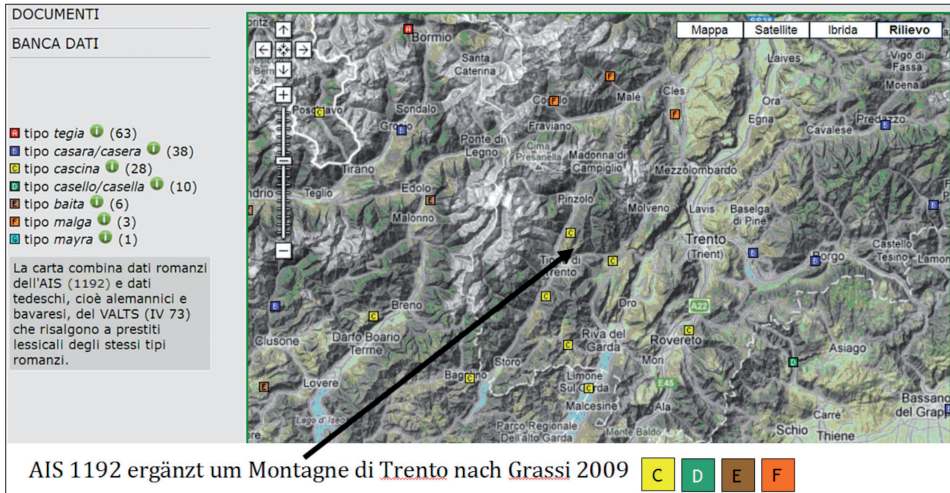
¹⁶ Von großem Wert sind hier vor allem für das Internet entwickelte Regionalprojekte, wie z.B. das von Hans GEISLER für die Val Chiavenna und das Bergell (<<http://alplomb.phil.hhu.de/>>). Grundsätzlich werden Kooperation und Datenaustausch mit allen laufenden Projekten angestrebt, die sich mit dem Alpenraum beschäftigen – wobei der Schwerpunkt zunächst auf sprachwissenschaftlich ausgerichteten Vorhaben liegt.

Die Aufgabe besteht also darin, einen bereits verfügbaren Altbestand von Daten (Typen [1] und [2]) in mehrfacher Hinsicht durch Erhebung eines Neubestands (Typ [3]) zu erweitern. Die entscheidende Funktion bei der Integration kommt der Georeferenziertheit der sprachlichen Daten zu, denn sie erlaubt ihre Verknüpfung mit *online* verfügbaren Geoinformationssystemen, wie z.B. *Google Maps* oder *Open Street Maps* und dadurch eine beliebig skalierbare kartographische Abbildung, die kontinuierlich durch neue Informationen in Gestalt von Primär- oder Metadaten erweitert werden kann. Dies gilt selbstverständlich auch für nichtsprachliche, aber georeferenzierbare Daten zur Geschichte, Demographie und Ethnographie, die in beliebig feiner Granulierung auch in multimedialer Gestalt (Ton, Bild) eingebunden werden können. Die Georeferenzierung leistet somit einen substanziellen und unentbehrlichen Beitrag, der die Sprachgeographie, sozusagen, in eine “Geolinguistik 2.0” verwandelt (cf. KREFELD/LÜCKE 2014).

Auch diese vielversprechende Perspektive wirft jedoch wiederum neue Probleme auf. Eine grundsätzliche Frage betrifft die Kombinierbarkeit von Daten aus unterschiedlichen Forschungs- und Erhebungstraditionen. So erlaubt es die Retrodigitalisierung [1], georeferenzierte Atlas- und Wörterbuchdaten (insofern sie ebenfalls georeferenzierbar sind) nebeneinanderzusetzen, um ein möglichst engmaschiges Netz von Aufnahmeorten zu schaffen. Lokal sehr ergiebig für diesen Zweck ist zum Beispiel Corrado GRASSIS Wörterbuch des Ortsdialekts von Montagne di Trento (GRASSI 2009). Wenn wir die auf Karte 2 belegten Bezeichnungstypen dort suchen und um einen auf dem Ausschnitt von Karte 2 nicht ausgewiesenen Typ ergänzen, ergibt sich das in Karte 5 dargestellte Bild.

Wie man sieht, passt der im Wörterbuch erfasste Ortsdialekt einerseits genau ins räumliche Atlasbild, insofern er den charakteristischen Typ seiner Umgebung (= , *cascina*) ebenfalls kennt; die Wörterbuchdokumentation ist jedoch

Im Hinblick auf die vielen Anknüpfungspunkte der Sprachwissenschaft zu anderen Disziplinen ist jedoch auch an Kooperation und Datenaustausch mit nicht-sprachwissenschaftlichen Institutionen und Projekten gedacht (etwa Archäologie, Ethnographie, Demographie, Wirtschaftswissenschaften, Geographie, Statistik). Rein praktisch ist entsprechender – selbstverständlich bidirektional geplanter – Datenaustausch durch Einsatz der modernen Medien kein substantielles Problem. Für die wichtige und anspruchsvolle Aufgabe des Abgleichs bzw. der Standardisierung von Datenformaten und -kodierungen der auszutauschenden Daten sowie zur Kontaktpflege und als Ansprechpartner wurde für die gesamte Projektlaufzeit mindestens eine Wissenschaftlerstelle beantragt.



Karte 5: Erweiterung der Karte 2 um georeferenzierte Wörterbuchbelege aus Montagne di Trento.

andererseits grundverschieden, da sie anstatt einzelner Lexeme ein kleines, in sich differenziertes Feld verwandter Bezeichnungen belegt (C, D, E, F).

Dieses Ergebnis scheint kein Zufall zu sein, denn es lässt sich durch andere Wörterbücher bestätigen (z.B. DE AGOSTINI in Vorb.). Die Divergenz resultiert daraus, dass die Daten in den traditionellen Sprachatlanten ganz überwiegend und nicht selten ausschließlich in onomasiologischer Perspektive präsentiert werden. Dieses Prinzip, das ausgehend von der Sache bzw. dem Konzept nach der Bezeichnung fragt, liegt auch der Erhebung zugrunde. Zwar werden die abgefragten Sachbegriffe im Atlas bis zu einem gewissen Grade auch kontextualisiert, wie bei der SENNHÜTTE im Blick auf das Almwesen. So nennt die Legende der zu Grunde liegenden AIS-Karte (1192) die folgenden, onomasiologisch (informationstechnisch gesagt: ontologisch) relevanten Schlagwörter, die einen komplexen *frame* konstituieren:

Ausnutzung des Graswuchses auf der Alpweide [...] unter der Hut von Hirten weidendes Vieh [...] Verwertung der gewonnenen Milch an Ort und Stelle [...] Zweck des alpwirtschaftlichen Betriebs [...] Einzelalping [...] genossenschaftliche Betriebe [...] Alpgenosse [...] Häuschen [...] melkt [...] Kühe [...] Milch [...] meist massive steinerne Sennhütte [...] Käser ('Senn') [...] Hilfspersonal [...] Käse [...] Butter [...] Zieger [...] Milch-, Käsekeller [...] Alpstall [...] Ziegmilch [...] Herde [...] Älpfamilie [...] Männer [...] Schutzhütte für das Vieh [...] Schweinestall [...] Milch-, Käsekeller meist tiefhängend oder angebaut oder in einem separaten Häuschen [...] mit laufendem Wasser oder an laufende Gewässer angebaut [...]

Daraus lässt sich etwa der folgende ontologische Rahmen für die Erfassung des Almwortschatzes ableiten; er orientiert sich an fünf sachlichen Teilfeldern und kann beliebig präzisiert werden:

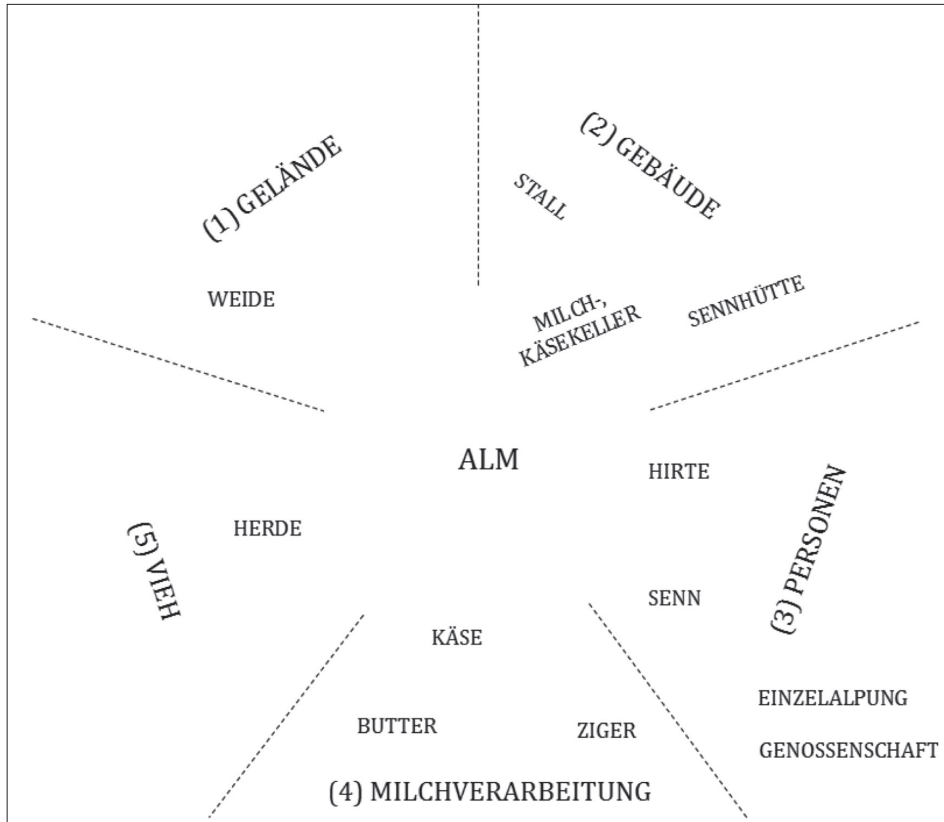


Abb. 2: Elementarer ontologischer *frame* zu ALM.

Aber vor diesem *frame* wird sofort klar, dass die Sachkategorie auf der AIS-Karte selbst, d.h. SENNHÜTTE, unspezifiziert ist; die Legende teilt ja mit, dass es unterschiedliche Arten von Sennhütten gibt ("meist massive steinerne Sennhütten") und dass auf manchen Almen mit mehreren Gebäuden zu rechnen ist ("Milch-, Käsekeller, [...] Alpstall [...], Schutzhütte für das Vieh [...]; Schweinestall [...] Milch-, Käsekeller meist tiefliegend oder angebaut oder in einem separaten Häuschen"). Es ist also durchaus unklar, ob der genannte Ausdruck womöglich (auch) einen speziellen der jeweiligen Gebäudetypen bezeichnen kann.

Ganz anders im Wörterbuch von C. GRASSI, wo die Bedeutungen der komplementären Bezeichnungen detaillierter angegeben werden; hier folgt die Präsen-

tation (wie die Erhebung wohl auch) dem semasiologischen Prinzip, das von der Bezeichnung ausgeht und nach dem Konzept bzw. der Sache fragt:

Typ **C**: “**cašina**: sf. – (alp.) casera: *la cašina endó che stava i vachèr e i lavorava el lat* (edificio della malga) dove stanno i malgari e lavoravano il latte)” (GRASSI 2009, 90)

Typ **G**: “**cašèl**: sm. sg. (pl. **cašèi**) – (cas.) (disus.) l’edificio dove si raccoglieva il latte e si producevano burro, formaggi e ricotta, caseificio” (op.cit., 89)

Typ **B**: “**baita**: sf. – 1. capanna di montagna [...] 2. capanna solitamente di legno, anche semplice baracca, adibita a legnaia [...]” (op.cit., 29)

Typ **E**: “**malga**: sf. – (alp.) pascolo di montagna ed edifici di servizio ad esso annessi” – “Poste ad alta quota, spesso al di sopra della vegetazione d’alto fusto, sono costituite da terreno e pascolo, da una *cašina*, edificio dove si lavora il latte e dove trovano riparo gli addetti, e dallo *stalón*, l’edificio destinato ad accogliere il bestiame in caso di cattivo tempo.” (op.cit., 274)

Bildet man dieses kleine Wortfeld auf den ontologischen Rahmen ab, so ergibt sich das in Abb. 3 dargestellte Schema.¹⁷

Man beachte, nebenbei, dass *malga* im Dialekt von Montagne di Trento (= *malga* 3) eine weitere Bedeutung hat, die durch den AIS ebenfalls nicht erfasst wurde und die in meronymischer, d.h. in einer Teil-Ganzes-Relation zu den auf Karte 3 wiedergegebenen Bedeutungen “Sennhütte” (= *malga* 1) und “Kuhherde” (= *malga* 2) steht.¹⁸ Der Ausdruck kann also – in unterschiedlichen Ortsdialekten – drei ontologische Ausschnitte unseres Rahmens besetzen; berücksichtigt man die weitverbreitete Ableitung *malgár/margár* “Senn” (cf. AIS 1198 CACIAO, SENN), auch “Kuhhirte” (cf. AIS 1187 IL VACCARO, KUHHRIRTE)¹⁹, sind es sogar vier. Andere Ableitungen wie *malghès* “Kuhhirte” und insbesondere das häufige *malgheria/margheria* “Kuhherde”,²⁰ kommen hinzu, so dass die Wortfamilie mit der vorrömischen Wurzel im gesamten romanischen Alpenraum zum Kernwortschatz des Almwesens gehört.

¹⁷ Die durchgezogenen schwarzen Linien markieren den Referenzbereich der rot eingetragenen Lexeme.

¹⁸ Auch dt. *Alm* bezeichnet analog dazu 1. “Weidegelände im Gebirge” und 2. “Gebäude auf dem Weidegelände”.

¹⁹ Für Montagne di Trento schreibt GRASSI (2009, 275): “**malghèr**: sm. – (all.) (alp.), chi conduce la mandria alla *malga* (v.), malghese”. Damit ist offenbar der Senn, nicht der Kuhhirte gemeint, denn letzterer heißt hier *vachèr* (op.cit., 547).

²⁰ Anders ist die Bedeutung in Montagne di Trento: “**malghèra**: sf. – (all.) il luogo dove gli animali al pascolo venivano radunati o si radunano spontaneamente per il riposo quotidiano” (op.cit., 275).

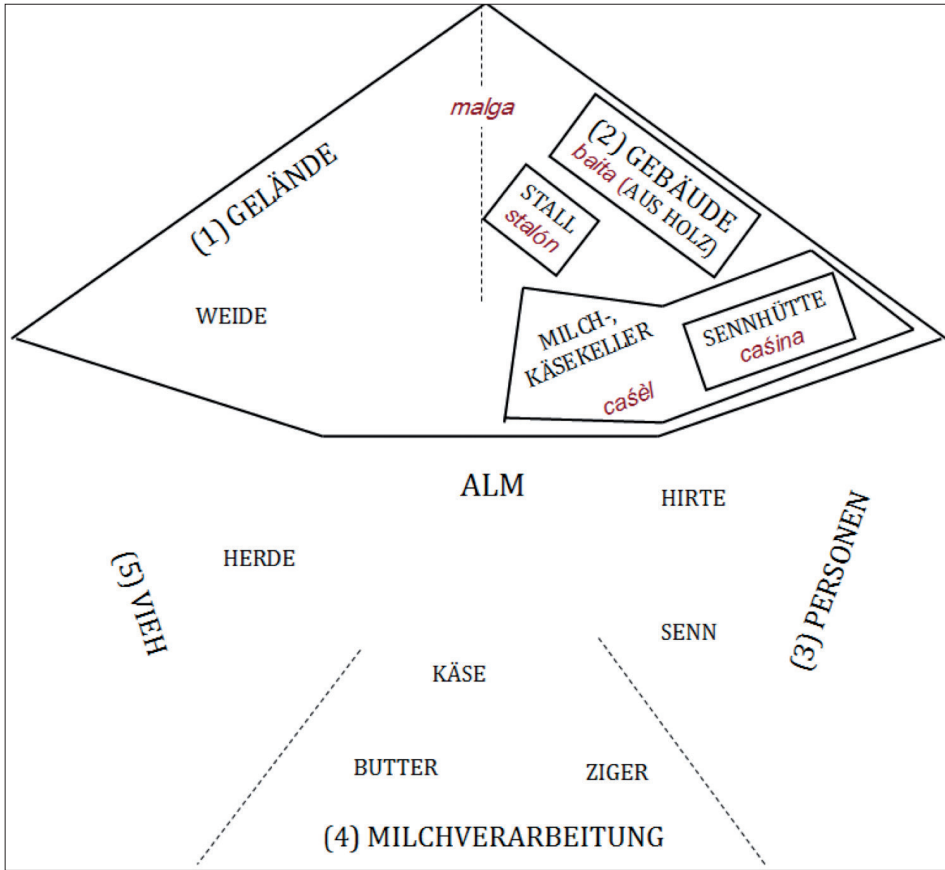


Abb. 3: Bezeichnungen für ALMGEBÄUDE aus dem Dialekt von Montagne di Trento im ontologischen *frame*.

Unser punktuelleres Beispiel zeigt, dass es grundsätzlich wünschenswert und möglich ist, Atlasdaten und Wörterbuchdaten zusammenzubringen; es zeigt aber auch die nicht weniger grundsätzliche Problematik dieses Unterfangens, denn semasiologische Wörterbücher zielen explizit auf maximale Bedeutungs-differenzierung ab und sind tendenziell dem Risiko der Überspezifizierung ausgesetzt, wogegen onomasiologisch erhobene Atlasdaten – ganz im Gegensatz – eher zur Unterspezifizierung neigen, so dass Vergleichbarkeit nicht unbedingt garantiert ist. Die georeferenzierte Kombination dieser beiden zusammengehörigen Quellen von Datenaltbeständen bedarf deshalb einer kritischen methodischen Kontrolle, die durch die Erhebung geeigneter Datenneubestände erfolgen soll.

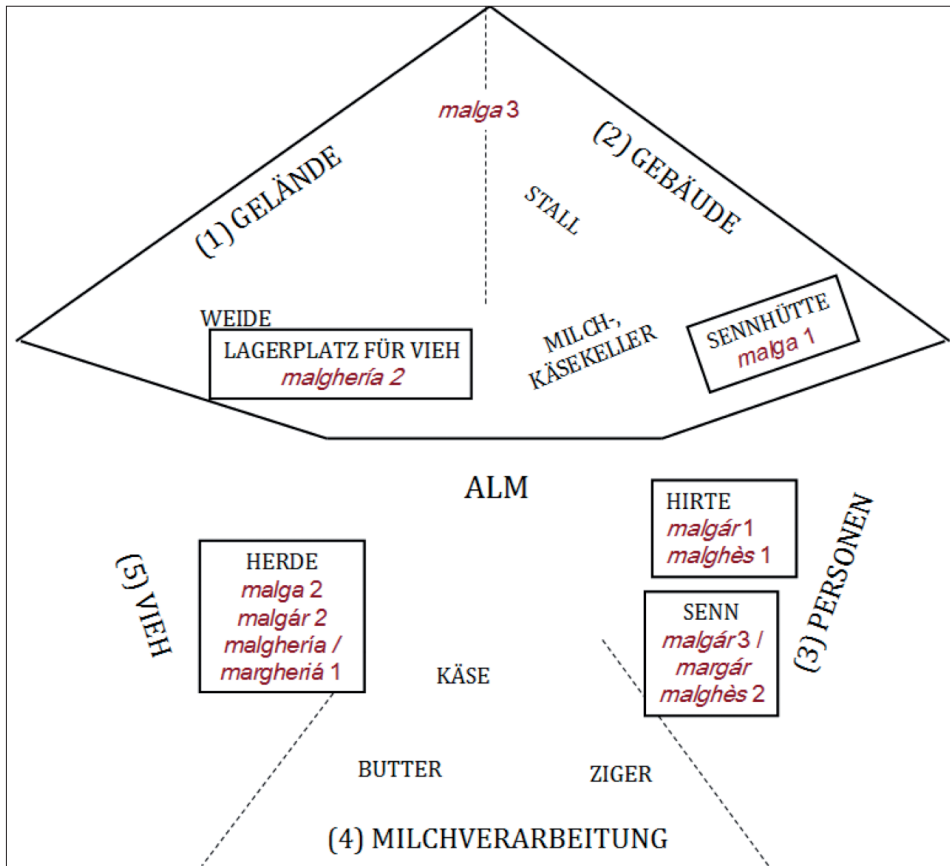


Abb. 4: Der über den gesamten romanischen Alpenraum verteilte Typ *malga* und seine Derivate.

3. Crowdsourcing – macht Zusammengehöriges vergleichbar

Mit der medialen Revolution und speziell mit der Einrichtung des so genannten Web 2.0 haben sich der Sprachwissenschaft mittlerweile ganz neue und vielversprechende Optionen eröffnet, entsprechende Verfahren zu implementieren. Ein erhebliches, erst ansatzweise erprobtes Potential bieten speziell die *social media*, da sie versprechen, große Zahlen von Probanden zu aktivieren.

Mit *VerbaAlpina* soll ein entsprechendes sprachübergreifendes Portal aufgebaut werden, das die Datenaltbestände in linguistisch und ethnographisch kommentierter Weise dokumentiert und darüber hinaus interessierte und dialektkompetente Internetnutzer des Alpenraums, also die relevante “Crowd”, zu systematischer Datenanreicherung einlädt. Dabei können, allgemein gesagt, stets zwei alternative Optionen angeboten werden, je nachdem, ob die Eingabe neuer

Daten einem Raster vorgegebener Kategorien folgt oder nicht; beide Strategien schließen sich jedoch keineswegs aus, denn sie können durchaus komplementär eingesetzt werden. In der hier skizzierten sprachgeographischen Perspektive sind weiterhin zwei Anwendungen zu unterscheiden, die einerseits auf die Hinzufügung weiterer georeferenzierter Ortspunkte und andererseits auf die Ergänzung der in den bereits erfassten Ortspunkten schon verfügbaren Materialien abzielen.

3.1 Erweiterung des Ortsnetzes

Die Möglichkeit, das Netz der sprachgeographisch schon erschlossenen Gebiete zu erweitern, bedarf keiner Rechtfertigung; sie ist im Blick auf manche Gegenden, insbesondere in Österreich, wo noch keine Atlanten vorliegen, von wirklich grundlegender Bedeutung (cf. Karte 1). Die Option mit Vorgaben präsentiert dem Benutzer die Möglichkeit, seine Daten auf einer bereits eingerichteten kartographischen Darstellung des Alpenraums einzutragen; kleinstmögliche Geo-Einheit ist in diesem Fall die politische Gemeinde. Entsprechende Routinen mit kartographischer Nutzerführung wurden für das wissenschaftlich allerdings ganz anders ausgerichtete Projekt *Metropolitalia* (cf. <<http://www.metropolitalia.org/newstatement.html>>, letzte Abfrage 07.10.2014) bereits eingerichtet. Die Funktionalität besteht in der Eingabe von sprachlichen Daten in ein Fenster und ihre Georeferenzierung durch Anklicken eines Punktes auf einer interaktiven und mit einer Datenbank verknüpften Karte. (Allerdings ist die Prozedur im erwähnten Projekt in eine Spielumgebung eingebunden, die für *Verba Alpina* nicht benötigt wird.) Die folgende Abbildung gibt einen Eindruck von der Bildschirmoberfläche:



Abb. 5: Kartographische Benutzerführung bei der Dateneingabe.

Die Ortspunktwahl und Georeferenzierung des Dateneintrags kann ebenso gut über eine Liste erfolgen:

Aggiungi l'espressione che ritieni tipica del luogo scelto.

L'espressione significa:

Provincia: Trento V Comune: Dorsino V

✓ Invia la tua espressione

Abb. 6: Listenbasierte Benutzerführung bei der Dateneingabe.

Es ist allerdings gelegentlich sinnvoll und je nach Fragestellung auch wichtig, auf kleinere Einheiten als die politische Gemeinde zurückzugreifen, etwa bei weit verstreuten Ortsteilen oder bei unterschiedlich ausgeprägter demographischer Dynamik einzelner Orts- oder Stadtteile. In diesem Fall können Daten auch über beliebig fein granulierbare Koordinaten georeferenziert und in die Datenbank eingebunden werden.

3.2 Anreicherung der Primär- und Metadaten

Die Diskussion der Beispiele in Kapitel 2 hat bereits gezeigt, wie der Aufbau des Datenneubestands, jenseits der Ortsnetzproblematik, ausgerichtet sein sollte. Die Option mit kategorialen Vorgaben ist hier von besonderer Wichtigkeit, um die Inkonsistenzen der Altbestände zu kompensieren. Die Benutzer werden konsequent in eine onomasiologische und eine semasiologische Richtung geführt. Aus der Perspektive der Semasiologie werden sie mit der Aufgabe konfrontiert, Ausdrücke und Bedeutungen des schon vorhandenen Datenbestands für ihren jeweiligen Herkunftsort zu bestätigen oder gegebenenfalls zu modifizieren. Für die Erhebungsorte des Altbestands eröffnet sich hiermit eine wichtige diachronische Perspektive, da die Daten des Altbestands ja zum Teil bereits vor mehr als 90 Jahren erhoben wurden. Selbstverständlich werden solche Sprachdaten, z.B. des AIS, niemals ersetzt, sondern im Fall einer Bestätigung durch das Bestätigungsdatum und im Fall der Modifikation durch aktuelle Bedeutungszuweisungen oder Gebrauchsinformationen (z.B. “veraltet” oder “unbekannt”) ergänzt.

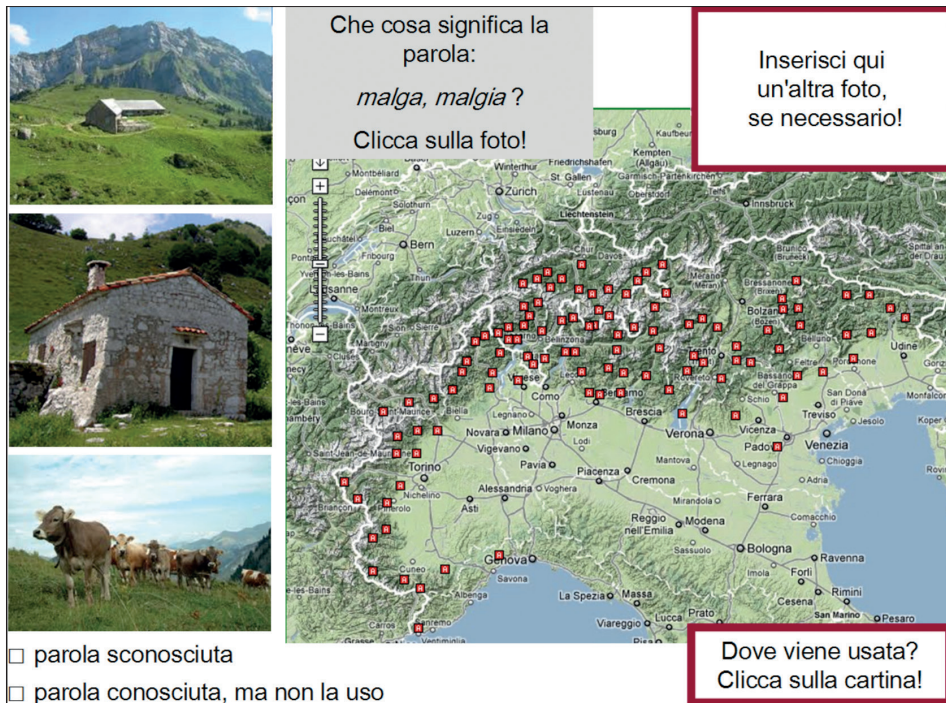


Abb. 7: Oberfläche zur Datenanreicherung.

Über die Evaluation der vorgegebenen Einträge hinaus haben alle Benutzer wiederum die Möglichkeit, fehlende einschlägige Ausdrücke oder Sachbegriffe (bzw. Fotos) einzutragen, so wie es in den Abb. 5 und 6 für Ausdrücke gezeigt wird.

3.3 Validierung von Benutzereingaben

Ein großes Problem beim Einsatz von *Crowdsourcing* besteht darin, dass jeder Nutzer des Internet zunächst anonym ist. Aus der Perspektive des vorgestellten Projektes spielt die konkrete Identität der Internetnutzer keine Rolle; es ist jedoch unverzichtbar, ihre Beiträge zum Datenbestand des Projektes zu bewerten. Für die Validierung anonymer Beiträge haben sich im Internet bereits verschiedene Konzepte etabliert, deren teils komplementärer Einsatz im Rahmen von *Verba Alpina* geplant ist.

Naheliegend und leicht realisierbar ist die Verwendung von *Cookies*, d.h. kleinen Dateien, die auf den *Clients* der Webseitenbesucher abgelegt werden und die bei späteren, neuerlichen Besuchen eine Identifizierung ermöglichen, oder auch das

Angebot an die Nutzer, sich unter freiwilliger Angabe verschiedener persönlicher Daten auf dem *VerbaAlpina*-Portal zu registrieren. Mittelfristig kann – und soll – sich auf diese Weise *VerbaAlpina* nebenbei zu einer Plattform entwickeln, auf der sich Spezialisten und interessierte Laien über spezifische Themen aus dem alpenländischen Sprach- und Kulturraum miteinander austauschen.

Ein bewährter Ansatz für Datenvalidierung ist die Überprüfung und Bestätigung von einmal eingegebenen Daten durch eine unabhängige dritte Instanz. Dies erfolgt ganz einfach durch Analyse des Datenbestandes: Hat ein Internetnutzer beispielsweise die Behauptung aufgestellt, das Wort *malga* bezeichne im Ort Tuenno im Trentino eine Kuhherde, so wird dies in dem Moment als gültig markiert, da ein weiterer Nutzer dieselbe Behauptung aufstellt. Das Verfahren wiederholt im Grunde die unter 3.2 beschriebene Vorgehensweise bei der Evaluierung älterer Atlas- oder Wörterbuchdaten. Die Kumulation von Bestätigungen einer bestimmten Behauptung führt zur Entstehung eines Gültigkeits- oder Relevanzindex, der sich kartographisch durch analoge Veränderung der Punktsymbolgröße und/oder der Farbsättigung des Symbols abbilden lässt. In der zentralen Datenbank würde sich dies in etwa folgendermaßen abbilden. Zunächst würde die Ersteingabe der entsprechenden Behauptung registriert und mit dem Relevanzindex 0 versehen:

user	timestamp	konzept	bedeutung	ort	relevanzindex
nutzer123	2013-10-21 13:17:22	malga	kuhherde	Tuenno	0

Abb. 8: Datenbankeintrag mit erstmaliger Verortung des Begriffs *malga* mit der Bedeutung "Kuhherde" im Ort Tuenno.

Spätere kongruente Eingaben anderer Nutzer resultieren in einer Steigerung des Relevanzindexes, wobei dieser stets die Anzahl der entsprechenden Datenbankeinträge wiedergibt:

user	timestamp	konzept	bedeutung	ort	relevanzindex
nutzer123	2013-10-21 13:17:22	malga	kuhherde	Tuenno	3
nutzer11771	2014-02-12 22:50:43	malga	kuhherde	Tuenno	3
nutzer1822	2014-02-28 19:30:07	malga	kuhherde	Tuenno	3
usw.	usw.	malga	kuhherde	Tuenno	usw.

Abb. 9: Validierung der Verknüpfung *malga*-“Kuhherde”- Tuenno durch analoge Datenbankeinträge anderer Internetnutzer.

In der kartographischen Darstellung könnte sich der Wandel des Relevanzindexes beispielsweise so abbilden lassen:

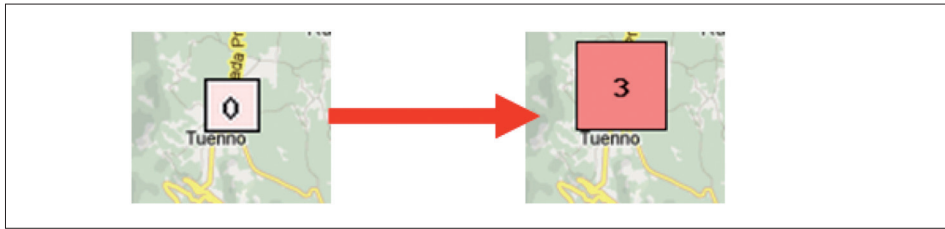
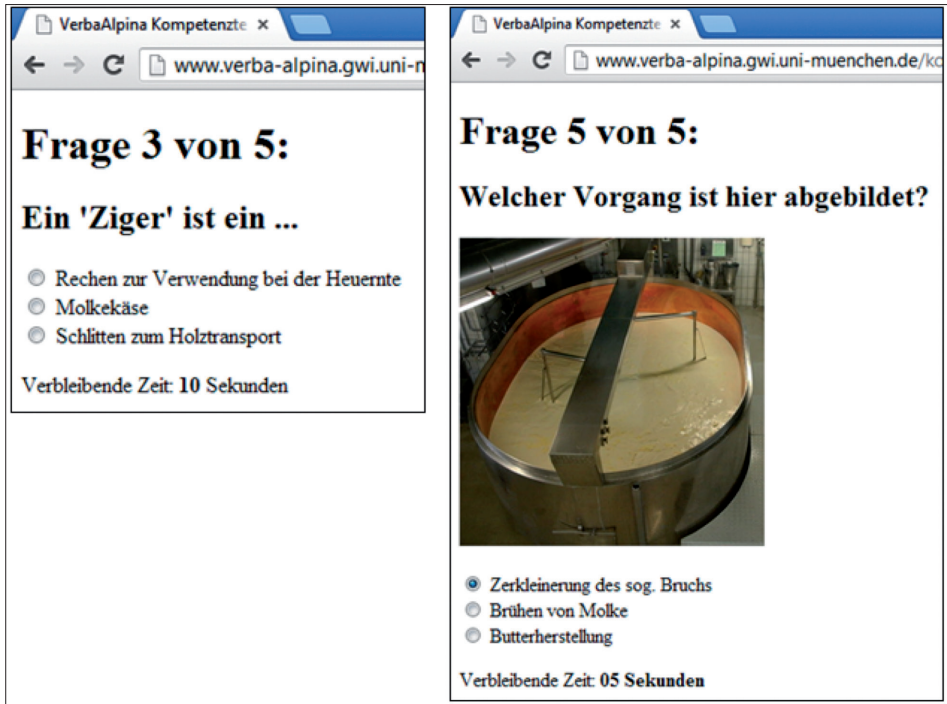


Abb. 10: Kartographische Abbildung des Relevanzindexes.

Im Sinne einer systematischen Datenerhebung wäre es auch sinnvoll, den Internetusern gezielt noch nicht validierte Usereinträge zur Beurteilung vorzulegen.

Für die Einschätzung der von Internetusern beigesteuerten Daten erscheint auch die Formung einer Art von "Kompetenzprofil" der User sinnvoll. In jüngster Zeit haben sich im Internet Strategien etabliert, die auch im Rahmen von *Verba Alpina* Verwendung finden könnten. Interessant ist in diesem Zusammenhang ein Konzept, das unter anderem im Projekt "Galaxy Zoo" bzw. in einer Reihe davon abstammender Tochterprojekte, alle versammelt unter dem Dach von <zooniverse.org>, Verwendung findet. In diesen Projekten werden den anonymen Freiwilligen aus der Internetgemeinde, die z.B. Klassifizierungs- und Transkriptionsaufgaben erfüllen sollen, gelegentlich mehr oder weniger offenkundige Kompetenztests vorgelegt, die Rückschlüsse auf die Qualität der von ihnen beigetragenen Daten erlauben. Diese Kompetenztests basieren jeweils auf unzweideutig beantwortbaren Fragen und erlauben eine Klassifizierung der Nutzer hinsichtlich Ernsthaftigkeit, spezifischem Erfahrungshorizont und Lokalisierung, letzteres unabhängig von Geburts- oder Wohnort.

Auch die Einbindung von Bildern oder Tonaufnahmen in einen solchen Test ist problemlos realisierbar. Um zu verhindern, dass die Probanden die richtige Antwort im Internet recherchieren, empfiehlt sich zumindest bei bestimmten Fragen die Einbindung eines Zeitlimits:

Abb. 11: Kompetenztest.²¹

Es wäre zu erwägen, auf Basis dieser kleinen Kompetenztests ein für die registrierten Benutzer sichtbares “Ranking” zu etablieren, eventuell mit – halb scherzhaft gemeinten – Kategorien von, beispielsweise, “Küstenbewohner” über “Gebirgstourist” bis “Senn”. Die Kompetenztests sollen – ebenso wie das *Framework* des gesamten Internetportals – jeweils mindestens in den Standardsprachen Deutsch, Französisch, Italienisch und Slowenisch sowie unter Umständen in weiteren im Alpenraum verbreiteten Verkehrssprachen verfügbar sein.

3.4 Standardisierung von Benutzereingaben

Neben der Validierung von *Crowdsourcing*-Daten bedarf die transkriptionelle Standardisierung dieser Daten einer gründlichen Planung; denn die Erfassung nicht standardisierter Varietäten sowie die Erfassung der phonetischen Realisie-

²¹ Das Bild zeigt die Zerkleinerung der dickgelegten Milch mit Hilfe der KÄSEHARFE (Bildquelle: <http://commons.wikimedia.org/wiki/File:Production_of_cheese_1.jpg>, letzter Abruf: 04.03.2013).

Wie bezeichnen Sie folgendes/n Objekt/Vorgang?



Bitte wählen Sie zunächst einen Sachbegriff, der zum Dargestellten am besten passt:

Gebäude ▲

Almhütte

Milchkeller

Stall

Milchverarbeitung

Käse ▼

Zum Begriff 'Almhütte' wurden bereits folgende Schreibvarianten eingegeben:

Almhüttñ ▲

Almhittñ

Oimhittñ

...

Wählen Sie eine Variante aus oder geben Sie eine neue Bezeichnung so ein, wie man in der Orthographie der ausgewählten Hochsprache den Wortklang nachbilden würde:

Oimhidñ

Orthographie/ortografia: ☐ dt. ☐ fr. ☐ friul ☐ it. ☐ lad. ☐ rum.gr ☐ slo.

Abb. 12: Eingabe von dem Wortklang nachempfundenen Wörtern ("laien-phonetische" Eingabe).

rung (unabhängig von der Existenz einer Standardorthographie) stellen zweifellos eine besondere Herausforderung dar. Eine wissenschaftlichen Ansprüchen genügende phonetische Transkription kann und darf einem Laien, d.h. einem linguistisch nicht gebildeten Sprecher nicht abverlangt werden; da *Verba Alpina* lexikologisch ausgerichtet ist, wird impressionistische phonetische Präzision auch gar nicht erwartet. Es sollte daher genügen, wenn man die Informanten auffordert, den Wortklang mit den Mitteln der hochsprachlichen Orthographie nachzubilden, wobei der Nutzer die jeweilige Referenzorthographie angeben müsste. In diesem Fall könnte man die Auswahl tatsächlich auf die Standardsprachen beschränken, da man davon ausgehen kann, dass alle Informanten in mindestens

einer dieser Sprachen orthographische Kompetenz besitzen. Die entsprechende Eingabemaske könnte in etwa so aussehen wie in Abb. 12 dargestellt.

Eine zusätzliche Absicherung erfährt diese Eingabemethode durch die Verknüpfung der eingegebenen Bezeichnung mit Elementen der vorgegebenen – gleichzeitig aber auch erweiterbaren – lexikalischen Ontologie. Im Einzelnen werden sich in der Durchführung manche, vorderhand nicht vorhersehbare Detailprobleme ergeben.

Ergänzend ist daran gedacht, den Informanten die Möglichkeit zu geben, Beiträge via Mikrofon als Audioaufnahmen beizusteuern. Dieses Verfahren böte außerdem die Möglichkeit, die eingesandten Audioaufnahmen zu einem späteren Zeitpunkt von Spezialisten phonetisch transkribieren zu lassen bzw. die einmal eingegebene Transkription anhand der Audioaufnahme zu überprüfen. Selbstverständlich ist eine nachgelagerte Zweitkategorisierung durch Dritte oder eine Korrektur der Primärkategorisierung möglich. Technische Möglichkeiten zur Tonaufnahme über das Internet bestehen bereits heute, wobei derzeit noch der Rückgriff auf die *Flash*-Technologie erforderlich ist. In nächster Zukunft wird es jedoch möglich sein, entsprechende Funktionen mit HTML(5)-eigenen Mitteln zu realisieren.

4. Bibliographie

ADIS = KREFELD, Thomas/LÜCKE, Stephan: *Atlante linguistico digitale dell'Italia e della Svizzera meridionale*: <<http://www.adis.gwi.uni-muenchen.de>>.

AIS = JABERG, Karl/JUD Jakob (eds.): *Sprach- und Sachatlas Italiens und der Südschweiz*, Zofingen 1928–1940, 8 voll.; [Neudruck: Nendeln 1971].

ALAVAL = KRISTOL, Andres (ed.): *Atlas linguistique audiovisuel du Valais romand*: <<http://www2.unine.ch/dialectologie/page-8174.html>>.

ALD-I: GOEBL, Hans/BAUER, Roland/HAIMERL, Edgar (eds.): *Atlant linguistich dl ladin dolomitich y di dialec vejins, 1ª pert* / *Atlante linguistico del ladino dolomitico e dei dialetti limitrofi, 1ª parte* / *Sprachatlas des Dolomitenladinischen und angrenzender Dialekte, 1. Teil*, Wiesbaden 1998, 7 voll.; [sprechender Sprachatlas: <<http://ald.sbg.ac.at/ald/ald-i/index.php>>].

ALD-II = GOEBL, Hans et al. (eds.): *Atlant linguistich dl ladin dolomitich y di dialec vejins, 2ª pert* / *Atlante linguistico del ladino dolomitico e dei dialetti limitrofi, 2ª parte* / *Sprachatlas des Dolomitenladinischen und angrenzender Dialekte, 2. Teil, II*, Strasbourg 2012, 7 voll.

ALEPO = CANOBBIO, Sabina/TELMON, Tullio: *Atlante linguistico ed etnografico del Piemonte occidentale*, Torino 2003–.

- ALF: GILLIÉRON, Jules/EDMONT, Edmond (eds.): *Atlas linguistique de la France*, Paris 1902–1910, 9 voll.
- ALI = MASSOBRIO, Lorenzo et al. (eds.): *Atlante linguistico italiano*, Roma 1995–.
- ALJAN = MARTIN, Jean-Baptiste/TUAILLON, Gaston: *Atlas linguistique et ethnographique du Jura et des Alpes du nord*, Paris 1971, 1978, 1981, 3 voll.
- ALL = GARDETTE, Pierre: *Atlas linguistique et ethnographique du Lyonnais*, Paris 1950, 1952, 1956, 3 voll.
- ALP = BOUVIER, Jean-Claude: *Atlas linguistique et ethnographique de la Provence*, Paris 1975, 1979, 1986, 3 voll.
- ASLEF = PELLEGRINI, Giovan Battista: *Atlante storico-linguistico-etnografico friulano*, Padova 1974–1986, 6 voll.
- ASV = GEIGER Paul/WEISS, Richard: *Atlas der schweizerischen Volkskunde/Atlas de Folklore suisse/Atlante di Folclore svizzero*, Basel 1950–1995, 2 Teile, Karten und Kommentar.
- DE AGOSTINI, Ferruccio (in Vorb.): *Vocabulari dal dialèt de Gurdúna*.
- DECURTINS Alexi: *Lexicon romontsch cumparativ. Sursilvan – tudestg*, Chur 2012.
- DELI = CORTELAZZO, Manlio/ZOLLI, Paolo: *Dizionario etimologico della lingua italiana*, Bologna 1979.
- DURO, Aldo: *Il Vocabolario Treccani*, Roma 1997², 5 voll., <<http://www.treccani.it/vocabolario/>> .
- ELSPASS, Stephan/KÖNIG, Werner (eds.): *Sprachgeographie digital. Die neue Generation der Sprachatlanten (mit 80 Karten)*, Hildesheim/Zürich/New York 2008.
- GEISLER, Hans: *Alpinromanischer Wortschatz – Anfizonza Lombardo-Ladina*, <<http://alplomb.phil.hhu.de/>>.
- GRASSI, Corrado: *Dizionario del dialetto di Montagne di Trento*, San Michele all'Adige 2009.
- KBS = KÖNIG, Werner: *Kleiner Bayerischer Sprachatlas*, München 2006; *Sprechender Sprachatlas von Bayern*: <<http://sprachatlas.bayerische-landesbibliothek-online.de>>.
- KREFELD, Thomas: *Rezension zu* ELSPASS/KÖNIG 2008, op.cit., in: “Zeitschrift für Dialektologie und Linguistik”, 76, 2009, 351–353.
- KREFELD, Thomas/LÜCKE, Stephan: *Geolinguistik 2.0. – Ein aktueller Bericht aus der Werkstatt der webbasierten Sprachgeographie*, in: TOSQUES, Fabio (ed.), *20 Jahre digitale Sprachgeographie*, Berlin 2014, 133–154.
- REW = Meyer-Lübke, Wilhelm: *Romanisches etymologisches Wörterbuch*, Heidelberg 2009⁷.
- SAO = Adalbert-Stifter-Institut des Landes Oberösterreich (ed.): *Sprachatlas von Oberösterreich*, Linz 1998–.
- SAT = GABRIEL, Eugen/FUNK, Edith/SEIDELMANN, Erich: *Sprachatlas von Tirol*; [nicht publiziertes Material, zugänglich über das Bayerische Wörterbuch an der Bayerischen Akademie der Wissenschaften].
- SBS = KÖNIG, Werner/WEILMANN, Hans: *Sprachatlas von Bayerisch-Schwaben*, Heidelberg 1996–.
- SCHORTA, Andrea: *Rätisches Namenbuch*, Bd. 2: *Etymologien*, Bern 1964.
- SDS = HOTZENKÖCHERLE, Rudolf et al.: *Sprachatlas der Deutschen Schweiz*, Bern 1962–2003, 9 voll.
- SLA = SKOFIC, Jožica (ed.): *Slovenski lingvistični atlas 1. Človek – telo, bolezni, družina*, Ljubljana 2011–.
- SOB = EICHINGER, Ludwig M. (ed.): *Bayerischer Sprachatlas / Regionalteil VI: Sprachatlas von Oberbayern*, Heidelberg 2005–.

TPPSR = GAUCHAT, Louis/JEANJAQUET, Jules/TAPPOLET, Ernest: *Tableaux phonétiques des patois suisses romands*, Neuchâtel 1925.

TSA = KÜHEBACHER, Egon: *Tirolischer Sprachatlas*, Innsbruck 1965, 1969, 1971, 3 voll.

VALTS = GABRIEL, Eugen: *Vorarlberger Sprachatlas mit Einschluss des Fürstentums Liechtenstein, Westtirols und des Allgäus*, Bregenz 1985–2004, 5 voll.

VIVALDI = KATTENBUSCH, Dieter/TOSQUES, Fabio (eds.): *Vivaio Acustico delle Lingue e dei Dialetti d'Italia*, <<http://www2.hu-berlin.de/vivaldi/index.php>>.

Résumé

L raion alpin é na region dret interessanta, nia enultima dal pont de veduda linguistich. Viadedò n sotfonz de na omogeneité locala y etnografica él n che-der rich de fassetes dla varieté linguistica, caraterisé dal contat de trei families linguistiches (romanich, germanich, slavich). Ence sce les enrescides à na tradi-zion longia ne vala empone da se informé sun dut l raion y sun sia stratificazion storica y linguistica, problem dessegur ence gaujé da sie spediciament politich y scientifich. L projet digital *VerbaAlpina* <<http://www.verba-alpina.gwi.uni-muenchen.de/>> à l fin de mete adum te n gran mosaich l vocabolar spezifich dles Alpes bele dant man ti ciamps dla “etnografia”, dla “fauna”, dla “flora”, dla “contreda” y dl “temp” y de l prejenté tla plu gran trasparenza te internet sciche atlant online. L material bele publiché te atlanc y te vocabolars dess chilò vegnì integré, controlé y prezisé con les metodes dl *crowdsourcing*. Na banca de dac relazionale é l strument zentral: si dac pò vegnì consultés tres internet, ensi che la struttura geolinguistica dl raion se mostra tla cartografia de paroles y de cosses. De gran importanza é chilò i areai che va soura i confins linguistics fora.