RegioLingua

Zeitschrift für regionale Sprache und Literatur

Herausgegeben von
Doreen Brandt, Marina Frank, Robert Langhanke, Nicole Palliwoda und Jeffrey Pheiff

Jahrgang 3 (2026)

Heft 1

Magnus Breder Birkenes, Robert Engsterhold, Hanna Fischer und Jeffrey Pheiff (Hg.)
Digitale Dialektologie. Plattformen und Apps zu den deutschen Regionalsprachen

Stefan Rabanus

AlpiLinK. Korpus und Web-Portal

Zitationsvorschlag:

Rabanus, Stefan (2026): AlpiLinK. Korpus und Web-Portal. In: Birkenes, Magnus Breder/Engsterhold, Robert/Fischer, Hanna/Pheiff, Jeffrey (Hg.). *Digitale Dialektologie. Plattformen und Apps zu den deutschen Regionalsprachen*. = *RegioLingua. Zeitschrift für regionale Sprache und Literatur* 3 (1), 7–16. DOI: 10.17192/regiolingua.2026.3.1.9120.

https://journals.uni-marburg.de/regiolingua/
Kontakt: regiolingua@uni-marburg.de
ISSN: 2944-2133

Die Zeitschrift *RegioLingua* erscheint online über den Publikationsserver der Universitätsbibliothek Marburg unter der Creative Commons Lizenz CC-BY 4.0.

STEFAN RABANUS

AlpiLinK. Korpus und Web-Portal

AlpiLinK. Corpus and web portal

Zusammenfassung: AlpiLinK (*Alpine Sprachen im Kontakt*) ist ein Projekt zur Untersuchung und Valorisierung der Dialekte und Minderheitensprachen des italienischen Alpenbogens. In einer Datenerhebung mit Crowdsourcing-Methoden, die auch als Instrument für Citizen Science-Aktivitäten vor allem in den Schulen diente, wurde ein großes Korpus aus vorwiegend Tonaufnahmen geschaffen, das die vergleichende Beschreibung der romanischen, germanischen und slawischen Varietäten des Untersuchungsgebiets sowie die Analyse ggf. sprachkontaktinduzierten Sprachwandels ermöglicht, vor allem in Bezug auf die Morphosyntax.

Schlagwörter: Crowdsourcing, Citizen Science, Dialekte, Minderheitensprachen, Norditalien, Südtirol, Morphosyntax, Sprachkontakt, Sprachwandel, Sprachkartographie.

Abstract: AlpiLinK (*Alpine languages in contact*) is a project for the investigation and promotion of the dialects and minority languages in northern Italy. Data were collected with crowdsourcing methods which also acted as an instrument for Citizen Science activities mainly in schools. The large corpus, which primarily features audio recordings, allows for the comparative description of the Romance, Germanic and Slavic varieties of the area of investigation as well as the analysis of potentially contact-induced language change, mainly with respect to morpho-syntactic features.

Keywords: crowdsourcing, citizen science, dialects, minority languages, northern Italy, South Tyrol, morpho-syntax, language contact, language change, language mapping.

1. Gegenstand

Gegenstand von AlpiLinK (*Alpine Sprachen im Kontakt*) sind die Dialekte und Minderheitensprachen der Regionen des italienischen Alpenbogens, also von Piemont, dem Aostatal, der Lombardei, von Trentino-Südtirol, Venetien und Friaul-Julisch Venetien, darunter von besonderem Interesse für die deutsche Dialektologie die Tiroler Dialekte Südtirols und die historisch deutschen Minderheitensprachen:[1] Walserdeutsch, Fersentalerisch, Zimbrisch, Plodarisch, Zahrerisch, Tischelwangerisch und Kanaltal-Deutsch. Über die Web-Plattform des Projekts (https://alpilink.it/de/) wurden Sprachdaten gesammelt, die

[1] Grundlage für die Einstufung als Minderheitensprache ist das italienische Staatsgesetz 482 von 1999, das allerdings nur undifferenziert von *popolazioni germaniche* „germanischen Bevölkerungsgruppen“ spricht. Regionale Normen spezifizieren die Gruppen genauer. Die im Text genannte Klassifizierung in sieben Sprachen stammt von den Minderheitengemeinschaften selbst, die Walserdeutsch und Zimbrisch (trotz geringer Sprecherzahlen) in weitere Subvarietäten differenzieren. Ein Kurzabriss der Merkmale dieser Minderheitensprachen findet sich in Rabanus et al. (2019). Mit dem Staatsgesetz 482 werden auch Frankoprovenzalisch, Ladinisch, Friulanisch, Okzitanisch und die Varietäten der *popolazioni slovene* „slowenischen Bevölkerungsgruppen“ als Minderheitensprachen anerkannt.

RegioLingua. Zeitschrift für regionale Sprache und Literatur, 3 (2026), H. 1, 7–16.
DOI: 10.17192/regiolingua.2026.3.1.9120

auf der AlpiLinK-Plattform selbst und über ein Zenodo-Repositorium als AlpiLinK-Korpus bereitgestellt werden.

Das Ziel von AlpiLinK ist die digitale Dokumentation der norditalienischen Dialekte und Minderheitensprachen, also aller Varietäten, die verschieden von der italienischen, deutschen, französischen und slowenischen Standardsprache sind, welche im Gesamtgebiet oder in einzelnen Regionen (Südtirol, Aostatal, Friaul-Julisch Venetien) als Amtssprachen fungieren. In sprachwissenschaftlicher Hinsicht ist die Datensammlung vor allem für die Untersuchung von kontaktinduziertem Sprachwandel konzipiert. Ein zweites, dem sprachwissenschaftlichen gleichrangiges Ziel ist es, das Wissen um die Dialekte und Minderheitensprachen und ihre Reputation in der Gesellschaft zu vergrößern und gleichzeitig das Selbstwertgefühl der Minderheitengemeinschaften zu erhöhen. Eine wichtige Rolle dafür spielen Kooperationsprojekte mit den Schulen im Rahmen des Teilprojekts VinKiamo: Oberschüler waren als „Forschungsassistenten" aktiv in die Datensammlung eingebunden (Bertollo/Rabanus 2023; i. Dr.; Siviero et al. 2025; Alber et al. 2025b, 133–144).[2] Die gesammelten Daten können im Unterricht sowohl zur Thematisierung von Mehrsprachigkeit im Allgemeinen und den Dialekten und Minderheitensprachen Norditaliens im Besonderen als auch zur Heranführung der Schüler an Methoden der Datenanalyse dienen.

Zielgruppen von AlpiLinK sind folglich sowohl die wissenschaftliche Fachgemeinschaft (Dialektologie, Typologie, Sprachwandelforschung, Mehrsprachigkeitsforschung) als auch die Minderheitengemeinschaften und die allgemeine Öffentlichkeit. AlpiLinK führt in Bezug auf Ziele und Zielgruppen das Forschungsprojekt VinKo (*Varietäten im Kontakt*, 2018–2022) weiter, das diesen Ansatz für ein Teilgebiet (nur die Regionen Trentino-Südtirol und Venetien) entwickelt und erprobt hat (vgl. zu VinKo Kruijt et al. 2023a; 2023b). AlpiLinK wurde 2022–2025 vom italienischen Ministerium für Universität und Forschung als ‚Projekt relevanten nationalen Interesses' finanziell gefördert.[3] Beide Projekte, VinKo und AlpiLinK, wurden vom Autor dieses Beitrags geleitet.

2. Umsetzung

Die Daten des AlpiLinK-Korpus wurden zwischen dem 11. Juli 2023 und dem 4. Juli 2025 per Crowdsourcing mit einem Online-Fragebogen über die AlpiLinK-Plattform gesammelt. Die Erhebung war zweigeteilt: Im ersten Teil wurden Sozialdaten wie Alter, Geschlecht und Sprachrepertoire der ansonsten anonymen Sprecher abgefragt, zusätzlich Selbsteinschätzungen zur Kompetenz und Verwendung der vertretenen Varietät, außerdem eine kurze Stellungnahme zur individuellen Sprachbiographie. Im zweiten Teil wurden mit verschiedenen Aufgabentypen (Übersetzung aus der Standardsprache, Bildbeschreibung, Satzvervollständigung, Variantenauswahl) die eigentlichen Sprachdaten

[2] Der Name *VinKiamo* ist eine spielerische Ableitung vom Namen des weiter unten genannten Projekts VinKo. Während die Sammlung der Sprachdaten im engeren Sinne abgeschlossen ist, gibt es nach wie vor die Möglichkeit, Photographien zur Dokumentation der visuellen Mehrsprachigkeit hochzuladen, was weiterhin Projektarbeit in den Schulen ermöglicht, siehe AlpiLinK-Portal, Sektion „Linguistic Landscape", Link: https://alpilink.it/de/linguistic-landscape/ (02.11.2025).

[3] *AlpiLinK. German-Romance Language Contact in the Italian Alps: documentation, explanation, participation. Progetto di ricerca di rilevante interesse nazionale (PRIN)*, Ausschreibung 2020, Projektnr. 2020SYSYBS.

elizitiert, zum allergrößten Teil als Tonaufnahmen. Alle Daten sind georeferenziert über die offiziellen ISTAT-Namen der italienischen Verwaltungsgemeinden, mit den Mittelpunktkoordinaten (Längen- und Breitengrad) der Gemeinden und mit den GIDs, die die Gemeindegeographien im REDE SprachGIS indizieren.[4] Die Untersuchungsvariablen sind zum größten Teil morphosyntaktisch. Zu Details des Fragebogens siehe Rabanus (i. Dr.). Tabelle 1 zeigt eine Übersicht der gesammelten Daten, in denen die Tiroler Dialekte und die Gruppe der historisch deutschen Minderheitensprachen gesondert ausgewiesen sind. Diese deutschen Varietäten stellen nur knapp ein Viertel der Gesamtdaten dar, was aber in absoluten Zahlen dennoch fast 15.000 Tonaufnahmen entspricht und damit eine sehr hohe Dichte von Tonaufnahmen für ein geographisch kleines und zusätzlich nur dünn besiedeltes hochalpines Gebiet bedeutet. Die romanischen Varietäten, die mit gut drei Vierteln den Großteil der Daten ausmachen, sind in acht Gruppen klassifiziert: einerseits die Minderheitensprachen Okzitanisch, Frankoprovenzalisch, Ladinisch und Friulanisch, andererseits die piemontesischen, lombardischen, venetischen und trentinischen Dialekte ohne Status als Minderheitensprache. Die slawischen Minderheitensprachen entlang der Grenze zu Slowenien sind nur durch wenige Daten zum Resianischen und Kanaltal-Slowenischen vertreten, was aber dennoch interessante typologische Vergleiche und Beobachtungen zum kontaktinduzierten Wandel ermöglicht.

Kategorie		Tiroler Dialekte		Deutsche Minderheiten		Romanische Varietäten		Slawische Minderheiten		Gesamt
		Anzahl	*%*	*Anzahl*	*%*	*Anzahl*	*%*	*Anzahl*	*%*	*Anzahl*
Sprachdaten	*Audio*	9.178	15,9	5.049	8,7	43.327	74,8	331	0,6	57.885
	Text	2.003	16,5	207	1,7	9.928	81,8	6	< 0,1	12.144
Sprachbiographie	*Audio*	129	12,1	101	9,4	833	77,8	7	0,7	1.070
Gesamt		11.310	15,9	5.357	7,5	54.088	76,1	344	0,5	71.099

Tabelle 1: AlpiLinK-Korpus, Version 1.2.1. Die Prozentangaben beziehen sich auf den Anteil an der Gesamtzahl pro Kategorie

Alle Daten sind frei zugänglich und unter Einhaltung der Regeln der Creative Commons-Lizenz „Nicht kommerziell-Share Alike 4.0 International“ (CC BY-NC-SA 4.0) verwendbar. Die Teilnehmer der Crowdsourcing-Datenerhebung haben explizit der wissenschaftlichen Nutzung und Weiterverbreitung der Daten gemäß dieser Lizenz zugestimmt, auch der Nutzung ihrer Tonaufnahmen, die als persönliche Daten gelten, weil die Stimme die Identifizierung der Person ermöglichen kann. AlpiLinK stellt die Daten über zwei Kanäle zur Verfügung, die aktualisiert wurden und werden, um einerseits neue (bis zum Sommer 2025 aufgezeichnete) Tonaufnahmen zu integrieren und andererseits fehlerhafte oder inadäquate Tonaufnahmen zu löschen. Alle Tonaufnahmen werden manuell kontrolliert, die Kontrollen werden voraussichtlich bis Anfang 2027 andauern, weshalb das Korpus erst dann seine Endform erreichen wird. Bis zur Version 1.2.1 (publiziert am 4. Februar 2026)

[4] ISTAT steht für *Istituto Nazionale di Statistica*, d. h. nationales italienisches Statistikinstitut, welches die offiziellen Geographiedaten Italiens pflegt. Die Mittelpunktkoordinaten sind nicht direkt der ISTAT-Datenbank entnommen, sondern aus den im REDE SprachGIS übernommenen Geographien im Koordinatenreferenzsystem EPSG:4326 berechnet.

liegt das AlpiLinK-Korpus in insgesamt 16 Versionen vor, die alle vollständig auf Zenodo verfügbar sind.[5] Das Korpus enthält alle Daten und Metadaten (als flac-Tondateien und csv-Textdateien) und die für ihre Analyse notwendigen Zusatzinformationen in englischer Sprache. Jede Tondatei ist mit einem eindeutigen Code aus Stimulusnummer, Varietätensigle und Informantennummer – z. B. I01_frp_U0860 für die frankoprovenzalische (frp) Beschreibung des Bildstimulus I01 (siehe Abb. 1) durch die Informantin U0860 (ein 18-jähriges Mädchen aus Cogne im Aostatal) – identifizierbar und zitierbar. Trotz der prinzipiell freien Zugänglichkeit für alle ist das Korpus für die wissenschaftliche Fachgemeinschaft konzipiert und ohne korpuslinguistische Kompetenz kaum sinnvoll nutzbar. Auf dem AlpiLinK-Portal (https://alpilink.it/de/) – technisch ein WordPress-Blog – gibt es dagegen einen Zugang zum größten Teil der Tonaufnahmen über eine per iframe in WordPress eingebundene interaktive Karte in der Sektion „Anhören & Entdecken“,[6] die wie alle anderen Sektionen in den vier Sprachen Italienisch, Deutsch, Englisch und Französisch verfügbar ist. Es handelt sich um eine mit der Leaflet-Bibliothek erstellte Karte, auf der die Grenzen der Verwaltungsgemeinden in GeoJSON-Layern als Polygone dargestellt werden, welche nach Auswahl eines Stimulus (per Ausklappmenü) die Farbkodierung der Sprachvarietäten annehmen, denen die Tonaufnahmen der Gemeinde zugehören. Nutzer können in der Karte frei zoomen und zu vorwiegend für sie interessanten Regionen navigieren. Per Mausklick auf das Gemeindepolygon öffnet sich ein Fenster mit einem Audioplayer zum Abspielen der Tonaufnahmen zum gewählten Stimulus. Die Karte ist für die Nutzung seitens der allgemeinen Öffentlichkeit konzipiert und im Schulkontext erprobt.

3. Ergebnisse

Die erste systematische Auswertung von AlpiLinK-Daten für das gesamte Untersuchungsgebiet findet sich bei Alber et al. (2025a) zur Vorhersagbarkeit der Muster zur Kurznamenbildung. Weitere Ergebnisse sind auf Konferenzen vorgestellt worden und teilweise im Druck (zum Beispiel Rabanus i. Dr. zur Verwendung des Diminutivs und zur Kongruenz bei Possessivpronomen), eine umfangreiche systematische Auswertung ist unter dem Arbeitstitel *Kleiner Sprachatlas der norditalienischen Dialekte und Minderheitensprachen* derzeit im Gange. Zur Illustration des Erkenntnispotenzials soll hier die geolinguistische Auswertung des AlpiLinK-Korpus in Bezug auf die Verwendung eines Lokativs mit dem Verbkonzept ‚schneiden‘ dienen, die in Abbildung 2 visualisiert ist.

[5] Link zum Gesamtkorpus: https://doi.org/10.5281/zenodo.8360169 (02.11.2026). Alle Veränderungen in Bezug auf die vorhergehende Version sind in der readme-Datei der jeweiligen Version dokumentiert.
[6] 48.028 der 57.885 Sprachdaten-Tonaufnahmen sind über die Karte abspielbar. Nicht verfügbar sind textbasierte Daten, Sozialdaten der Sprecher, die Tonaufnahmen der Sprachbiographien und Tonaufnahmen von Stimuli, die nicht für das gesamte Untersuchungsgebiet, sondern nur für Teilgebiete erhoben wurden.



Abbildung 1: AlpiLinK-Bildstimulus I01

Zur Beschreibung der in Abbildung 1 dargestellten Szene verwenden Sprecher Tiroler Dialekte größtenteils das Verb *aufschneiden* (z. B. *a Mon schneid die Kaminwurz auf*, I01_tir_U1784, ‚ein Mann schneidet die Kaminwurz [Wurstsorte] auf') mit dem trennbaren Präfix *auf-*, das eine nach oben gerichtete bzw. öffnende Bedeutung hat (vgl. Fleischer/Barz 2012: 404), in der Karte in Abbildung 2 grün symbolisiert. In den romanischen Varietäten des Westens (Piemont, Aostatal) und des Ostens (Friaul-Julisch Venetien, zentrales und östliches Venetien) dominieren einfache Verben ohne Lokativ, etwa dialektale Varianten von *tagliare* ‚schneiden', z. B. in Gemona del Friuli *l'om al taiá salam* (I01_fur_U0915), wörtlich ‚der Mann er schneidet Wurst', in der Karte schwarz. Viele romanische Varietäten des Trentino, des westlichen Venetiens und der Lombardei fügen dem Verb dagegen ein lokatives Adverb hinzu, das meist eine distanzierende bzw. nach unten gerichtete Bedeutung hat (vgl. Fleischer/Barz 2012: 399–400; Cordin 2011: 105), z. B. in Trient *uno che taia zo luganega* (I01_tre_U0464), wörtlich ‚einer der schneidet ab Luganega (Wurstsorte)', in der Karte rot symbolisiert.

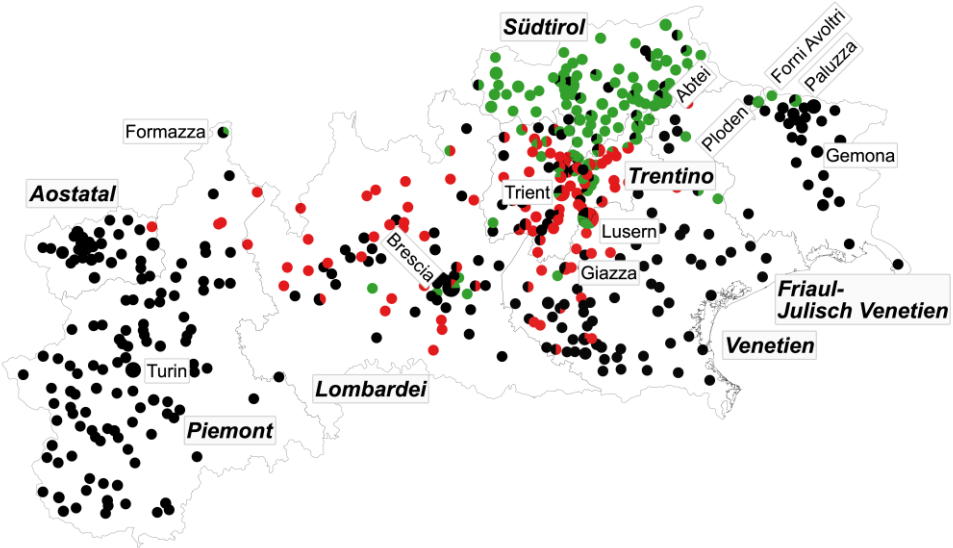

Abbildung 2: Karte der Verteilung der morphosyntaktischen Varianten für das Verbkonzept ‚schneiden' in Beschreibungen der in Abbildung 1 gezeigten Szene. Die Größe der Symbole ist proportional zur Quantität der Daten pro Ortspunkt. Grün: Lokativ mit öffnender Bedeutung (‚auf'); rot: Lokativ mit distanzierender Bedeutung (‚ab'); schwarz: kein Lokativ, nur Verb. Die Grenzlinien repräsentieren die norditalienischen Regionen und autonomen Provinzen. Datenbasis: AlpiLinK-Korpus 1.1.5

Auffällig ist nun, dass in den ladinischen Varietäten Südtirols, deren Sprecher in vielen Fällen auch Deutsch bzw. Tiroler Dialekt auf muttersprachlichem Niveau beherrschen, und in den trentinischen Varietäten an der Grenze zu Südtirol ein Lokativ mit öffnender, nach oben gerichteter Bedeutung verwendet wird (in der Karte grün), z. B. im Ladinischen des Südtiroler Abteitals *sö* (Entsprechung zu ital. *su* ‚auf'), z. B. *l ël taié sö l scialam* (I01_lld_U0237), wörtlich ‚der Mann schneidet auf die Wurst'. Dasselbe gilt für die beiden einzigen friulanischen Belege mit Lokativ: Es handelt sich um *taiá su* in Paluzza (I01_fur_U0798) und *pestá su* in Forni Avoltri (I01_fur_U0892). Beide Orte liegen in unmittelbarer Nähe zur österreichischen Grenze und zu den deutschen Sprachinseln Ploden und Tischelwang (Teil der Gemeinde Paluzza), wo wie in Südtirol in der Regel das Präfix *auf-* verwendet wird (alle genannten Orte sind auf der Karte in Abbildung 2 mit Namen verzeichnet). Umgekehrt ist im Zimbrischen, dessen Sprecher in der Regel auch kompetent bzw. sogar dominant im lokalen venetischen Dialekt sind, ein hoher Anteil an Lokativen mit distanzierender bzw. nach unten gerichteter Bedeutung belegt (*abe* ‚ab', *nidar* ‚nieder'). Diese Situation ist das Ergebnis eines sprachkontaktinduzierten morphosyntaktischen Wandels, welcher sprachgruppenübergreifende Raumstrukturen gebildet hat. Das tirolisch-deutsche Muster mit öffnender Bedeutung (in der Karte grün symbolisiert) ist auf die romanischen Kontaktvarietäten, und das romanische Muster mit distanzierender Bedeutung (rot symbolisiert) auf das Zimbrische übertragen worden, wobei jede Sprache ihre spezifischen Wortformen bewahrt: Phonologisches Material wird nicht entlehnt, es liegt also *pattern replication* vor. Im Zimbrischen von Giazza und teilweise von Lusern hat auch die Syntax kontaktinduzierten Wandel erfahren, weil der Lokativ kein Verbpräfix mehr ist, sondern ein Adverb, das wie in den romanischen Kontaktdialekten immer rechts vom lexikalischen Verb steht. Ein Beispiel aus dem Zimbrischen von Lusern ist der Infinitiv in der Progressivkonstruktion *dar mon is no tzo hakha nidar an toko salado* (I01_cim_U0014), wörtlich ‚der Mann ist nach (‚dabei') zu hacken nieder ein Stück Wurst' genau wie im trentinischen Dialekt von Novella *sto om l'è drio a taiar zo na lugiangia* (I01_tre_U0061), wörtlich ‚dieser Mann er ist nach zu schneiden ab eine Luganega'.[7] Das AlpiLinK-Korpus ermöglicht einen solchen Nachweis kontaktinduzierten Wandels aufgrund der Anlage als Korpus mit sprachübergreifend oder, bei Bildbeschreibungen wie hier im Beispiel, sprachunabhängig elizitierten Varianten. Weitere vergleichbare Analysen sind in Arbeit.

4. Ausblick

Die Kontrolle der bereits in der Version 1.2.1 publizierten und auch über die interaktive Karte zugänglichen Korpusdaten ist noch nicht abgeschlossen. Die Endversion des AlpiLinK-Korpus wird daher voraussichtlich erst Anfang 2027 veröffentlicht. Das Web-Portal soll auch darüber hinaus von der Veroneser Arbeitsgruppe gepflegt werden. Aktualisierungen betreffen besonders zwei Sektionen: In „Ergebnisse" werden Kurzfassungen

[7] Auch im Tischelwangerischen gibt es erste Anzeichen für syntaktischen Wandel durch die Aufgabe der Verbklammer bei einigen Informanten, z. B. *der mon schneit auf in salam* (I01_tis_U0648) ‚der Mann schneidet auf die Wurst', während im Plodarischen die Verbklammer bewahrt ist: *der Mario schneit in salam auf* (I01_plo_U0722). Siehe dazu vergleichbare Beobachtungen von Poletto/Tomaselli (2018: 122).

wissenschaftlicher Artikel in für die allgemeine Öffentlichkeit zugänglicher Form (und in den drei Sprachen Italienisch, Deutsch, Englisch) und Beschreibungen ausgewählter Phänomene mit direktem Link zu Tondateien angeboten (zum Beispiel die „Schaufenster der Regionalsprachen Norditaliens“: dasjenige zu den Progressivkonstruktionen enthält auch eine instruktive Karte). In „Unsere Varietäten“ finden sich Kurzcharakterisierungen der 18 Varietätengruppen mit Literaturhinweisen und Listen (möglichst im Internet verfügbarer) sprachbezogener Ressourcen wie Wörterbücher, Textsammlungen oder grammatischen Kompendien, die kontinuierlich aktualisiert werden.

Die interaktive Karte ist der Ausgangspunkt für ein Folgeprojekt mit dem Namen corDATA, was ein Akronym für *Coordinamento per il raccordo di database georeferenziati* ‚Koordinierung der Verknüpfung georeferenzierter Datenbanken‘ ist (https://progetto-cordata.it). Es handelt sich um eine Plattform, auf der Sprachdaten aller Arten und Formate über ihre geographischen Koordinaten aufeinander beziehbar sind. Referenzpunkt sind wie bei der interaktiven AlpiLinK-Karte die italienischen Verwaltungsgemeinden, bei corDATA für das gesamte italienische Staatsgebiet. Such-, Darstellungs- und Exportfunktionen können in Bezug auf Gemeinden, Provinzen, Regionen oder vom Nutzer in der Karte mit Freihandtools ausgewählte Gebieten genutzt werden. Ziel ist es, sowohl bereits digital publizierte Datenbanken als auch und vor allem solche Datensätze zugänglich zu machen, die entweder (a) noch nie publiziert worden sind (es gibt in Italien zahlreiche Sprachatlasprojekte, die bisher nicht zu Kartenwerken geführt haben), (b) noch nie digital publiziert worden sind und (c) deren digitale Publikation nicht mehr online ist, weil die Webseiten nicht mehr gepflegt werden konnten (z. B. ALIQUOT und VIVALDI; vgl. zur Problematik der Langzeitsicherung digitaler Daten Kruijt/Rabanus 2025). Der Prototyp der corDATA-Plattform[8] verknüpft die Audiodaten von AlpiLinK und die textkodierten Daten (Transkriptionen von Wortformen, Paradigmatafeln) der großen Datenerhebung von Manzini/Savoia (2005), für die es ebenfalls eine separate Web-Plattform gibt (https://manzinisavoia.changes.unimi.it). Die Wahl dieser beiden Projekte für den Prototyp der corDATA-Plattform ist darin begründet zu zeigen, dass eine Verknüpfung über ein rein geographisches Referenzsystem möglich und sinnvoll auch in den Fällen ist, in denen weder der Datentyp (Audio vs. Text) noch das theoretische Referenzsystem bzw. die Definition der Variablen übereinstimmt (AlpiLinK ist deskriptiv-typologisch ausgerichtet, Manzini/Savoia 2005 generativ). CorDATA versteht sich als Meta-Datenbank, bei der Autorschaft und Spezifizität der Ausgangsprojekte expliziert und getreu abgebildet werden. Es werden nur Daten von Projekten integriert, die den Schutz der personenbezogenen Daten und die Beachtung des Copyrights garantieren. Inhaltliche Verknüpfungen (etwa projektübergreifende Suchmöglichkeiten nach morphosyntaktischen Kategorien, z. B. ‚Subjektpronomen‘, ‚expletiver Artikel‘ oder ‚Negationsmarker‘) sind erwünscht, allerdings nicht prioritär und angesichts der erheblichen Schwierigkeiten, sich auf für die unterschiedlichen theoretischen Referenzsysteme gleichermaßen adäquate Kategorien zu einigen, wohl nur auf längere Sicht und mit besonderer Finanzierung möglich. Die

[8] Der Prototyp der corDATA-Plattform wurde von Carlo Zoli und dem Autor dieses Beitrags am 8. September 2026 auf dem Kongress der SIDG am Forschungszentrum Deutscher Sprachatlas in Marburg vorgestellt.

corDATA-Initiative ist in einem Vertrag zwischen den Universitäten Verona, Florenz und Bozen verankert, mit dem die dauerhafte Zugänglichkeit der Plattform sichergestellt und der bei Einzelprojekten, die von zeitlich begrenzten Projektfinanzierungen und individuellen Initiativen abhängen, stets drohende Offline-Status vermieden werden soll.

Referenzen

Alber, Birgit/Arndt-Lappe, Sabine/Kokkelmans, Joachim (2025a): The Predictability of Name Truncation: Factoring in Language Change. In: *Catalan Journal of Linguistics* 24 (1), 7–39. DOI: 10.5565/rev/catjl.467.

Alber, Birgit/Bonelli, Angelica/Kokkelmans, Joachim/Pilsbacher, Anna Katharina (2025b): Crowdsourcing and Citizen Science in Linguistics: New Ways to Study Non-Standard and Minority Languages. In: *L'Analisi Linguistica e Letteraria* 38, 127–148.

AlpiLinK = Rabanus, Stefan/Kruijt, Anne/Alber, Birgit/Bidese, Ermenegildo/Gaeta, Livio/Raimondi, Gianmario (2023ff.): *AlpiLinK. German-Romance language contact in the Italian Alps: documentation, explanation, participation*. In Zusammenarbeit mit Paolo Benedetto Mas, Sabrina Bertollo, Serena Bissolo, Angelica Bonelli, Dario Capelli, Jan Casalicchio, Raffaele Cioffi, Patrizia Cordin, Michele Cosentino, Silvia Dal Negro, Ilaria Driussi, Alexander Glück, Joachim Kokkelmans, Adriano Murelli, Andrea Padovan, Aline Pons, Matteo Rivoira, Marta Tagliani, Caterina Saracco, Alessandra Tomaselli, Ruth Videsott, Alessandro Vietti und Barbara Vogt. https://alpilink.it [Stand: 02.11.2026].

AlpiLinK-Korpus = Rabanus, Stefan/Kruijt, Anne/Alber, Birgit/Bidese, Ermenegildo/Gaeta, Livio/Raimondi, Gianmario (2026): *AlpiLinK Corpus 1.2.1*. In Zusammenarbeit mit Paolo Benedetto Mas, Sabrina Bertollo, Serena Bissolo, Angelica Bonelli, Dario Capelli, Jan Casalicchio, Raffaele Cioffi, Patrizia Cordin, Silvia Dal Negro, Ilaria Driussi, Sara Erriu, Alexander Glück, Joachim Kokkelmans, Adriano Murelli, Andrea Padovan, Aline Pons, Matteo Rivoira, Marta Tagliani, Caterina Saracco, Alessandra Tomaselli, Ruth Videsott, Alessandro Vietti und Barbara Vogt. DOI: 10.5281/zenodo.15524879.

Bertollo, Sabrina/Rabanus, Stefan (2023): VinKiamo: ein Citizen-Science-Projekt für Schulen zur Förderung von (sprach-)übergreifenden Kompetenzen: In: *Apprentissage des langues et systèmes d'information et de communication* 26 (1). DOI: 10.4000/alsic.7076.

Bertollo, Sabrina/Rabanus, Stefan (i. Dr.): Promoting minority languages at school: VinKiamo in Northern Italy. In: Gaeta, Livio/Raimondi, Gianmario (Hg.): *Safeguarding Varieties in Contact. Scaffolding Identities and Planning Minority Languages in Europe*. Berlin/Boston: De Gruyter.

Cordin, Patrizia (2011): *Le costruzioni verbo-locativo in area romanza. Dallo spazio all'aspetto*. Berlin/Boston: De Gruyter.

Fleischer, Wolfgang/Barz, Irmhild (2012): *Wortbildung der deutschen Gegenwartssprache*. 4. Auflage; völlig neu bearbeitet von Irmhild Barz unter Mitarbeit von Marianne Schröder. Berlin/Boston: De Gruyter.

Kruijt, Anne/Cordin, Patrizia/Rabanus, Stefan (2023a): On the validity of crowdsourced data. In: Pustka, Elissa/Quijada Van den Berghe, Carmen/Weiland, Verena (Hg.): *Corpus Dialectology*. Amsterdam/Philadelphia: Benjamins, 10–33. DOI: 10.1075/scl.110.01kru.

Kruijt, Anne/Rabanus, Stefan (2025): From VinKo to AlpiLinK: web-based long-term storage and accessibility of information. In: Krefeld, Thomas/Lücke, Stefan (Hg.): *Berichte aus der digitalen Geolinguistik* (IV). (Korpus im Text, 17). https://www.kit.gwi.uni-muenchen.de/?p=106395&v=1 [Stand 01.11.2026].

Kruijt, Anne/Rabanus, Stefan/Tagliani, Marta (2023b): The VinKo Corpus. Oral data from Romance and Germanic local varieties of Northern Italy. In: Kupietz, Marc/Schmidt, Thomas (Hg.): *Neue Entwicklungen in der Korpuslandschaft der Germanistik. Beiträge zur IDS-Methodenmesse 2022*. Tübingen: Narr Francke Attempto, 203–212. DOI: 10.24053/9783823396024.

Manzini, Maria Rita/Savoia, Leonardo Maria (2005): *I dialetti italiani e romanci: morfosintassi generativa*. Alessandria: Edizioni dell'Orso.

Poletto, Cecilia/Tomaselli, Alessandra (2018): Grammatische Konvergenz. Entwicklungspfade deutscher Sprachinseln im Nordosten Italiens. In: *Germanistische Linguistik* 239–240, 117–143.

Rabanus, Stefan (i. Dr.): Alpine Languages in Contact (AlpiLinK): from crowdsourcing data to linguistic maps. In: Gaeta, Livio/Raimondi, Gianmario (Hg.): *Documenting Varieties in Contact. Multidisciplinary Approaches to Minority Languages in Europe*. Berlin/Boston: De Gruyter.

Rabanus, Stefan/Bidese, Ermenegildo/Dal Negro, Silvia (2019): Deutsch als Minderheitensprache in Italien. In: Herrgen, Joachim/Schmidt, Jürgen Erich (Hg.): *Sprache und Raum. Ein internationales Handbuch der Sprachvariation. Bd. 4: Deutsch*. Berlin/Boston: De Gruyter Mouton (Handbücher zur Sprach- und Kommunikationswissenschaft, 30.4), 1096–1114. DOI: 10.1515/9783110261295-042.

Siviero, Emily/Alber, Birgit/Kokkelmans, Joachim (2025): Dialektforschung und Schule: Das Projekt VinKiamo Südtirol. In: Hochstadt, Christiane/Schweigkofler-Kuhn, Anny (Hg.): *Dialekt in der Lehre. Sprachdidaktische und varietätenlinguistische Perspektiven*. Tübingen: Narr Francke Attempto, 165–181. https://www.narr.de/dialekt-in-der-lehre-1090-1 [Stand: 02.11.2026].