

9 Aufbereitung der Daten

Hier nehme ich ganz Obwalden zusammen. Die Lungerer haben sicher andere Ausdrücke als die Sarner, aber ich würde es zusammennehmen.

Draw-a-map-task Nahraum, Proband PB4 mit Bezug auf *Obwalden*

Die Daten, die vorliegend interessieren (vgl. hierzu Tab. 3), wurden für die Auswertung und Interpretation je unterschiedlich aufbereitet. Die Aufbereitungsschritte werden nachfolgend diskutiert und damit transparent gemacht. Ein Grossteil der aufbereiteten Daten befinden sich in einer sql-basierten Datenbank, die für das *Ländere*ⁿ-Projekt eingerichtet wurde und die mit einem Passwortzugang einsehbar sind (vgl. <http://www3.unifr.ch/laenderen/> (letzter Zugriff 10.06.2019)).

9.1 Sozialdaten

Die Sozialdaten, die in Orientierung an das Personalblatt des SDS erhoben wurden (vgl. Kap. 8.4.1), wurden zu grossen Teilen direkt (Name, Geburtsjahr und Geburtsort der Probanden) in ein Excel-File übertragen. Manche wurden zusätzlich in Kategorien überführt: Die Informationen zur Mobilität der Probanden bspw. wurden je nach Mobilitätsradius der Probandinnen und Probanden mit Ziffern versehen (Hauptbezugspunkt Wohnort = Kategorie 1, Hauptbezugspunkt Innerschweiz = Kategorie 2, Hauptbezugspunkt Schweiz = Kategorie 3).

9.2 Objektsprachliche Daten

Die objektsprachlichen Daten wurden einerseits mittels Fragebogeninterview, andererseits im freien Gespräch erhoben; beide Gesprächssituationen wurden aufgezeichnet (vgl. dazu Kap. 8.4.2 und Kap. 8.4.3). Aus dem Fragebogeninterview wurden einzig die objektsprachlichen Realisierungen der Probanden transkribiert, aus dem freien Gespräch zusätzlich zu den einschlägigen Realisierungen die gesamte Gesprächssituation. Die Transkription erfolgte in Orientierung an Dieth (1986). Um die transkribierten objektsprachlichen Realisierungen für qualitative Zwecke analysierbar zu machen, wurden sie in einem zweiten Schritt kategorisiert. Die kategorisierten Daten liegen allesamt in Excel-Files vor. Überdies sind sie über die sql-Datenbank greifbar, die unterschiedliche Abfragemög-

lichkeiten bietet (Abfrage nach Phänomen, Person, Ort sowie Sozialdaten). Um sie auch für quantitative Auswertungen zu nutzen, wurden die kategorisierten Daten zusätzlich recodiert.

9.2.1 Kategorisierung

Die Kategorisierung der objektsprachlichen Daten erfolgte in Orientierung an die beschriebenen Dialektbefunde im SDS. Tab. 4 bietet einen Überblick über alle kategorisierten Variablen; sie verzeichnet zunächst 32 lautliche Variablen (30 Variablen zur Lautqualität, zwei Variablen zur Lautquantität), danach zwei morphologische und schliesslich vier lexikalische Variablen.

Für jede dialektale Variable, die vorliegend von Interesse ist, wurden Varianten bestimmt, die laut SDS gängigen Unterwaldner Varianten entsprechen; für mhd. *û* beispielsweise die Varianten *ui*, *uu*, *üü*, *oi/öi*, für mhd. *uo* die Varianten *io* und *ue* usw. Die objektsprachlichen Realisierungen der Probandinnen und Probanden wurden anschliessend diesen Antwortkategorien zugeteilt. Während dieses Prozesses kam es zu Anpassungen und zu Erweiterungen des Kategoriensystems: Das Kategoriensystem von mhd. *û* beispielsweise wurde um die Kürzen *u* und *ü* erweitert, das Kategoriensystem von mhd. *uo* um die Variante *üe*. Überdies wurden für jede Variable die Antwortkategorien *freie Antwort*, *Variable nicht genannt* und *andere Variable genannt* geschaffen, um realisierte Varianten, die von den vorgesehenen Kategorien abweichen (und für die nicht extra eine neue Kategorie geschaffen werden sollte), ebenfalls klassifizieren zu können: *Freie Antwort* bedeutet dabei, dass die Variante, die realisiert wurde, zwar im Lexemgerüst auftritt, das vorgesehen war, die Variante selbst aber nicht vorgesehen war; *Variable nicht genannt* meint, dass die Variable gar nicht realisiert wurde; *andere Variable genannt* meint, dass hier das Lexemgerüst nicht dem entspricht, das erwartet wurde, sondern dass die Probandin/der Proband ein anderes Lexem gebrauch hat. Dieses Vorgehen – die Konzeption des Kategoriengerüsts in Orientierung an den SDS und dessen induktive Erweiterung – erlaubte, möglichst viele der realisierten Varianten möglichst angemessen zu kategorisieren.

Zur Einteilung der von den Probandinnen und Probanden realisierten Varianten in die jeweiligen Antwortkategorien seien noch einige kritische Reflexionen angefügt. Was die Einteilung der Varianten angeht, die mittels Fragebuch erhoben wurden, ist zu problematisieren, dass diese Einteilung von zwei verschiedenen Personen vorgenommen wurde: Trotz genauer Absprachen wird kaum zu vermeiden sein, dass die gehörten Varianten z. T. unterschiedlich klassifiziert wurden. Bezüglich der Einteilung der Varianten, die im freien Gespräch

Tab. 4: Übersicht über die kategorisierten Variablen mit ihren jeweiligen Antwortkategorien

Variable	Abstrahierte Variante	Beispiel
mhd. <i>î</i> vor Konsonant	<i>ii, ei, i</i>	<i>lis, Eis, Is</i>
mhd. <i>î</i> im Auslaut	<i>ii, ei, i</i>	<i>gsii, gsei, gsi</i>
mhd. <i>î</i> im Hiatus		<i>schneie, schniie, schneiee</i>
mhd. <i>û</i> vor Konsonant	<i>uu, ui, oi/öi, üü, u, ü</i>	<i>Muus, Muïs, Moïs/Möis, Müüs, Mus, Mûs</i>
mhd. <i>û</i> im Auslaut	<i>uu, ui, oi/öi, üü, u, ü</i>	<i>duu, dui, doi/döi, düü, du, dü</i>
mhd. <i>û</i> im Hiatus		<i>baue/boue, buu(w)e, bul(w)e, bol/öi(w)e, büü(w)e</i>
mhd. <i>iu</i>	<i>üü, ii, ei, ui</i>	<i>Müüs, Mûis, Meis, Muïs, Mûs</i>
mhd. <i>iu</i> im Hiatus		<i>röje, riuwe, reiwe, riye, ruje, rei(j)e, rüü(j)e, roje</i>
mhd. <i>ou</i>	<i>ou, oi, ai, au/äu, öi</i>	<i>Boum, Boim, Baim, Baum/Bäum, Böim</i>
mhd. <i>ou + w</i>	<i>ou, oi, ai, au/äu, öi</i>	<i>Frou, Froi(w), Frai(w), Frau/Fräu, Fröi(w)</i>
mhd. <i>ou</i> Sonderfall <i>glauben</i>	<i>ou, oi, ai, au/äu, öi</i>	<i>gloube, gloibe, glaibe, glaube/gläube, glöibe</i>
mhd. <i>öu + w</i>		<i>höie, heu(u)e, häiwe, häie, hoie</i>
mhd. <i>öu</i> vor Nasal	<i>öi, ai, oi</i>	<i>Böimli, Baimli, Boimli</i>
mhd. <i>uo</i>	<i>ue, io, üe</i>	<i>guet, giot, gtiot</i>
mhd. <i>üe</i>	<i>üe, ie</i>	<i>müed / mied</i>
aobd. <i>iu</i>	<i>äi, äü, öi, ie, üü, oi</i>	<i>täif, täüf, töif, tief, tüüf, toif</i>
aobd. <i>iu</i> (fliegen, lügen)		<i>fliige, fleige, fluige, fläige, flüüge</i>
aobd. <i>iu + w</i>		<i>Chnâi(w), Chneu(w), Chnii(w), Chnai, Chnûü, Chnöi, Chnoi</i>
mhd. <i>e</i> (Primärumlaut) vor Nasal + Konsonant	<i>e, ä</i>	<i>eng/läng; Hent/Händ</i>
Sekundärumlaut vor Nasal + Konsonant	<i>e, ä</i>	<i>Hendsche / Händsche</i>
mhd. <i>rn</i>		<i>gärn, gäre, gären</i>
mhd. <i>l</i>	<i>l, u</i>	<i>-el, -u</i>
(Nicht)Umlaut von mhd. <i>u</i>	<i>ü/i, u</i>	<i>über/iber, uber</i>
Entrundung mhd. <i>iu</i>	<i>entrundet, nicht entrundet</i>	<i>üü/üi/ü, ii/ei</i>
Entrundung mhd. <i>ö</i>	<i>entrundet, nicht entrundet</i>	<i>ö, e</i>

Entrundung mhd. <i>oe</i>	<i>entrundet, nicht entrundet</i>	öö, ee
Entrundung mhd. <i>öu</i>	<i>entrundet, nicht entrundet</i>	variiert je nach Kontext (<i>öu</i> ; <i>öu</i> + <i>w</i> ; aobd. <i>iu</i>)
Entrundung mhd. <i>ü</i>	<i>entrundet, nicht entrundet</i>	<i>ü, i</i>
Entrundung mhd. <i>üe</i>	<i>entrundet, nicht entrundet</i>	<i>üe, ie</i>
Rundung mhd. <i>e</i>	<i>ungerundet, gerundet</i>	<i>e, ö</i>
Dehnung von alter Kürze im einsilbigen Wort	<i>kurz/lang</i>	<i>Glas/Glaas</i>
Dehnung von alter Kürze in offener Silbe	<i>kurz/lang</i>	<i>Grabe/Graabe</i>
Diminutiv	<i>-li, -eli, -ili</i>	<i>-li, -eli, -ili</i>
Normalverb 3. Pl. Ind. Präs.	<i>-id, -ed, -e</i>	<i>machid, mached, mache; schlafid, schlafed, schlafe</i> usw.
'sehr'		<i>fest, huere, meined, sehr, so, usinnig</i>
'immer'		<i>eisder, eisdig, eischer, tuschuur, allpott, allwill, geng, immer</i>
'noch'		<i>no, nu</i>
'hinunter'		<i>abe, appe</i>

erhoben wurden, gibt es v. a. einen Problembereich, und zwar, dass die Varianten, die zu klassifizieren waren, nicht bereits feststanden (wie in der Fragebuch-situation), sondern nachträglich identifiziert werden mussten. Dabei mussten unterschiedliche Entscheidungen über die Aufnahme resp. den Ausschluss von Varianten getroffen werden, die z. T. nicht unproblematisch waren. Vier Umstände, die diese Entscheidungen erschwerten, sollen hier herausgegriffen und thematisiert werden: Sie sollen auf einer Metaebene auch illustrieren, dass beim Aufbereiten der objektsprachlichen Daten mit der nötigen Sorgfalt vorgegangen wurde. Eine dieser Entscheidungen betraf die Identifizierung und Aufnahme neuer Laute in das Datenkorpus. Während im Fragebogeninterview nur alte Laute abgefragt wurden, kamen im spontanen Gespräch auch Laute vor, die als neue (entlehnte) Laute charakterisiert und kategorisiert werden konnten: mhd. *î* bspw., das einen alten langen Laut darstellt, wurde im Fragebogeninterview u. a. anhand des Lexems *zîte* 'Zeit' erhoben. Im spontanen Gespräch kamen nun Lexeme wie *poliitisch* 'politisch' und *relatiiv* 'relativ' vor, die ebenfalls lange *i*-Laute beinhalten, die aber nicht als alte, sondern als neue Längen charakterisiert werden müssen. Neue Längen, die sich von ihrer lautlichen Form her analog verhalten wie alte Längen, wurden wie die alten Längen den historischen Klassen zugeteilt. Es wurde also nicht zwischen alten und neuen Längen unterschieden. Eine weitere Entscheidungskette drehte sich um das Phänomen der Diphthongierung und betrifft wiederum die Problematik von alten und neuen Lauten. Im Untersuchungsgebiet ist u. a. die Diphthongierung von mhd. *û* zu *ui* und jene von mhd. *î* zu *ei* ausgewiesen. Auch hier musste bezüglich der neuen Laute entschieden werden, welche Varianten bzw. welche Lexemumgebungen überhaupt Grundvoraussetzungen für allfällige Diphthongierungsprozesse bieten. Bestimmt wurde, dass grundsätzlich jeder Laut, der lang ist, eine Grundlage zur Diphthongierung bietet, mit der Einschränkung, dass weniger morphologische Substanz mit weniger Chancen auf Diphthongierung gleichzusetzen ist. Varianten wie *Feirma* 'Firma' (diphthongierte Variante von *Fiirma*) oder auch *Hambuirger* 'Hamburger' (diphthongierte Variante von *Hambuurger*) sind Fälle von neuen, langen und zusätzlich neuen, diphthongierten Lauten, die in den Spontandaten auftreten und in die Kategorien mhd. *î* und mhd. *û* aufgenommen wurden. Die letzte Entscheidungskette, die hier thematisiert werden soll, betrifft das Phänomen der *l*-Vokalisierung. Auch bezüglich dieser Variable, die ein neueres Phänomen im Untersuchungsgebiet darstellt, mussten Kontexte bestimmt werden, in denen die Grundvoraussetzungen für eine Vokalisierung erfüllt zu sein schienen. In Anlehnung an Christen (2001) werden dabei folgende Vokalisierungskontexte berücksichtigt: VLC (*Geld*), VL## (*viel*), kurz: L (*Himmel*), VLLV (*Keller*). Realisierungen von anlautendem und intervokalischem *l* werden vernachlässigt.

9.2.2 Recodierung

Nachdem die objektsprachlichen Daten nach den beschriebenen Kriterien kategorisiert waren, wurden sie so recodiert, dass sie sich zusätzlich zur Beantwortung der vorliegend interessierenden quantitativen Forschungsfragen eignen (vgl. dazu im Detail Kap. 13.1). Dafür wurden die Antwortkategorien – die sich durch je unterschiedliche und unterschiedlich viele Varianten auszeichnen – in die Metakategorien *basisdialektal*, *grossräumig*, *kleinräumig* überführt, um quantitative Aussagen darüber machen zu können, zu welchen Teilen die Probandinnen und Probanden die fraglichen Variablen basisdialektal, grossräumig oder kleinräumig realisieren. Die Entscheidung darüber, in welche Kategorie eine Variante fällt, wurde in Orientierung an der im SDS ausgewiesenen lokalen Norm gefällt. *Basisdialektal* meint dabei, dass die Antwortkategorie in Bezug auf den entsprechenden Ortspunkt im SDS als basisdialektale Variante ausgewiesen ist; *kleinräumig* bedeutet, dass die Antwortkategorie nicht in Bezug auf den entsprechenden Ortspunkt, dafür aber in Bezug auf andere Ortspunkte, die im Untersuchungsgebiet liegen, als basisdialektale Variante ausgewiesen ist, gleichzeitig aber nicht auch für den Grossraum der Schweiz als basisdialektale Varianten belegt ist; *grossräumig* meint schliesslich, dass die Antwortkategorie nicht in Bezug auf den entsprechenden Ortspunkt, dafür aber in Bezug auf Ortspunkte, die – innerhalb des Deutschschweizer Dialektraumes – für einen Großteil der Ortspunkte als basisdialektale Variante ausgewiesen ist. Dieser Prozess der Recodierung soll zusätzlich an einem Beispiel verdeutlicht werden: Auf der Grundlage der bereits in Kap. 9.2.1 erwähnten dialektalen Variablen mhd. *î*, mhd. *û* und mhd. *uo* wird in Tab. 5 dargestellt, wie die Recodierung für die Ortspunkte Lungern (UW 10) und Engelberg (UW12) konkret umgesetzt wurde (vgl. für eine Übersicht über die Recodierung aller Variablen die Tabelle im digitalen Anhang unter 7).

Die Variable mhd. *î* etwa ist im SDS für den Ortspunkt Lungern in der Variante *ii* (bzw. an unbetonter Stelle als *i*) und für den Ortspunkt Engelberg in der diphthongierten Variante *ei* als basisdialektale Variante ausgewiesen. Infolgedessen wurde für den Ortspunkt Lungern die Variante *ei* als kleinräumige Variante codiert, da diese Variante nicht für Lungern selbst, wohl aber für andere Orte im Untersuchungsgebiet als basisdialektale Variante gilt (nicht aber für den Grossraum der Schweiz). Für den Ortspunkt Engelberg hingegen wurden die Varianten *ii* und *i* als grossräumige Varianten qualifiziert, da die nicht diphthongierte Entsprechung von mhd. *î* die Normvariante vieler Schweizer Dialekte darstellt. Hierbei ist transparent zu machen, dass die Kurzformen der Laute – also bspw. *i* – immer in die gleiche Kategorie eingeteilt wurden wie die Langformen: In Bezug auf den Ortspunkt Lungern etwa gelten sowohl *i* als auch *ii* als basisdialektale Variablen, in Bezug auf den Ortspunkt Engelberg gelten sowohl

Tab. 5: Einblick in den Prozess des Recodierens anhand ausgewählter Variablen und Ortschaften

Obertyp	Untertyp	Antwort- kategorie Beispiel	Antwort- kategorie abstrakt	Recodierung Lungern UW 10	Recodierung Engelberg UW 12
<i>î</i>	mhd. <i>î</i> vor Kons.	<i>lis</i>	<i>ii</i>	basisdialektal	grossräumig
<i>î</i>	mhd. <i>î</i> vor Kons.	<i>Eis</i>	<i>ei</i>	kleinräumig	basisdialektal
<i>î</i>	mhd. <i>î</i> vor Kons.	<i>Is</i>	<i>i</i>	basisdialektal	grossräumig
<i>û</i>	mhd. <i>û</i> vor Kons.	<i>Muus</i>	<i>uu</i>	grossräumig	grossräumig
<i>û</i>	mhd. <i>û</i> vor Kons.	<i>Muis</i>	<i>ui</i>	basisdialektal	kleinräumig
<i>û</i>	mhd. <i>û</i> vor Kons.	<i>Mois/Möis</i>	<i>oi/öi</i>	kleinräumig	basisdialektal
<i>û</i>	mhd. <i>û</i> vor Kons.	<i>Müüs</i>	<i>üü</i>	kleinräumig	kleinräumig
<i>û</i>	mhd. <i>û</i> vor Kons.	<i>Mus</i>	<i>u</i>	grossräumig	grossräumig
<i>û</i>	mhd. <i>û</i> vor Kons.	<i>Müs</i>	<i>ü</i>	kleinräumig	kleinräumig
<i>uo</i>	mhd. <i>uo</i>	<i>guet</i>	<i>ue</i>	grossräumig	grossräumig
<i>uo</i>	mhd. <i>uo</i>	<i>giot</i>	<i>io</i>	basisdialektal	kleinräumig
<i>uo</i>	mhd. <i>uo</i>	<i>güet</i>	<i>üe</i>	kleinräumig	basisdialektal

ii als auch *i* als grossräumige Varianten. Gerade letztere Codierung kann dahingehend kritisiert werden, als ein kurzes *i* nicht die Idealvoraussetzungen erfüllt, um diphthongiert zu werden; so werden unbetonte undiphthongierte *i*-Laute im SDS für den Ortspunkt Engelberg denn auch z. T. als basisdialektale Laute ausgewiesen. Die Einteilung wurde vorliegend trotzdem so vorgenommen mit der Begründung, dass es als Bearbeiterin manchmal nur schwer möglich war, betonte von unbetonten Lauten zu unterscheiden; dadurch, dass mehrere Personen an der Transkription der Daten beteiligt waren, wurde dieser Umstand noch erschwert. Die nächste Beispielvariable mhd. *û* ist in Bezug auf Lungern mit der basisdialektalen Form *ui* ausgewiesen, *uu* entspricht demgegenüber der Normalform des Schweizer Grossraums. *Oi/öi* ist als Engelberger Form belegt, *üü* als Urner Variante, weshalb diese beiden Formen als kleinräumige Formen eingetragen wurden. In Bezug auf den Ortspunkt Engelberg sieht es etwas anders aus: Hier ist *oi/öi* als basisdialektale Variante belegt, *ui* und *üü* wurden als kleinräumige Varianten eingestuft, weil sie im übrigen Unterwalden und im angrenzenden Uri verbreitet sind. *Uu* kann auch in Bezug auf Engelberg als grossräumige Form eingestuft werden. Was schliesslich die Variable mhd. *uo* anbetrifft, so gilt deren Reflex *ue* sowohl für Lungern als auch für Engelberg als grossräumige Variante; *io* ist die basisdialektale Variante für Lungern (und eine kleinräumige für Engelberg), *üe* ist die basisdialektale Variante für Engelberg (und eine kleinräumige für Lungern).

Es liegt auf der Hand, dass die Recodierung der Basiskategorien in die Metakategorien *basisdialektal*, *kleinräumig*, *grossräumig* im Einzelfall nicht einfach und eindeutig war: Am kritischsten dabei war, dass die Varianten, die der SDS als basisdialektale Varianten von Ortspunkten ausweist, einer gewissen Variation unterliegen. So werden etwa in Bezug auf mhd. *î* – zumindest, wenn man die Kommentare zu den Karten genau liest – für gewisse Ortspunkte z. T. auch nicht diphthongierte Formen als basisdialektale Formen ausgegeben. Solche Angaben – die bei der Recodierung der Daten die grösste Herausforderung darstellten – wurden im Prozess des Aufbereitens möglichst genau studiert und versucht in der Recodierung angemessen zu berücksichtigen, was dann z. B. zur Folge haben konnte, dass an gewissen Orten mehrere Varianten als basisdialektale Varianten ausgewiesen wurden.

Der Mehrwert dieser Recodierung besteht nun darin, dass man alle interessierenden dialektalen Variablen in der Hinsicht vergleichbar macht, als die jeweiligen Antwortkategorien – die in der Basiskategorisierung voneinander abweichen – miteinander vereinbar sind. Es ist also möglich, eine Aussage darüber zu treffen, inwiefern die vorliegend interessierenden Variablen sich hinsichtlich der Häufigkeit, mit der sie von den Probandinnen und Probanden basisdialektal, grossräumig und kleinräumig realisiert werden, ähneln oder eben unterscheiden. Die Homogenisierung ermöglicht fernerhin, dass der Gebrauch der Variablen nicht nur hinsichtlich der Unterschiede zwischen den Variablen, sondern auch im Hinblick auf situative Unterschiede (Fragebuchdaten vs. Spontandaten, vgl. Kap. 8.4.2) und aussersprachlicher Einflussgrössen (Bildung, Ort und Ortsloyalität, vgl. Kap. 8.2 und 8.4.6) untersucht werden kann.⁷³

9.3 Draw-a-map-tasks

Die zwei verschiedenen *draw-a-map-tasks*, die in vorliegender Studie verwendet wurden, und die Rückschlüsse erlauben über die *mental maps* der Probandinnen und Probanden, förderten zwei Datentypen zu Tage: handgezeichnete Karten und Kommentare zu diesen handgezeichneten Karten. Diese Datentypen wurden je unterschiedlich aufbereitet.

⁷³ Für die konkreten statistischen Berechnungen (vgl. Kap. 13) wurde die Codierung *basisdialektal*, *grossräumig*, *kleinräumig* weiter in die binäre Codierung *basisdialektal* und *grossräumig* überführt, was den quantitativen Umgang mit den Daten erleichterte. Alle Varianten, die in vorliegender Codierung als *kleinräumig* eingestuft wurden, wurden der Gruppe *basisdialektal* zugeteilt.

9.3.1 Handgezeichnete Karten

Mittlerweile hat es sich eingebürgert, für die Aufbereitung von handgezeichneten Karten (vgl. für ein Beispiel Abb. 10) mit geografischen Informationssystemen (GIS) zu arbeiten (vgl. für ältere Aufbereitungsmethoden Stoeckle 2014, S. 110–114). Ein geografisches Informationssystem kann dabei als «rechnergestütztes System» definiert werden, mit dem «raumbezogenen Daten/Informationen [...] digital erfasst und redigiert, verwaltet und reorganisiert, analysiert sowie alphanumerisch und grafisch präsentiert werden» können (Bill 2010, S. 8). Der Vorteil der Bearbeitung raumbezogener Daten auf Basis eines GIS liegt darin, dass die Daten georeferenziert (d. h. mit Koordinaten versehen und so mit der erdräumlichen Oberfläche in Beziehung gesetzt) sind und dadurch mit beliebig vielen weiteren Informationen verknüpft werden können. Überdies erlaubt das System die digitale Weiterverarbeitung und Berechnung der Daten (vgl. dazu auch Stoeckle 2014, S. 115). Vorliegend wurde mit dem Programm ArcGIS Desktop 10.5 (Copyright 1999–2016 Esri Inc.) des Softwareherstellers ESRI gearbeitet.⁷⁴

Zur digitalen Aufbereitung (vgl. für methodische Details der digitalen Kartenaufbereitung Stoeckle 2014, S. 117–123) wurden die handgezeichneten Karten⁷⁵ in einem ersten Schritt eingescannt, ins Programm ArcGIS integriert und dort georeferenziert. Der Prozess des Georeferenzierens wurde händisch getätigt: Drei bekannte Punkte der eingescannten Karten – die Ortspunkte Sarnen, Stans und Engelberg – wurden von Hand angewählt und mit bereits georeferenzierten Daten assoziiert. Anschliessend wurden die von den Probandinnen und Probanden markierten Gebiete abgezeichnet und als Vektordaten gespeichert: Jedes Dialektgebiet wurde so in ein sogenanntes Polygon überführt (vgl. für Informationen zu unterschiedlichen raumbezogenen Datentypen Stoeckle 2014, S. 114–117). Den Polygonen wurden danach eine Reihe von Attributen zugeordnet, die vorliegend von Interesse sind: die von den Probanden geäusserten Gebietsbezeichnungen, die damit verbundenen sprachlichen wie auch nicht sprachlichen Assoziationen. Nach diesem ersten Schritt der Aufbereitung lagen alle Daten, die mittels *draw-a-map-task* gesammelt wurden, in digitalisierter Form vor, inhaltlich aber nach wie vor in ihrer Rohform.

Der nächste Schritt der Aufbereitung bestand darin, aus den digitalisierten Karten *heatmaps* zu erstellen. Eine *heatmap* wird generiert aus der Überlagerung

⁷⁴ Ebenfalls in Betracht gezogen wurde die Open-Source-Applikation QGIS (<http://www.qgis.org/de/site/> (letzter Zugriff 10. 06. 2019)).

⁷⁵ Es sei hier angemerkt, dass diese Art der Aufbereitung vorliegend einzig für den *draw-a-map-task* zum Nahraum vorgenommen wurde. Der *draw-a-map-task* zum Grossraum, der in dieser Arbeit nur am Rande Beachtung findet, wurde anders aufbereitet.



Abb. 10: Handgezeichnete Karte der Probandin PB27 aus Lungen (reproduziert mit Bewilligung von swisstopo (BA19016))

der ermittelten Polygone und lässt so zu Tage treten, welche Polygone sich mehr, welche sich weniger stark decken. Solchen verschiedenen «Grade[n] an Überlagerung» (Stoeckle 2014, S. 119) lässt sich entnehmen, welche Sprachraumkonzepte stärker, welche weniger stark interindividuell repräsentiert sind. Sind die Berechnungen zur Überlagerung gemacht (vgl. hierzu Stoeckle 2014, S. 119–122), lassen sich den Häufigkeitswerten, mit denen sich die Polygone überlagern, unterschiedliche Farbwerte zuordnen, womit eine übersichtliche Darstellung der Form und Häufigkeit der sich überschneidenden Polygone gelingt.

Erstellt wurden die *heatmaps* mit ArcGIS Desktop 10.5 (Copyright 1999–2016 Esri Inc.). Die Geobasisdaten stammten vom Bundesamt für Landestopografie (swisstopo), als Koordinatensystem wurden die Schweizer Landeskoordinaten (CH1903_LV03) genutzt. Für die Klassifizierung der Daten der *heatmaps* wurde auf den automatischen Klassifizierungsalgorithmus nach Jenks/Caspall zurückgegriffen. Dieser Algorithmus sucht nach «natürlichen Unterbrechungen» in den Daten, anhand derer sich Klassengrenzen festlegen lassen:

Klassen vom Typ «Natürliche Unterbrechungen» basieren auf natürlichen Gruppierungen innerhalb der Daten. Es werden Klassengrenzen identifiziert, die ähnliche Werte möglichst gut gruppieren und die Unterschiede zwischen den Klassen maximieren. Es werden Grenzen an den Stellen gesetzt, wo die Daten relativ große Unterschiede aufweisen. (ArcGIS Pro, 2017)

Bei bis zu fünf Überschneidungen der Polygone wurde mit der jeweiligen Anzahl Klassen gearbeitet (2 Überschneidungen = 2 Klassen), ab 5 bis und mit 19 Überschneidungen wurden 5 Klassen erstellt. Ab 20 Überschneidungen wurden zur besseren Lesbarkeit 6 Klassen erstellt. Für die Präsentation wurde die Color Ramp Bright Red gewählt, die die Polygone umso stärker rot einfärbt, je mehr sie sich überschneiden, und umso schwächer, je weniger sie sich überschneiden (Transparenz: 30 %). Zudem wurde immer der gleiche Kartenmassstab gewählt, um auch diesbezüglich eine optische Vergleichbarkeit zu gewährleisten.

Mit diesen Aufbereitungsschritten wurde eine optimale Grundlage geschaffen, um Abfragen zu machen in Hinblick auf mental repräsentierte Sprachraumkonzepte bei einzelnen wie auch bei mehreren Sprechern, sowohl, was die räumliche Dimension (Flächen), als auch, was die inhaltliche Dimension (weitere Attribute) angeht. Insgesamt ist die Aufbereitung der handgezeichneten Karten zu *heatmaps* natürlich nicht unproblematisch, v. a., weil man als Aufbereitende/r das visuelle Bild, das sich aus den Karten ergibt, massgeblich beeinflussen kann. Vorliegend wurden diese visuellen Aufbereitungsschritte darum möglichst transparent gemacht. Nicht unproblematisch ist indes auch der Gesamtanspruch, den man mit solchen Aufbereitungsverfahren und anschliessenden Interpretationen verfolgt: nämlich, dass man aus *heatmaps* sprachraumstrukturierende Praktiken

der Probandinnen und Probanden ableiten kann. Diesbezüglich soll Folgendes noch einmal deutlich gemacht werden: Insgesamt muss davon ausgegangen werden, dass die wahrgenommenen Grenzen durch das jeweilige Raumverständnis der Probandinnen und Probanden erheblich geprägt sind, was deren Vergleichbarkeit einschränkt (vgl. dazu etwa Anders 2010b, S. 184). Zudem sind die handgezeichneten Karten der Probandinnen und Probanden an die Zeit gebunden, in welcher sie angefertigt wurden: Wahrscheinlich würden die Zeichnungen der Probandinnen und Probanden anders aussehen, hätte man sie zu einem anderen Zeitpunkt anfertigen lassen. Lameli et al. (2008, S. 64) schlagen aus diesem Grund vor, bei den Entitäten, die aus den handgezeichneten Karten abgeleitet werden können, eher von «Projektionen» als von «Repräsentationen» zu sprechen. Dennoch wurde auch in vorliegender Arbeit von der Methodik Gebrauch gemacht, handgezeichnete Karten zu *heatmaps* aufzubereiten, um daraus – mit der entsprechenden Vorsicht – interindividuelle räumliche Kategorien der Probandinnen und Probanden abzuleiten (vgl. dazu im Detail Kap. 10). Den eben thematisierten kritischen Argumenten gegen ein solches Vorgehen kann damit begegnet werden, dass individuelle Unterschiede in der räumlichen Wahrnehmung gerade durch den interindividuellen Fokus, d. h. durch die Anzahl verarbeiteter Karten, gemildert werden kann. Gegen den Punkt der Veränderbarkeit des projizierten räumlichen Wissens auf der Karte ist einzuwenden, dass grosse Teile gerade interindividuell geteilten Wissens über die Zeit stabil bleibt, weshalb die handgezeichneten Karten diesbezüglich auch eine gewisse Konstanz aufweisen.

9.3.2 Kartenkommentare

Nebst den handgezeichneten Karten förderten die *draw-a-map-tasks* als zweiten Datentyp Gespräche zu den handgezeichneten Karten zu Tage. Diese Gespräche – die für das Verständnis der räumlichen Alltagskategorien der Probandinnen und Probanden zentral sind (vgl. dazu im Detail Kap. 11) – wurden auf Grundlage der Audioaufnahmen inhaltlich transkribiert (vgl. für einen Einblick in die Daten die Transkripte zum *draw-a-map-task* zum Nahraum im digitalen Anhang unter 8).

Anschliessend wurden die Transkripte auf folgende Aspekte hin durchgesehen: Wie bezeichnen die Probandinnen und Probanden die von ihnen gezeichneten Gebiete? Welche sprachlichen Merkmale verbinden sie damit? Welche weiteren Wissensinhalte verbinden sie damit? Informationen zu diesen Fragen wurden tabellarisch aufgenommen und in ArcGIS übertragen: Dort konnten die Informationen mit den eingezeichneten Gebieten digital verknüpft werden. Überdies finden sich diese tabellarischen Auflistungen auch in der Datenbank.

9.4 Ortsloyalitätstest

Der Ortsloyalitätstest, der misst, wie stark sich die Probanden mit ihrer *in-group* identifizieren, wurde auf zwei Arten aufbereitet. Einerseits wurde über den Mittelwert⁷⁶ eruiert, wie die Ortsloyalität pro Proband im Durchschnitt beschaffen ist. Dieser Mittelwert pro Proband floss dann etwa in der quantitativen Auswertung der objektsprachlichen Daten in das Regressionsmodell ein (vgl. dazu im Detail Kap. 13.1.3.2).

Des Weiteren wurde mit den Daten eine Clusteranalyse durchgeführt, da interessierte, welche sinnvollen Gruppen von Probanden sich *data-driven* auf der Grundlage des Ortsloyalitätstests bilden lassen. Die Clusteranalyse gruppiert Untersuchungsobjekte dabei in natürliche Gruppen, sogenannte Cluster. Dabei soll jedes Cluster in sich möglichst homogen sein, während sich die Cluster untereinander möglichst stark voneinander unterscheiden sollen. Clusteranalytische Verfahren haben explorativen Charakter, da man keine inferenzstatistischen Rückschlüsse auf die Grundgesamtheit macht, sondern datengetrieben eine Struktur zu entdecken versucht. Die Forschenden spielen hierbei eine wichtige Rolle, da das Ergebnis unter anderem von der Wahl des Proximitätsmasses und des Clustering-Algorithmus beeinflusst wird (vgl. für detailliertere Informationen Brosius 2013, S. 711–744). Zur Gruppierung der Daten wurde das Proximitätsmass der *quadrierten euklidischen Distanz* und der Algorithmus *average linkage* verwendet. Auf dieser Basis konnten drei Ortsloyalitätsgruppen eruiert werden, die statistisch signifikant voneinander verschieden sind ($F(2,57) = 217.3$, $p = .000$, $n = 60$): eine wenig (3.1 Punkte), eine eher (5.2 Punkte) und eine sehr ortstloyale Probandengruppe (6.4 Punkte von min. 1 und max. 7 Punkten). Von diesen Ortsloyalitätsgruppen wurde im deskriptiven Teil der quantitativen Untersuchung Gebrauch gemacht: Die Boxplots zur Variation im Gebrauch der vorliegend interessierenden Sprachdaten wurden u. a. über diese Ortsloyalitätsgruppen erstellt (vgl. z. B. Kap. 13.1.3.1).

9.5 Zusammenfassung

Nebst der Erhebung der Daten bildet die Aufbereitung der Daten einen wichtigen Aspekt der empirischen Untersuchung. Vorliegend sind folgende Datentypen relevant und verlangten je unterschiedliche Aufbereitungsmethoden: die Sozial-

⁷⁶ Den Mittelwert (auch arithmetisches Mittel) erhält man, indem man alle Beobachtungswerte addiert und die Summe durch die Anzahl der Beobachtungen dividiert.

daten, die objektsprachlichen Realisierungen, die handgezeichneten Karten und dazugehörigen Kommentare und die Ergebnisse des Ortsloyalitätstests. Die Sozialdaten zu den 60 Probandinnen und Probanden wurden z. T. in Rohform, z. T. in numerischen Kategorien in Tabellen zusammengetragen. Die objektsprachlichen Realisierungen der Probandinnen und Probanden zu den 38 Dialektvariablen wurden ausgehend von den aufgezeichneten gesprochensprachlichen Äusserungen transkribiert und mittels eines an den Belegen des SDS orientierten und induktiv erweiterten Kategoriensystems geordnet: Somit konnte für jeden Probanden/für jede Probandin verzeichnet werden, welche dialektalen Varianten die Probandinnen und Probanden zu welchen Anteilen realisieren, was eine vollumfängliche qualitative Auswertung der Daten ermöglicht. Um die objektsprachlichen Realisierungen der Probandinnen und Probanden auch quantitativ auswerten zu können, wurden die Sprachdaten zusätzlich recodiert: Und zwar wurde für jede dialektale Variante bestimmt, inwiefern sie – für die acht sich in sprachlicher Hinsicht unterscheidenden Orte – pro Ort als basisdialektale, grossräumige oder kleinräumige Variante eingeschätzt werden kann. Die Referenzgrösse bildete hierbei stets der SDS. *Basisdialektal* meint, dass die Variante für den jeweiligen Ortspunkt als lokal gültige Variante ausgewiesen ist, *grossräumig* bedeutet, dass die Variante nicht in Bezug auf den fraglichen Ortspunkt, dafür aber für den Grossteil der Schweiz als Normvariante ausgewiesen wird, *kleinräumig* meint, dass die Variante nicht für den fraglichen, wohl aber für andere Orte im Untersuchungsgebiet belegt ist. Diese Recodierung der Daten ermöglichte, die Realisierungsvarianten der einzelnen Variablen zu homogenisieren und dadurch vergleichbar zu machen. Auf dieser Grundlage werden statistische Auswertungen möglich, die sonst nicht machbar gewesen wären. Die sprachkonzeptionellen Daten, die vorliegend interessieren, entstammen allesamt den *draw-a-map-tasks*. Die 60 handgezeichneten Karten, die im Rahmen des *draw-a-map-tasks* zum Nahraum entstanden, wurden anhand des Programms ArcGIS digital aufbereitet und in *heatmaps* überführt: Solche *heatmaps*, die mittels Farbgebung visualisieren, wo und wie stark sich die handgezeichneten Gebiete der Probandinnen und Probanden überschneiden, ermöglichen zu eruieren, wie die *mental maps* der Probandinnen und Probanden ausgestaltet sind. Nebst den handgezeichneten Karten förderten die *draw-a-map-tasks* als zusätzlichen Datentyp Gespräche zu diesen handgezeichneten Karten zu Tage. Diese 120 Gespräche wurden alle inhaltlich transkribiert, die relevanten Informationen wurden für einen besseren Überblick zusätzlich tabellarisch ausgezogen. Die Daten zum Ortsloyalitätstest schliesslich wurden so aufbereitet, dass für jeden Probanden ein Ortsloyalitätsindex angegeben werden konnte. Zudem wurden die Probandinnen und Probanden über eine Clusteranalyse in drei Ortsloyalitätsgruppen aufgeteilt. Die aufbereiteten Daten zum Ortsloyalitätstest flossen v. a. in den statistischen Teil der

Arbeit ein. Viele der aufbereiteten Daten sind – mit einem entsprechenden Passwortzugang – über eine sql-basierte Datenbank greifbar: Man kann dort etwa Einblick nehmen in die aktuellen objektsprachlichen Verhältnisse im Untersuchungsgebiet, die pro Ort und unter Angabe der zugrundeliegenden Daten in Kuchendiagrammen ausgegeben werden. Auch die handgezeichneten Karten und dazugehörigen Kommentare sind einsehbar; in ihrer Rohform wie auch in aufbereiteter Form.