

apropos

[Perspektiven auf die Romania]

Sprache/Literatur/Kultur/Geschichte/Ideen/Politik/Gesellschaft

Con parole tue
Dai parlanti a VerbaAlpina attraverso il Crowdsourcing

Beatrice Colcuc & Anna Rodella

apropos [Perspektiven auf die Romania]

hosted by Hamburg University Press

2022, 9

pp. 187-212

ISSN: 2627-3446

Online

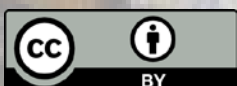
<https://journals.sub.uni-hamburg.de/apropos/article/view/1877>

Zitierweise

Colcuc, Beatrice & Anna Rodella. 2022. „Con parole tue. Dai parlanti a VerbaAlpina attraverso il Crowdsourcing.“ *apropos [Perspektiven auf die Romania]* 9/2022, 187-212.

doi: <https://doi.org/10.15460/apropos.9.1877>

Except where otherwise noted, this article is licensed under a Creative Commons Attribution 4.0 International license (CC BY 4.0)



Beatrice Colcuc & Anna Rodella

Con parole tue

Dai parlanti a VerbaAlpina attraverso il crowdsourcing

Beatrice Colcuc

è ricercatrice presso VerbaAlpina,
Ludwig-Maximilians-Universität
München.

beatrice.colcuc@lmu.de

Anna Rodella

è stata collaboratrice presso
VerbaAlpina, Ludwig-Maximilians-
Universität München.

arodella@live.it

Parole chiave

VerbaAlpina – Crowdsourcing – Digital Humanities – Geolinguistica – Digitalizzazione

1. Introduzione

Con i suoi 1300 km di distensione territoriale, la catena delle Alpi rappresenta il sistema montuoso più importante d'Europa. Dal punto di vista topografico ed etnografico, le Alpi appaiono sostanzialmente compatte: cultura, tradizioni, flora, fauna e formazioni paesaggistiche si ritrovano lungo tutto il crinale alpino dalla Francia alla Slovenia passando per la Svizzera, il Liechtenstein, l'Austria, la Germania e l'Italia. Decisamente più eterogenea si mostra invece la situazione linguistica, legata non solo alle lingue nazionali (francese, tedesco, italiano, retoromano e sloveno), ma soprattutto relativa all'assetto dialettale dell'area. La varietà di idiomi parlati nel territorio alpino è straordinariamente grande e non solo ogni regione o ogni provincia, ma spesso ogni piccolo villaggio detiene una parlata locale dalle caratteristiche peculiari. Il pluralismo linguistico a livello sincronico, è frutto però di una storica unità linguistico-culturale che il progetto *VerbaAlpina* dell'Università di Monaco di Baviera tenta di analizzare e di mettere in rilievo.

Il presente contributo rappresenta la versione approfondita e aggiornata della relazione tenuta dalle stesse autrici in occasione del trentasettesimo *Deutscher Romanistentag*, organizzato dall'università di Augusta e tenutosi online dal 4 al 7 ottobre 2021. L'occasione del Romanistentag ha fornito un'importante opportunità

di presentare dati relativi alle attività e alle modalità di raccolta dei dati linguistici da parte di *VerbaAlpina* in una prospettiva *work in progress*. Come si evincerà dal presente scritto, *VerbaAlpina* opera una raccolta dati mediante una piattaforma di *crowdsourcing* interna e, al contempo, rivolgendosi a un portale esterno. Il presente contributo concentrerà la propria attenzione sulla piattaforma di *crowdsourcing* interna, portando a conoscenza del lettore il contesto, gli obiettivi e alcuni risultati scaturiti da una breve indagine sull'utilizzo della piattaforma da parte dei parlanti dei comuni dell'arco alpino. Dopo la presentazione generale del progetto e delle sue finalità, il contributo si concentra sul *crowdsourcing* quale modalità di raccolta dati, approfondendo alcuni aspetti specialmente legati a *VerbaAlpina*.

2. VerbaAlpina - Der alpine Kulturraum im Spiegel seiner Mehrsprachigkeit

VerbaAlpina è un progetto di ricerca della *Ludwig-Maximilians-Universität* (LMU) di Monaco di Baviera finanziato dalla *Deutsche Forschungsgemeinschaft* (fondazione tedesca per la ricerca) a partire da ottobre 2014 con una prospettiva di durata fino al 2026. Il progetto è stato concepito e viene portato avanti in collaborazione dall'*Institut für Romanische Philologie* e dall'*IT-Gruppe Geisteswissenschaften* (Centro di Tecnologia dell'Informazione per le Scienze Umane) della LMU, e si iscrive all'interno dell'area di studi delle *digital humanities* (umanistica digitale) (cf. Krefeld & Lücke 2014a; Krefeld 2019a).

VerbaAlpina è un atlante linguistico digitale dell'area alpina le cui attività sono state suddivise in tre fasi temporali: la prima fase (10/2014 - 10/2017) è stata dedicata al lessico relativo alla gestione dei pascoli alpini, concedendo particolare attenzione all'ambito della lavorazione del latte. La successiva seconda fase (11/2017-10/2020) ha visto la consacrazione agli ambiti lessicali della flora, della fauna, delle formazioni paesaggistiche e della meteorologia alpina. La terza fase di elaborazione, iniziata a novembre 2020, si concluderà a ottobre 2023 e ha come oggetto di indagine il lessico dell'ambiente di vita moderno, con un occhio di riguardo all'ecologia e al turismo nelle Alpi. Questa terza fase, seppur affine alle prime due in termini di concezione ed elaborazione, ha visto l'introduzione di alcune novità relative soprattutto alla raccolta dei dati e alle modalità di collocazione del progetto all'interno del territorio alpino di cui sarà fatta relazione in questo scritto.

2.1 Concezione e presentazione generale del progetto

La concezione e la delimitazione metodologica di *VerbaAlpina* sono state mosse dal desiderio di superare alcune limitazioni imposte sia dalla tradizione di ricerca in geolinguistica come anche dagli strumenti e dalle modalità di ricerca tipiche del periodo pre-digitale. Solitamente, gli atlanti linguistici si limitano a un'area linguistica o politica nazionale in particolare, oppure prendono in considerazione solo una varietà linguistica esistente all'interno di una determinata area. La conseguenza di tali limitazioni si profila nell'idea illusoria di tratti linguistici che si arrestano improvvisamente entro i confini dell'area presa in considerazione,

oppure di un monolinguisimo che non rispecchia l'effettiva situazione linguistica della zona.

Per *Verba Alpina*, il superamento delle limitazioni della geolinguistica tradizionale significa anzitutto non limitarsi a un'area linguistica o politica nazionale (intesa come Stato-Nazione) in particolare, ma si è voluto estendere l'area di ricerca a tutta la regione montuosa dell'arco alpino. Ciononostante, agli inizi del progetto si rendeva necessaria una delimitazione della zona sottoposta a ricerca: ci si è così affidati al perimetro della *Convenzione delle Alpi*¹ al fine di analizzare l'unità storico-culturale che caratterizza questa regione montuosa e poter estrapolare le similitudini e i punti di contatto che intercorrono tra le diverse varietà linguistiche e, di conseguenza – ma soprattutto –, tra le tre diverse famiglie linguistiche delle Alpi: la romanza, la germanica e la slava.

Proprio nell'ottica di osservare il grado di vicinanza delle varietà linguistiche delle tre famiglie linguistiche alpine, è risultato essenziale innanzitutto riunire il materiale necessario all'analisi. I dati lessicali sono forniti da un lato da atlanti linguistici, cartacei e digitali, pubblicati nel corso degli anni, ai quali si sono aggiunti anche dizionari dialettali locali, e dall'altro lato da dati provenienti dalla piattaforma di *crowdsourcing* di cui verrà fatta relazione di seguito. I dati contenuti negli atlanti cartacei hanno necessitato di un processo di digitalizzazione attraverso un sistema di trascrizione basato esclusivamente su caratteri ASCII (cf. Krefeld 2019b). Una volta inserite nella banca dati attraverso la trascrizione (oppure attraverso la piattaforma di *crowdsourcing*), le espressioni subiscono un processo di tokenizzazione, mediante il quale esse vengono suddivise in singole parole per poi essere successivamente tipizzate, ovvero raggruppate sotto un'unica forma (un tipo morfo-lessicale) rappresentativa di tutte le sottovarianti secondo le loro caratteristiche grammaticali. La tipizzazione del materiale lessicale rappresenta la *conditio sine qua non* per strutturare i dati e per facilitare la loro lettura e la loro analisi in termini linguistici. Un tipo morfo-lessicale viene definito dalla concordanza delle seguenti proprietà: famiglia linguistica, categoria lessicale, parola semplice vs. parola affissa, genere, tipo di base lessicale. Lo schema seguente riassume il funzionamento della tipizzazione:

¹ La *Convenzione delle Alpi* è un trattato internazionale sottoscritto dai Paesi alpini di Austria, Francia, Germania, Italia, Liechtenstein, Principato di Monaco, Slovenia e Svizzera e dalla Comunità economica europea, con l'obiettivo di garantire una politica comune per l'arco alpino. Dal punto di vista metodologico, essa possiede un importante significato per *Verba Alpina*, in quanto il progetto ha adottato, fissandola come area di ricerca, la definizione di 'regione alpina' basata sulle frontiere geografiche e amministrative delimitate proprio dalla *Convenzione delle Alpi*.

Varianti fonetiche	<i>barga</i>	<i>bark</i>	<i>bargúnj</i>	<i>margún</i>
Famiglia linguistica	roa	roa	roa	roa
Categoria lessicale	sub	sub	sub	sub
Affisso	-	-	+	+
Genere	f	m	m	m
Tipo morfo-lessicale	1	2	3	3

Tab. 1 | Sistema di tipizzazione del materiale lessicale di *VerbaAlpina*

Tuttavia, il raggruppamento del materiale linguistico in tipi morfo-lessicali non è sufficiente all'adempimento completo degli obiettivi accennati *supra*. Per mostrare i punti di coesione tra i diversi idiomi alpini si rende necessario e interessante un altro tipo di tipizzazione, ovvero un raggruppamento del materiale linguistico in tipi di base. Un tipo di base è rappresentato dalla radice lessicale, comune alle diverse attestazioni, espressa attraverso la prima forma storicamente attestata. Ad esempio, il tipo di base latino *butyru(m)* 'burro' rappresenta la prima forma storicamente attestata di diverse parole dialettali (ma anche appartenenti alle lingue nazionali) di origine germanica, romanza e slava quali *b'yri* (La Javie, Francia), *but'e:r* (Sonico, Italia), *Buttr* (Füssen, Germania), *putar* (Jesenice, Slovenia)², tutte relative al concetto BURRO. Importante, ai fini della ricostruzione dei percorsi etimologici, è la precisazione della tipologia di relazione storica tra la forma di base e la parola dialettale (e/o il tipo morfo-lessicale) in questione: mentre per le forme romanze, in questo caso, la relazione con la forma base lat. *butyru(m)* è diretta, le forme germaniche e slave sono identificabili come prestiti.

I dati linguistici di *VerbaAlpina* appaiono su una mappa interattiva (cf. mappa interattiva 2022b)) che dispone di diverse modalità di filtrare i dati in prospettiva onomasiologica e semasiologica. Sulla mappa sono visualizzabili anche dati toponomastici ed extra-linguistici. Accanto a questa forma di visualizzazione dei dati, *VerbaAlpina* propone anche il *Lexicon Alpinum* (cf. Krefeld & Lücke 2014b) come strumento di lettura dei dati: in esso sono contenuti tutti i tipi morfo-lessicali, tipi di base e i concetti di *VerbaAlpina*, solitamente provvisti di un commento di tipo storico-etimologico oppure etnolinguistico.

2.2 Sfide generali

VerbaAlpina è un atlante linguistico pensato per e basato sul web. Ciò significa che, da un punto di vista tecnico-informatico, esso ha la necessità di confrontarsi con specifiche questioni che riguardano la sua integrità e la sua durabilità nel tempo. La pagina web, sviluppata sulla base del *content management system Wordpress*, viene «congelata» ogni sei mesi. Si vengono quindi a creare delle versioni stabili non più modificabili. L'attuale versione è la numero 21/2 (ossia la seconda versione

² Per la consultazione della cartina interattiva che raccoglie tutte le forme dialettali legate al tipo di base lat. *butyru(m)* cf. mappa interattiva 2022a.

creata nel 2021). Le modifiche vengono apportate sempre alla versione di lavoro che prende il nome «xxx». In termini cartacei, ogni versione può essere paragonata a una nuova edizione pubblicata di *VerbaAlpina*.³ Fondamentale per *VerbaAlpina* è, al contempo, lo sviluppo di strategie per garantire la durabilità dei dati così come il loro accesso anche dopo la fine del progetto. Ancor prima che i principi FAIR fossero formulati (cf. Wilkinson & Dumontier et al. 2016), *VerbaAlpina* ha lavorato, anche in collaborazione con enti esterni, affinché il progetto fosse reso *findable* (progetti *Generic Research Data Infrastructure - GeRDI*, *eHumanities - interdisziplinär e Discover*⁴), *accessible* (rinunciando ai diritti d'autore, applicando una licenza CC-BY SA 4.0. e mettendo a disposizione un'API per l'accesso computerizzato ai dati; cf. Krefeld & Lücke 2014-d), *interoperable* (strutturando minuziosamente i dati, applicando metadati, identificatori e controllo di autorità) e di conseguenza *reusable*.⁵

3. Il crowdsourcing: dare spazio alla collettività nella ricerca (geolinguistica)

L'immagine tradizionale che si ha del ciclo della scienza (nel corso di questo scritto si farà riferimento specialmente alla geolinguistica) segue normalmente lo schema seguente (estremamente semplificato): un ricercatore (o un gruppo di ricerca) elabora un progetto, si reca presso gli informanti (o i probandi), raccoglie i dati di cui necessita, ritorna presso il suo laboratorio, prosegue con l'analisi di quanto raccolto e dà comunicazione dei risultati emersi alla comunità scientifica. In questa immagine, il ricercatore viene visto come la guida, la persona che decide cosa estrapolare, ma soprattutto, quando estrapolare dati dall'informante. Sempre estremizzando, si tratta dunque di un sistema in cui l'idea che si veicola è di un ricercatore detentore del sapere da un lato e di un informante fonte di informazione, ma privo di pensiero critico, dall'altro lato. Questo tipo di approccio, pur essendo largamente utilizzato oltre che molto tradizionale, non sempre è adatto a ricerche nelle quali vi sia l'esigenza di disporre di un cospicuo *dataset*. Nel momento in cui si rende necessaria la raccolta di un grande numero di dati, oppure in casi in cui i progetti di ricerca siano molto corposi, può essere proficuo un altro tipo di approccio, basato sull'apertura al grande pubblico. Non si tratta più solamente di «sfruttare» le competenze dei parlanti, ma di farli partecipare attivamente alla ricerca, di coinvolgerli attraverso le loro competenze, fornendo loro la possibilità di prestarsi a un qualcosa di socialmente utile. Si tratta, in un certo senso, della democratizzazione dell'attività di fare scienza. Accanto al ricercatore, cittadini non necessariamente formati in uno specifico ambito (ad esempio non formati in linguistica) diventano protagonisti partecipando al successo del progetto di ricerca.

³ Per una panoramica delle maggiori novità introdotte per ogni versione cf. Krefeld & Lücke 2014-c.

⁴ Cf. Tochtermann 2018-, Universitätsbibliothek Erlangen-Nürnberg 2018-, University Library LMU Munich 2021-2022.

⁵ Per un approfondimento sull'applicazione dei principi FAIR a *VerbaAlpina* si rimanda a Lücke 2020, Lücke 2021, Colcuc & Mutter 2020.

Le tecnologie di cui disponiamo rendono possibile e permettono un rapido sviluppo di questa maniera di fare scienza che, oggigiorno, è conosciuto sotto il nome di *crowdsourcing*. Martin, Lessman & Voß (2008, 1256) propongono la seguente definizione:

Crowdsourcing ist eine interaktive Form der Leistungserbringung, die kollaborativ oder wettbewerbsorientiert organisiert ist und eine große Anzahl extrinsisch oder intrinsisch motivierter Akteure unterschiedlichen Wissensstands unter Verwendung moderner IuK-Systeme auf Basis des Web 2.0 einbezieht.⁶

La parola *crowdsourcing* è formata dalla sincretisi delle due parole inglesi *crowd* 'folla' e *outsourcing* 'esternalizzazione', ovvero la pratica utilizzata da sempre più aziende di ricorrere a imprese o enti esterni per lo svolgimento di alcune parti dei propri processi di produzione. Dall'economia, dunque, il pensiero di esternalizzazione di alcune pratiche, si è esteso anche alla scienza, anche se, nel contesto accademico, sarebbe improprio pensare al *crowdsourcing* come a un fenomeno recente. Ovviamente le tecnologie attuali permettono, come già accennato, di praticare *crowdsourcing* in maniera più immediata, ma esempi di tale approccio sono documentati già nella storia. Si pensi ad esempio all'*Oxford English Dictionary*, la cui realizzazione poté contare sulla partecipazione attiva di volontari chiamati a catalogare il materiale e a fornire esempi concreti per l'uso delle parole (cf. Gilliver 2020; Ogilvie 2018) oppure all'enciclopedia libera *Wikipedia* (cf. Aa.Vv. 2001-).

Le inchieste linguistiche sono normalmente caratterizzate da una raccolta dei dati di tipo diretto oppure indiretto. Le prime si caratterizzano per il fatto che il ricercatore, solitamente munito di taccuino e/o di un registratore, si reca egli stesso nei punti di rilevamento da lui scelti al fine di raccogliere i dati ai quali è interessato. Questa modalità di raccolta dei dati è tipica della tradizione di ricerca romanza: si pensi, ad esempio allo *Sprach- und Sachatlas Italiens und der Südschweiz* (AIS) di Karl Jaberg e Jakob Jud (1928-1940). La tradizione di ricerca germanistica ha invece prediletto le inchieste di tipo indiretto mediante questionari solitamente inviati ai parlanti per posta.⁷ Tutto sommato dunque, adoperare il *crowdsourcing* come modalità di raccolta dei dati in geolinguistica significa mettere in campo una ricerca di tipo indiretto, ma sfruttando le potenzialità di internet.

I metodi di *crowdsourcing* si inseriscono nel più ampio contesto della cosiddetta *citizen science*, ossia la scienza fatta dai cittadini, nell'idea di facilitare e incoraggiare la partecipazione del grande pubblico ai processi di ricerca e di mettere a disposizione della comunità i risultati emersi (cf. ECSA 2013-).

I vantaggi dell'utilizzo di piattaforme e metodi di *crowdsourcing* sono numerosi: potendo contare teoricamente su più persone che si impegnano in un determinato progetto, è più verosimile che gli obiettivi siano raggiunti più rapidamente o che i

⁶ Traduzione italiana: Il *crowdsourcing* rappresenta una forma interattiva di fornitura di servizi, organizzata in modo collaborativo o competitivo, coinvolgente un gran numero di attori motivati estrinsecamente o intrinsecamente, che dispongono di conoscenze diverse. Essa è resa possibile tramite l'utilizzo dei moderni sistemi d'informazione e di comunicazione, sulla base del web 2.0 (traduzione Beatrice Colcuc).

⁷ Per un approfondimento sulle differenze tra raccolta dei dati diretta e indiretta si rimanda a Kunzmann & Mutter 2021.

problemi siano risolti con maggiore velocità e che, data la grande partecipazione e potendo contare su feedback diversificati, si arrivi in minor tempo a identificare eventuali criticità o incompletezze. Per quanto riguarda specificatamente i progetti di ricerca, l'utilizzo del *crowdsourcing* attraverso internet rende possibile una velocizzazione significativa dei tempi così come un risparmio di risorse economiche. Dal punto di vista dell'immaginario sociale, il *crowdsourcing* veicola un'idea più forte di collettività, facendo notare come sia solamente mediante la partecipazione e l'apporto di tutti che progetti, idee e ricerche possono essere completati. Un altro vantaggio è, in base al tipo di studio che si desidera condurre, l'anonimato: molto spesso le piattaforme di *crowdsourcing* funzionano attraverso la partecipazione anonima di persone e questo può andare a beneficio sia della qualità del dato, poiché gli informati potrebbero tendere a rivelare di più di quanto non avrebbero fatto in una situazione di faccia a faccia con il ricercatore, sia della quantità, poiché ci si potrebbe sentire più a proprio agio nel non fornire informazioni personali (si osservi la differenza quantitativa tra *crowder* registrati e non registrati su *VerbaAlpina* in 4.3.).

Senza dubbio, i metodi di *crowdsourcing* portano con sé anche alcuni svantaggi: la partecipazione di un pubblico (più o meno) indefinito di persone significa maggiore probabilità che sorgano domande, dubbi o difficoltà. È necessario dunque un accompagnamento costante da parte degli iniziatori al fine di evitare errori o incomprensioni. Oltre a ciò, il grande numero di persone fa sì che talvolta non si riesca ad avere il controllo totale delle procedure.

4. Crowdsourcing e VerbaAlpina

4.1 Contesto e piattaforme

Il panorama delle opere lessicografiche in area alpina è complesso e abbastanza dettagliato. Come accennato *supra*, il *dataset VerbaAlpina* è rappresentato da dati estrapolati da diversi atlanti linguistici, coprenti parti diverse del territorio alpino, pubblicati (per la maggior parte) su carta e riconducibili a diverse tradizioni di ricerca (germanistica, romanistica e slavistica). Di conseguenza, il materiale lessicale di cui si dispone è consistente. In un primo momento, tuttavia, questo insieme di dati si presenta abbastanza eterogeneo: gli atlanti linguistici della tradizione germanistica seguono infatti modelli e strutture diversi da quelli relativi a varietà linguistiche romanze o slave, non sempre contengono forme dialettali riguardanti gli stessi concetti oppure, ancora, sono stati pubblicati in momenti storici diversi. Oltre a ciò, gli atlanti linguistici e vocabolari dialettali tendono a racchiudere in essi forme che provengono da punti di rilevamento solitamente ben identificati dagli esploratori. Questo tipo di approccio, ormai tradizionale, comporta però un percettibile svantaggio: mentre le varietà linguistiche di alcune località, data la loro particolarità e/o la loro rilevanza scientifica, sono fatte spesso oggetto di studi, altre rischiano di essere tagliate fuori e di non essere né documentate, né considerate in sede scientifica. Fermo restando che con i metodi tradizionali di ricerca sul campo sarebbe impossibile rilevare la completezza dei dati per tutte le micro-località dei territori di ricerca e che quindi una selezione ponderata dei punti

è d'uopo, è altrettanto noto che le nuove tecnologie possono offrire la soluzione a (tanti) limiti della ricerca tradizionale. Per superare i limiti qui illustrati, a partire dal 2017 (versione 16/2), *VerbaAlpina* ha lanciato una piattaforma di *crowdsourcing*, sviluppata all'interno dello stesso, al fine di raccogliere parole provenienti da parlanti di tutti i comuni della realtà alpina. La piattaforma (cf. *crowdsourcing* 2022) presenta ai parlanti una sola domanda: si desidera sapere qual è la forma dialettale utilizzata in loco per riferirsi a determinate entità della vita reale. L'immagine seguente mostra come si presenta la piattaforma di *crowdsourcing* di *VerbaAlpina*:



1 | Finestra principale della piattaforma di *crowdsourcing* di *VerbaAlpina* senza risposte

Al centro della finestra si può identificare l'arco alpino e, al suo interno, i dati *crowd* registrati per ogni località. Si nota immediatamente come vi siano delle regioni ad alta intensità di dati (Alpi centrali) e altre a bassa intensità (Alpi occidentali e nord-orientali). Nella parte bassa della schermata si può leggere la domanda posta al *crowder*: si chiede di cliccare sul comune per il quale si desidera fornire delle parole, successivamente si chiede di scegliere un concetto dalla lista proposta. Nella prospettiva di un *crowder*, la piattaforma presenta uno stimolo testuale in lingua standard (per i linguisti il concetto) che deve essere «tradotto» nel dialetto del comune prescelto. Una volta inviata una parola, si potrà proseguire con altre, sempre seguendo lo stesso procedimento.



2 | Finestra principale della piattaforma di crowdsourcing di VerbaAlpina con risposta

Prima di arrivare alla finestra principale, vi sono due passi precedenti da eseguire: la scelta della lingua di navigazione (francese, italiano, tedesco, sloveno) e la scelta del proprio dialetto tra quelli proposti, oppure, in alternativa, il suo inserimento nella lista. La piattaforma offre inoltre la possibilità di registrarsi oppure di inviare i dati in forma anonima (in alto a destra), di indicare la propria età e dispone di una sezione di domande frequenti così come di una classifica dei *crowder* più attivi, dei comuni più rappresentati e dei concetti preferiti (nella parte sinistra della schermata). L'interesse primario di VerbaAlpina si rivolge alla distribuzione delle varianti nello spazio geografico senza però focalizzarsi sull'opposizione tra "lingua" e "dialetto". Il materiale raccolto attraverso il *crowdsourcing* non esclude dunque varianti regionali delle lingue nazionali che si collocano in posizione intermedia tra i dialetti e queste ultime. In quest'ottica, categorizzare le parole raccolte come "dialettali" oppure come scaturite dalle lingue nazionali risulta superfluo. Nondimeno, si è rilevata una prossimità lievemente più accentuata alla lingua nazionale per quanto riguarda le parole relative ai concetti della vita moderna rispetto a quelle relative ai domini concettuali dell'alpicoltura e della natura alpine.

4.2. Finalità del crowdsourcing

Gli obiettivi perseguiti da *VerbaAlpina* sono molteplici: si desidera innanzitutto raccogliere dati attuali per poter operare un confronto con i dati presenti negli atlanti linguistici e nei dizionari cartacei. La raccolta di nuovo materiale è utile anche per colmare eventuali lacune presenti nelle opere lessicografiche (nel caso in cui, ad esempio, un atlante non abbia rilevato la forma lessicale relativa a un concetto preciso per un determinato comune). Essendo il comune l'unità territoriale di rilevamento più piccola in *VerbaAlpina*, ognuno di essi rappresenta un punto di rilevamento aumentando dunque la copertura territoriale.

Da ultimo, non meno importante, anzi, fondamentale è il quarto obiettivo: *VerbaAlpina* utilizza la piattaforma di *crowdsourcing* al fine di raccogliere i dati lessicali per la terza fase del progetto. Mentre per le prime due fasi (alpeggio e

natura alpina) la base di dati più consistente era rappresentata da opere lessicografiche già pubblicate, per l'attuale terza fase, incentrata sulla vita moderna nelle Alpi, in particolare sui domini concettuali dell'ecologia e del turismo, non sono ancora state documentate forme lessicali dialettali in maniera massiccia in alcuna opera lessicografica. Per questo motivo, la partecipazione dei parlanti dei comuni alpini è fondamentale.

La piattaforma di *crowdsourcing* non è soltanto utile alla raccolta di dati meramente lessicali, ma anche sociolinguistici e relativi alle rappresentazioni dei parlanti. In particolare, l'analisi delle indicazioni del dialetto (nelle immagini soprastanti si nota la dicitura dialettale *veneto* in basso a sinistra) può dare un contributo allo studio delle rappresentazioni relative al nome della propria parlata.

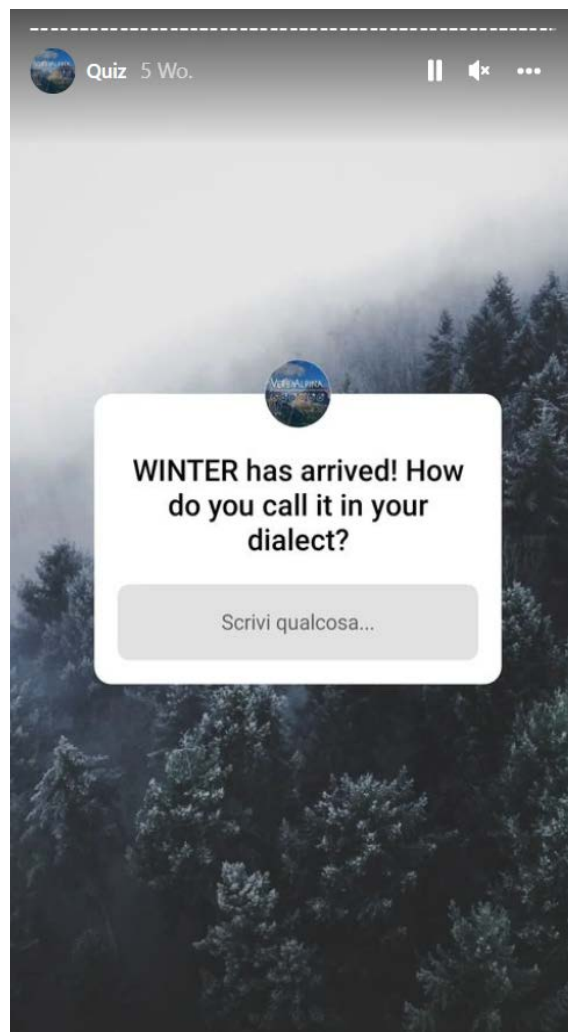
VerbaAlpina non pratica il *crowdsourcing* solo attraverso una piattaforma sviluppata in seno al progetto, ma ha trovato spazio anche sul portale di *citizen science Zooniverse* (cf. Zooniverse 2009-). Sono state proposte due attività: il *main task* chiedeva di corredare un insieme di 1212 stimoli di cartine dell' AIS con il proprio numero, mentre l'*optional task* prevedeva che i volontari della rete si impegnassero nella trascrizione del materiale degli atlanti utilizzando le regole di trascrizione di *VerbaAlpina*. Il progetto di *crowdsourcing* di *VerbaAlpina* su *Zooniverse* ha visto la partecipazione di 753 volontari in totale. Nonostante il successo generale dell'operazione, *VerbaAlpina* ha riscontrato alcune difficoltà nell'inserimento e nel lancio del progetto su *Zooniverse*. La creazione di un apposito profilo è un'operazione che richiede molta cura e molto tempo, il procedimento di revisione da parte di *Zooniverse* si è anch'esso prolungato nel tempo. Dalla prima richiesta di partecipazione al lancio effettivo di *VerbaAlpina* sono passati circa due anni e, una volta lanciato il progetto sulla piattaforma, sono occorse diverse difficoltà da parte dei *crowder* che hanno necessitato di un accompagnamento da parte di *VerbaAlpina*. Si è trattato soprattutto di domande relative alla trascrizione di specifiche attestazioni. L' AIS presenta infatti talune imprecisioni relative all'applicazione delle convenzioni di trascrizione stabilite dagli autori stessi. Altri problemi si sono presentati laddove la qualità dell'immagine delle cartine proposte non permettesse una chiara lettura. Oltre a ciò, *VerbaAlpina* ha ricevuto domande scaturite dal semplice fatto che alcuni utenti non avevano letto completamente e correttamente le regole da noi formulate.

4.3 Strategie e risonanza generale

A distanza di quattro anni dal lancio della piattaforma interna di *crowdsourcing*, sono possibili e necessarie alcune considerazioni sui successi ottenuti e sulle difficoltà riscontrate. Innanzitutto, è necessario sottolineare come la piattaforma, per portare risultati, abbia bisogno di un'assistenza continua. Naturalmente, affinché la partecipazione da parte dei parlanti sia costante nel tempo, vi è la necessità di far conoscere la piattaforma e questo implica il bisogno di pensare continuamente a modalità e strategie nuove e diversificate per pubblicizzare il progetto e la sua raccolta dati. L'attività di partecipazione da parte dei *crowder* è monitorata attraverso una statistica live (cf. statistiche live 2017-). Per ogni giorno di calendario, vengono segnalate quante parole entrano a far parte del *dataset* di

VerbaAlpina, da quali Paesi e da quanti *crowder*. L'altezza delle colonne è proporzionale al numero di parole inserite al giorno. Osservando i dati in un periodo di tempo più ampio, si nota come, ad ogni specifica attività da parte del progetto, corrisponda una partecipazione maggiore al *crowdsourcing*.

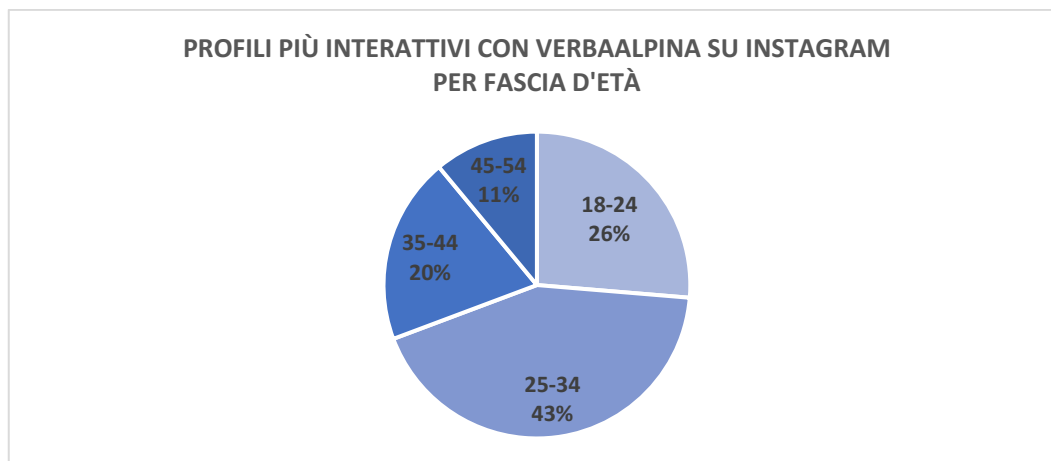
L'attività di *VerbaAlpina* per pubblicizzare il progetto e la sua piattaforma è cambiata nel corso degli anni, anche alla luce delle esperienze via via raccolte. Per la prima e la seconda fase ci si è affidati soprattutto a modalità di pubblicità stampate, specialmente articoli di giornale o di riviste e, parallelamente, sono stati creati dei flyer pubblicitari da distribuire all'interno dell'area di ricerca. *VerbaAlpina* ha inoltre preso parte anche a convegni popolari e ha rilasciato interviste radiofoniche (cf. *eco mediatica*). Contemporaneamente, si è proceduto a contattare, soprattutto via e-mail, enti e associazioni dell'arco alpino che si occupassero di lavorazione del latte e/o di natura con la richiesta di partecipare al *crowdsourcing* e/o di far conoscere il progetto nella propria cerchia di contatti. La terza fase del progetto ha visto un cambio di rotta per quanto riguarda le modalità di promozione del progetto: ai profili social di *VerbaAlpina* su *Facebook* e *Twitter*, aperti nel 2016, si è aggiunto nel 2020 un profilo *Instagram* (cf. *VerbaAlpina* su *Facebook* 2016-, *VerbaAlpina* su *Twitter* 2016-, *VerbaAlpina* su *Instagram* 2020-). La strategia per la diffusione delle informazioni riguardo a *VerbaAlpina* in questa fase ha visto (e vede tutt'oggi) protagonista internet, in particolare i *social network*. Le attività sui social sono di natura diversa: da un lato si ricercano attivamente contatti, specialmente pagine che condividono, almeno in parte, gli stessi interessi di *VerbaAlpina*; dall'altro lato si propone un ventaglio di contenuti divulgativi inerenti a *VerbaAlpina*, al mondo delle Alpi e ai temi cari al progetto. Per esempio, a luglio 2021 è stata ideata la rubrica "False myths to fight" 'falsi miti da sfatare' che si occupa di smontare idee false e luoghi comuni sui dialetti nell'ottica di fare chiarezza e di avvicinare i parlanti al mondo della dialettologia attraverso la sensibilizzazione. Su *Instagram* vengono proposte le cosiddette stories contenenti domande rivolte al pubblico. Una buona quantità dei dati raccolti a partire da maggio 2021 proviene proprio dalle risposte dei parlanti alle stories in cui viene chiesto loro, ad esempio, qual è la parola per INVERNO nel loro dialetto.



3 | Una story di *VerbaAlpina* su *Instagram*

Al fine di costruire una rete di contatti sempre più stabile e duratura, *VerbaAlpina* ha anche organizzato e partecipato a dirette su *Facebook* assieme ad altri progetti che si occupano di sensibilizzazione linguistica nelle Alpi.

Il cambio di strategia applicato con l'inizio della terza fase, ossia la (più) costante attività sui *social network*, ha evidenziato un maggiore interesse da parte di parlanti che hanno deciso di partecipare attivamente alla raccolta dati attraverso la piattaforma di *crowdsourcing*. A questo proposito è interessante soffermarsi brevemente sulla tipologia di pubblico che viene raggiunta grazie alle attività social di *VerbaAlpina*. Il grafico seguente mostra i profili che interagiscono maggiormente con la pagina di *VerbaAlpina* su *Instagram*, divisi per fasce d'età:

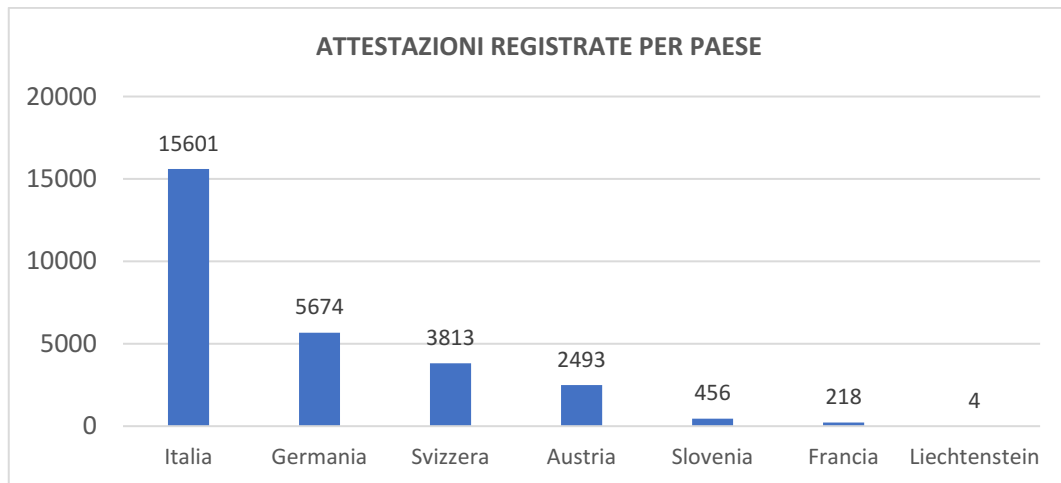


4 | Profili che interagiscono maggiormente con VerbaAlpina su Instagram per fasce d'età

Come accennato *supra*, la piattaforma di *crowdsourcing* concede ai *crowder* la libertà di creare un profilo e di dichiarare la propria età. Non essendo questo un dato che i *crowder* devono per forza fornire, *VerbaAlpina* non dispone di un numero netto relativo all'età media delle persone che partecipano alla piattaforma. Inoltre, i dati scaturiti dalle statistiche di *Instagram* sull'interattività con il profilo di *VerbaAlpina* non garantiscono obbligatoriamente una maggiore attività sulla piattaforma di *crowdsourcing*, ma suggeriscono prudentemente che gli informanti di *VerbaAlpina* potrebbero essere probabilmente più giovani rispetto agli informanti dei tradizionali atlanti linguistici.

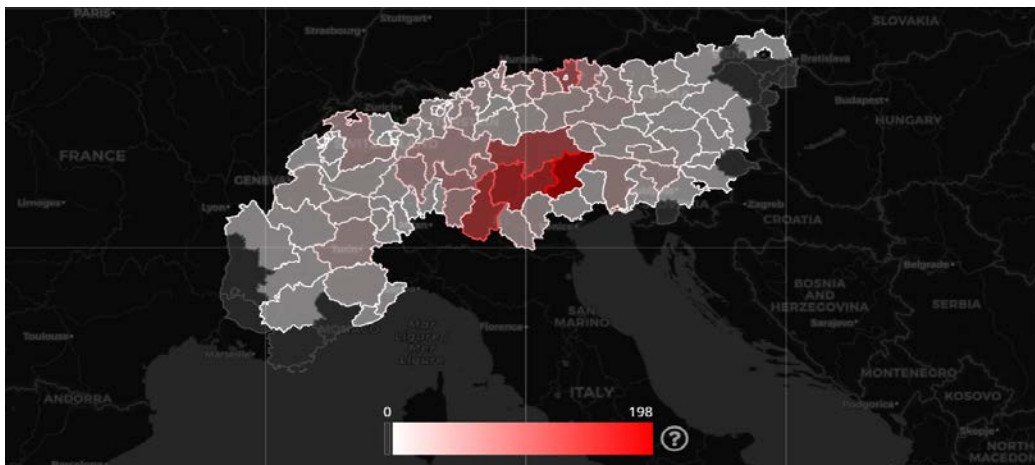
Dall'inizio dell'attività di *crowdsourcing* (dati aggiornati al 26 gennaio 2022), *VerbaAlpina* conta un totale di 1916 informanti, di cui 1668 non registrati e 248 registrati e ha raccolto 28.259 parole. Di queste, 15.601 provengono dall'Italia, 5674 dalla Germania, 3813 dalla Svizzera, 2493 dall'Austria, 456 dalla Slovenia, 218 dalla Francia e 4 dal Liechtenstein.⁸ Le parole raccolte attraverso il *crowdsourcing* hanno portato alla creazione di un totale di 4241 tipi morfo-lessicali nuovi (non presenti in atlanti linguistici), di cui 1794 per la prima fase (340 romanzi, 1384 germanici e 106 slavi), 1833 per la seconda fase (815 romanzi, 1001 germanici e 17 slavi) e 704 per la terza fase (611 romanzi, 93 germanici). Globalmente, i dati raccolti sono inoltre altamente qualitativi: su circa 40.000 tokens totali di cui si dispone (dati aggiornati a ottobre 2022), il numero di dati non validi è decisamente irrilevante (si tratta ad esempio di espressioni come "non so" oppure "non ricordo" inviate al posto della parola dialettale per un dato concetto).

⁸ Per consultare il numero di attestazioni registrate in costante aggiornamento cf. statistiche live.



5 | Numero di attestazioni registrate per Paese (dati aggiornati al 26 gennaio 2022)

La cartina seguente mostra invece le regioni NUTS⁹ dell'arco alpino che concentrano il maggior numero di *crowder* di *VerbaAlpina*: si tratta delle province italiane di Belluno, Trento, Brescia e Bolzano. A settentrione si distingue inoltre il Landkreis Rosenheim (per la versione interattiva della stessa cartina cf. mappa interattiva 2022c).



6 | Provenienza dei crowder di VerbaAlpina, secondo regioni NUTS (dati aggiornati al 26 gennaio 2022)

Per motivi di assetto interno del progetto, l'impegno per la promozione attraverso i *social network* si è inizialmente concentrato sull'Italia e, di contro, tale attività rafforzata ha incontrato fin da subito un forte interesse da parte del pubblico italiano.

⁹ Nomenclatura delle unità territoriali statistiche, si tratta di un sistema di suddivisione geografica del territorio dell'Unione europea a fini statistici.

4.4 Crowder

Si è accennato alla particolarità della terza fase di *VerbaAlpina*, caratterizzata dalla mancanza di documentazione lessicale in atlanti linguistici o dizionari dialettali per i domini concettuali della vita moderna alpina (ecologia e turismo), e dal fatto che, conseguentemente, *VerbaAlpina* ha deciso di impiegare la piattaforma di *crowdsourcing* come strumento principale della raccolta dati. Dopo tre anni di utilizzo della piattaforma, si è resa necessaria e significativa una breve indagine riguardante la percezione della piattaforma da parte dei *crowder*, il suo utilizzo e il rapporto tra *VerbaAlpina* e i parlanti al fine di poter considerare successi, criticità ed eventualmente apportare delle migliorie.

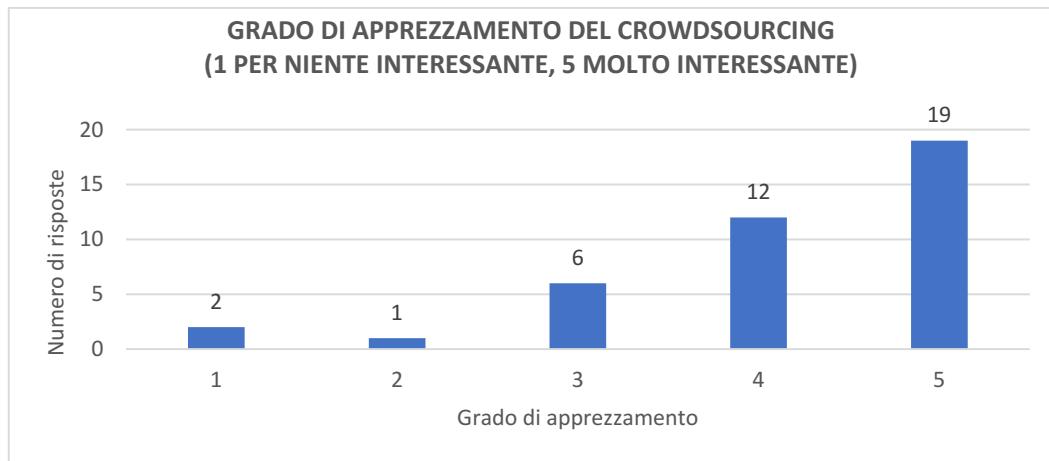
A questo scopo è stato creato un questionario online con diverse domande, suddivise in tre sezioni: la prima parte indaga il rapporto del *crowder* con *VerbaAlpina*, la seconda parte si concentra sulla valutazione dell'importanza attribuita dal *crowder* alla partecipazione ad attività collettive anche presso la propria comunità e la terza parte concerne il rapporto del *crowder* con il proprio dialetto. Il questionario è stato redatto nelle quattro lingue di *VerbaAlpina* (italiano, francese, tedesco e sloveno) e inviato a tutti i *crowder* registrati (di cui quindi *VerbaAlpina* possiede un indirizzo e-mail).¹⁰ L'osservazione delle risposte è servita per avere una panoramica generale di ciò che è apprezzato e ciò che invece andrebbe migliorato sulla piattaforma nell'ottica dei suoi utilizzatori.

In sintesi, il partecipante medio (registrato) dell'attività di *crowdsourcing* si profila come segue: ha tra i 26 e i 50 anni, dà un valore molto alto alla propria varietà dialettale e si interessa attivamente a questioni culturali della sua zona. È dell'idea che alle varietà dialettali si debba concedere più spazio a livello sociale, ad esempio attraverso l'insegnamento a scuola e considera importante che le piccole realtà linguistiche vengano documentate attraverso i progetti di ricerca.

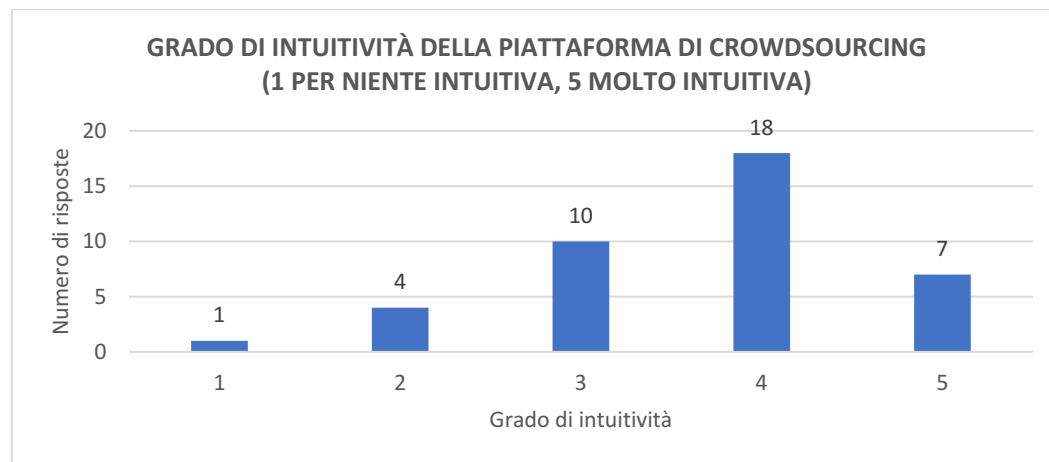
Ai fini della ricerca, per *VerbaAlpina* sono fondamentali tre aspetti dell'attività di *crowdsourcing*: in primo luogo, l'attività deve destare l'interesse dei parlanti e invogliarli a prendere parte al progetto; in secondo luogo, per stimolare le persone a partecipare o a continuare a partecipare, è necessario che la piattaforma sia rapida e di facile utilizzo. Inoltre, dal momento che l'idea generale del *crowdsourcing* è quella di creare una *community* che partecipi alle inchieste, è importante che i *crowder* che hanno apportato il loro contributo si sentano invogliati a invitare familiari, amici e conoscenti a prendere parte.

I dati cumulativi hanno mostrato che i riscontri relativi a tali aspetti sono, nel complesso, positivi. In linea generale, i *crowder* hanno trovato interessante la loro partecipazione e la piattaforma è risultata abbastanza intuitiva come è possibile notare dai due diagrammi sottostanti:

¹⁰ Per la versione italiana del questionario cf. Colcuc & Rodella 2021.

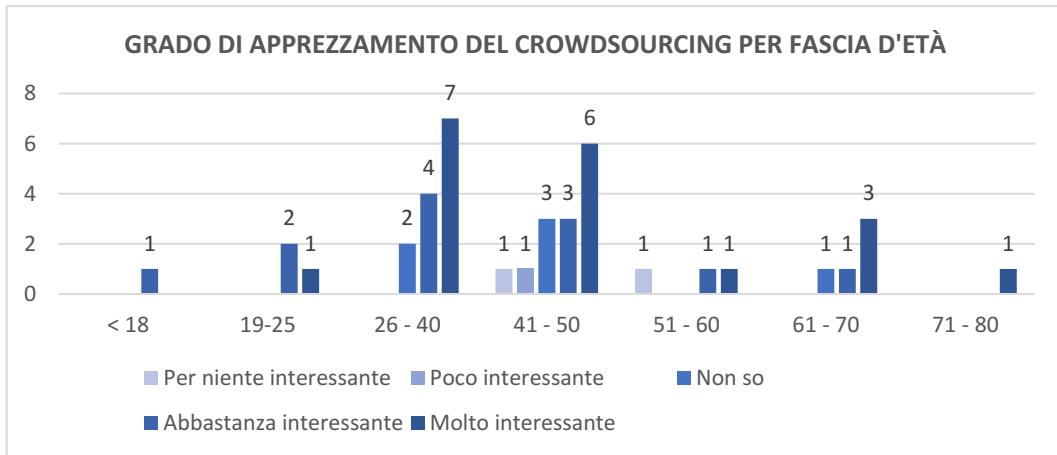


7 | Grado di apprezzamento della partecipazione al *crowdsourcing* di *VerbaAlpina*



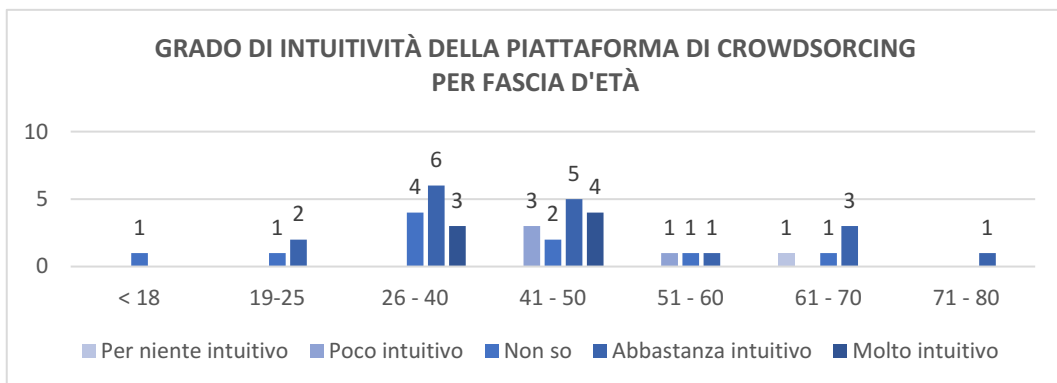
8 | Grado di intuitività della piattaforma di *VerbaAlpina*

I dati divengono più interessanti e incisivi nel momento in cui si vanno a incrociare con l'età e il titolo di studio dei partecipanti. In 4.3. si accennava alle fasce d'età che interagiscono maggiormente con la pagina *Instagram* di *VerbaAlpina* (fascia d'età tra i 25 e i 43 anni). Se, presi i dati del questionario, si nota come lo stesso gruppo dimostri di aver provato interesse nei confronti della partecipazione alla piattaforma. Su un totale di 13 partecipanti al questionario per la fascia 26-40, 11 hanno giudicato la loro partecipazione come abbastanza o molto interessante. Il dato è molto significativo, oltre che utile, per *VerbaAlpina*, poiché conferma che il progetto è riuscito a fare breccia in un gruppo di popolazione considerato generalmente distante dalle questioni legate ai dialetti e dal mondo della ricerca scientifica in dialettologia.



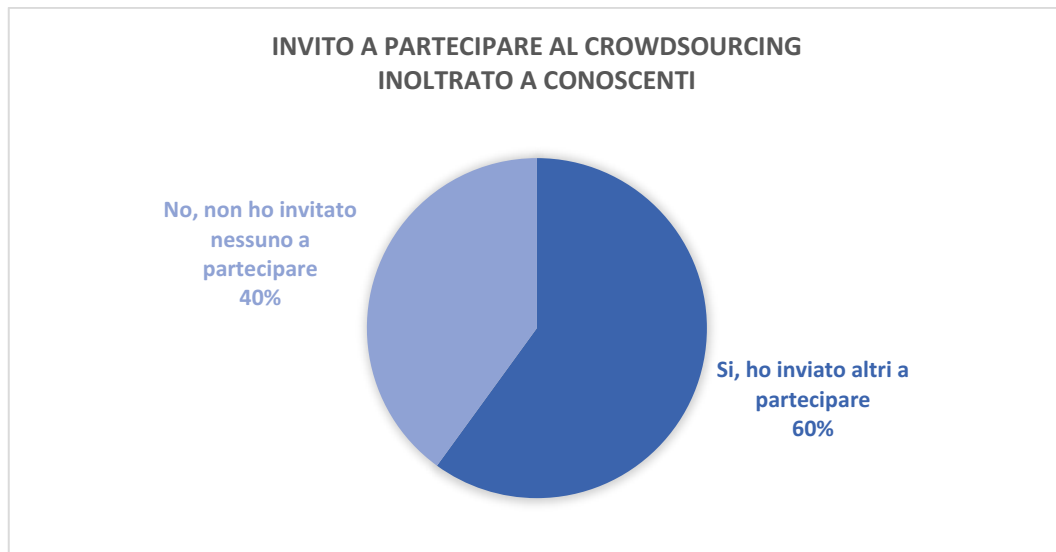
9 | Grado di apprezzamento della partecipazione al crowdsourcing per fascia d'età

Osservando i dati sul grado di intuitività della piattaforma in base all'età, si nota come tutte le fasce diano un giudizio sommariamente positivo. Ciononostante, idealmente, il livello di intuitività di una piattaforma che si pone come obiettivo la raccolta di dati lessicali, dovrebbe essere eccellente. Da considerare parallelamente a questo dato sono i *feedback* ricevuti da *VerbaAlpina* da alcuni utenti di *Facebook* a seguito di un'attività intensificata di ricerca di contatti: dopo aver pubblicato un breve testo di presentazione di *VerbaAlpina* insieme a un appello a partecipare, una decina di utenti, avendo riscontrato difficoltà di carattere tecnico, hanno contattato *VerbaAlpina* per ricevere delucidazioni.



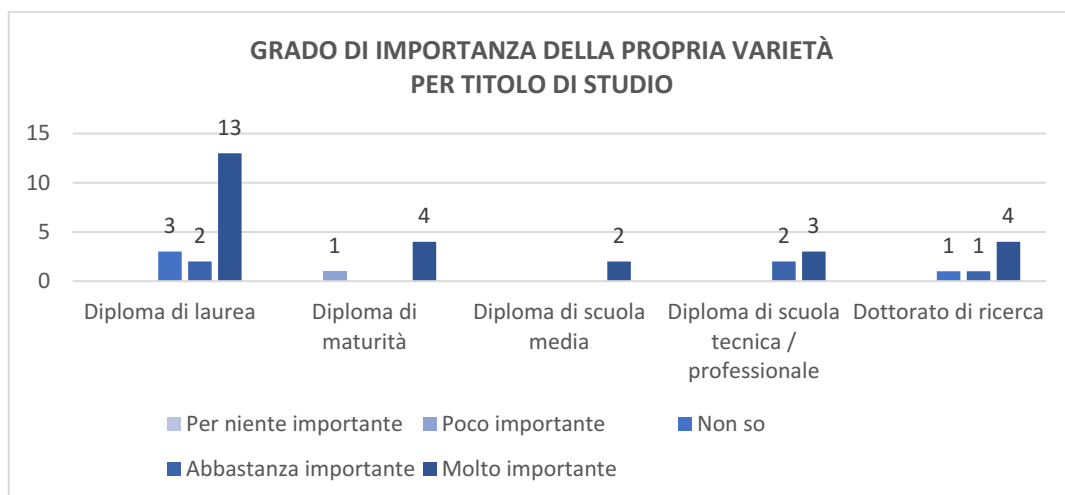
10 | Grado di intuitività della piattaforma di crowdsourcing di VerbaAlpina per fascia d'età

Per ciò che concerne la creazione di una community, il dato che traspare dalla breve inchiesta è che non tutti i *crowder*, dopo aver contribuito con le loro conoscenze, hanno invitato altre persone a partecipare.



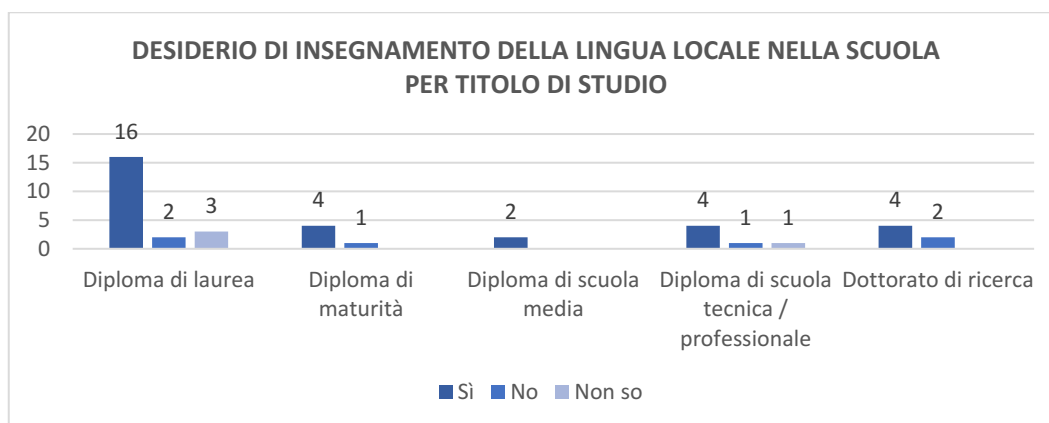
11 | Invito a partecipare al *crowdsourcing* inoltrato a conoscenti

Dopo aver mostrato e discusso i dati soprastanti, rimane un ultimo aspetto interessante da mettere in rilievo: il grado di istruzione. Si potrebbe (erroneamente) pensare che, in linea di massima, coloro che si sentono più legati e si interessano alle questioni a esso legate siano persone mediamente meno istruite. Le risposte al questionario hanno fornito una prospettiva diversa e interessante sulla questione. Su un totale di 40 partecipanti, 21 persone si sono dichiarate in possesso di un diploma di laurea. I dati relativi all'importanza concessa alla propria varietà, al desiderio di insegnamento della lingua locale nelle scuole e, più specificatamente, a *Verba Alpina*, il grado di apprezzamento della partecipazione al *crowdsourcing*, hanno mostrato come un percorso di istruzione completo non sia correlato a un disinteresse verso la propria lingua locale, anzi vi può essere un interesse e un attaccamento alla realtà dialettale molto forte. I diagrammi seguenti rendono i dati in maniera grafica:



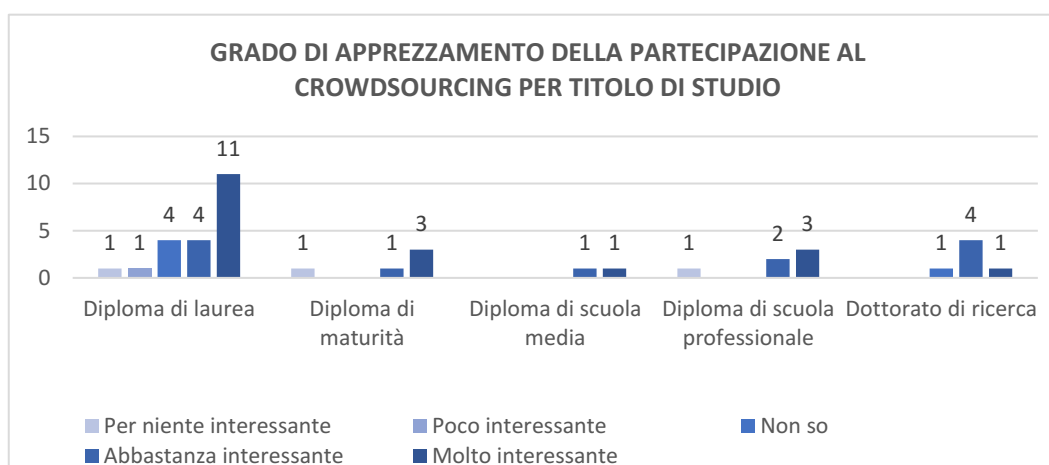
12 | Grado di importanza concesso alla propria varietà per titolo di studio

Si notino soprattutto le colonne relative alle persone in possesso di diploma di laurea e di dottorato di ricerca. Oltre a concedere importanza alla propria varietà, i parlanti sono dell'idea che si dovrebbe introdurre l'insegnamento della lingua locale nel percorso scolastico dei bambini.



13 | Desiderio di insegnamento della lingua locale nella scuola per titolo di studio

Per concludere, anche la stessa partecipazione alle inchieste di *Verba Alpina* ha riscosso successo presso i più istruiti. I giudizi dei possessori di laurea e di dottorato sono, come si vede, positivi.



14 | Grado di apprezzamento della partecipazione al crowdsourcing per titolo di studio

5. Considerazioni globali

Alla luce dei dati e delle informazioni apportate nel corso del presente scritto, alcune considerazioni riassuntive sono doverose. Innanzitutto, la decisione di sottoporre i *crowder* registrati a una breve inchiesta si è rivelata decisamente efficace, poiché è proprio sulla base dei dati raccolti che *Verba Alpina* ha potuto (e voluto) adoperarsi per apportare un miglioramento.

Si accennava all'inizio al fatto che la piattaforma di *crowdsourcing* non riesce a funzionare in maniera del tutto indipendente e che è necessario un continuo supporto e accompagnamento con attività di promozione da parte dei membri del

team di *VerbaAlpina*. Ciò che si può dire con certezza in questo frangente è che l'intuizione di concentrarsi sull'utilizzo dei *social media* come strumento di promozione è stata fondamentale. Oltre alla concreta partecipazione da parte di parlanti alpini attraverso l'invio di parole dialettali, *VerbaAlpina* ha riscontrato anche un crescente interesse da parte di testate giornalistiche e/o radiofoniche nei confronti dell'attività, segno che il progetto si sta interfacciando con sempre più persone in tutto il territorio alpino. Un altro vantaggio dell'utilizzo intensivo dei *social network* è, ai fini del progetto, il fatto che il pubblico giovane (e si è visto come vi siano anche molti giovani tra i *crowder* registrati), partecipando, apporta materiale attuale, aspetto centrale se l'obiettivo di *VerbaAlpina* è quello di fornire una prospettiva diacronica sui dati lessicali della regione alpina.

VerbaAlpina ha seriamente considerato i dati raccolti mediante il breve questionario e, dopo averli contestualizzati, si è proceduto ad elaborare delle strategie di comunicazione nuove che hanno riguardato soprattutto i *social network* e apportato delle migliorie di carattere tecnico che vanno nella direzione della *user-friendliness*. Per ciò che concerne le prime, mentre in un primo momento ci si era concentrati sul contattare pagine che avessero un legame con le Alpi con la preghiera di condividere un post o di chiamare i propri utenti a partecipare al *crowdsourcing*, ci si è attivati anche nei gruppi *Facebook* che raccolgono persone di determinate città o villaggi delle Alpi. Questo tipo di attività ha immediatamente e visibilmente aumentato la partecipazione e, al contempo, pubblicizzato il progetto direttamente presso i parlanti. Tuttavia, le novità principali lanciate a gennaio 2022 hanno riguardato l'integrazione di alcune nuove funzionalità della piattaforma di *crowdsourcing*. Tenuto conto dei *feedback* anonimi registrati relativamente al grado di intuitività mediante il questionario parallelamente ai commenti diretti ricevuti dagli utenti della rete (soprattutto su *Facebook*), è stato creato un breve video-tutorial che mostra il procedimento da seguire per inviare correttamente le parole, sono stati aggiunti alcuni effetti per rendere più visibili i passi da compiere (per esempio, prima si clicca su «comune» e poi su «concetto») e si è resa più visibile la scelta dei domini concettuali attraverso una breve spiegazione riguardante la possibilità di scelta.

5.1 Sfide

In primo luogo, *VerbaAlpina* deve cercare di attirare un maggior numero di parlanti, soprattutto provenienti dalle aree meno coinvolte, e convincerli a condividere il proprio sapere linguistico attraverso la piattaforma di *crowdsourcing* di *VerbaAlpina*. Per fare ciò, il progetto è costantemente impegnato a diffondere la propria presenza sul web attraverso l'uso attivo dei canali social e cercando di utilizzare tutte le lingue diffuse nell'arco alpino (con l'aggiunta dell'inglese).

Un fattore importante da tenere in considerazione è la «pigrizia» del parlante: si è notato che molti *crowder* sono disponibili a fornire parole, ma non sono sempre disposti a farlo da soli, non raramente per una semplice mancanza di dimestichezza con gli strumenti tecnologici e internet online, nonostante la piattaforma sia stata volutamente concepita in maniera semplice e risulti abbastanza immediata alla maggior parte degli utilizzatori. In questi casi è fondamentale operare una

mediazione, andando incontro alle esigenze di questi parlanti e colmando le loro lacune, per esempio rimanendo a disposizione tramite le e-mail o la messaggistica istantanea oppure inviando loro una lista di parole da tradurre nel proprio dialetto via e-mail che una persona del progetto andrà poi a inserire al loro posto.

Un'altra criticità riguarda la densità di parole registrate. Al momento, per *Verba Alpina* si verifica la compresenza di poche località ad alta densità di dati (province di Belluno, Brescia, province autonome di Trento e Bolzano e il *Landkreis Rosenheim*) e di moltissime aree a bassa densità. La situazione ideale, caratterizzata da un numero omogeneo di entrate in tutto l'arco alpino, è dunque lontana dall'essere raggiunta e l'aspetto sul quale è necessario incentrare gli sforzi è l'attività proprio nelle zone meno rappresentate (si veda il grafico in Fig. 5) cercando ad esempio di allacciare dei contatti in loco.

5.2 Opportunità

L'esperienza di *Verba Alpina* con il *crowdsourcing* ha portato alla luce le numerose opportunità che questo strumento è in grado di offrire.

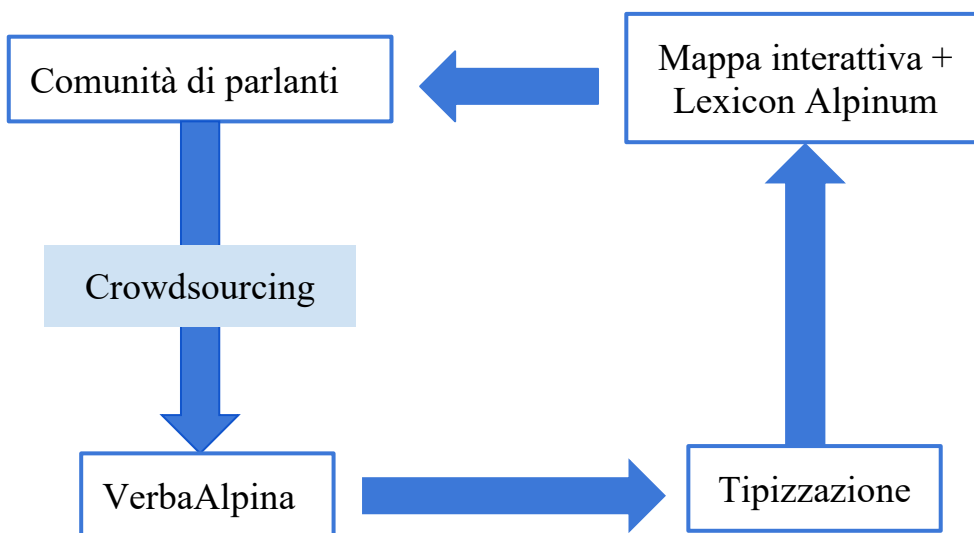
In primo luogo, l'apertura al grande pubblico per la raccolta di materiale lessicale, grazie a una piattaforma online, permette il superamento del limite di un informante per punto di rilevamento tipico della geolinguistica tradizionale. Attraverso il *crowdsourcing*, ogni punto di rilevamento, ossia ogni comune per *Verba Alpina*, può essere infatti rappresentato da più informanti.¹¹ Si tratta di un aspetto da non sottovalutare se si tiene conto dell'importanza della rappresentatività. In aggiunta, contrariamente a quanto accade(va) con gli atlanti linguistici tradizionali, dal momento che l'invio di dati lessicali è legato in primo luogo alla categoria territoriale del comune e non a una lingua o a una famiglia linguistica (almeno nella prospettiva del parlante), il *crowdsourcing* dà modo di registrare la realtà linguistica di un territorio dando conto della sua eventuale situazione di plurilinguismo. Riportiamo di seguito due esempi: per il comune di Livinallongo del Col di Lana (provincia di Belluno) si evidenziano due attestazioni, inviate da due utenti distinti, afferenti a due tipi morfo-lessicali diversi per il concetto BAITA (FABBRICATO, SEMPLICE, GESTITO SULL'ALPEGGIO), ossa *tablé* (tipo morfo-lessicale romanzo *tablé* di base lat. *tabulātum*) e *ciasota* (tipo morfo-lessicale romanzo *casotta* che si ricollega al tipo di base lat. *casa(m)*). Nel comune di Longarone (sempre in provincia di Belluno), si registrano due attestazioni per il concetto PANNA (STRATO DI GRASSO SOPRA IL LATTE CRUDO) che sono da un lato *cao* (tipo morfo-lessicale romanzo *capo*, con tipo di base lat. *caput*) e *brama* (tipo morfo-lessicale *brama*, con tipi di base prelat. *crama* e lat. *bruma*).¹²

Il *crowdsourcing* propone un metodo di ricerca creativo, che consente di indagare le potenzialità di mezzi non canonici, ma estremamente attuali come i *social media*,

¹¹ In termini teorici, non vi erano impedimenti che non permettessero la documentazione di forme raccolte da informanti diversi per punto di rilevamento, ma nella pratica, la piattaforma di *crowdsourcing* permette di eseguire questa operazione in maniera più semplice e rapida.

¹² Per visualizzare gli esempi sulla mappa interattiva si consulti il seguente link: <https://www.verbaalpina.gwi.uni-muenchen.de/it?page_id=27&noredirect=it_IT&db=221&tk=4533>.

per la raccolta di dati linguistici. Storie con quiz, rubriche, contenuti divulgativi, collaborazioni con altre pagine con obiettivi simili attraverso post o eventi sono tutti elementi potenzialmente fertili per la ricerca linguistica. Come accennato più volte, il *crowdsourcing* pone il parlante al centro, rendendolo (s)oggetto di studio privilegiato; inoltre, permette la raccolta di dati ecologici (ovvero non viziati dai vincoli e dalle sovrastrutture di un ambiente di laboratorio), reali e attuali, provenienti da parlanti spesso inconsapevoli della propria competenza dialettale e del proprio patrimonio linguistico locale. Attraverso il *crowdsourcing* l'accademia si avvicina alla comunità dei parlanti e si pone al suo servizio, creando uno scambio proficuo tra la conoscenza linguistica del parlante e il rigore del metodo di studio.



15 | Schematizzazione del rapporto tra *VerbaAlpina* e la comunità di parlanti¹³

5.3 Limitazioni

L'idea originale del *crowdsourcing* è l'integrazione di una comunità di interessati nei processi di elaborazione di un dato progetto. Nella teoria non si tratta quindi del mero invio di dati (nel caso di *VerbaAlpina*, dati linguistici), bensì la partecipazione potrebbe (o dovrebbe) estendersi anche ad altri ambiti come per esempio l'utilizzo della cartina interattiva per scopi personali, il salvataggio di cartine, la partecipazione attiva attraverso la promozione del progetto nella propria cerchia di conoscenze. Come è già stato evidenziato *supra*, la piattaforma di *crowdsourcing* di *VerbaAlpina*, pur funzionando egregiamente dal punto di vista tecnico, non è in grado di sostentarsi da sola e di garantire un flusso di dati costante se non vi è attività di accompagnamento continuo da parte del progetto stesso. Una difficoltà per la creazione di tale community è rappresentata dal fatto che *VerbaAlpina* si basa sulla partecipazione completamente volontaria di persone che decidono di impiegare il proprio tempo per aiutare la scienza. Nell'assetto sociale attuale, il tempo è divenuto una risorsa preziosa che tende a mancare e la maggior parte delle persone operano scelte di pragmaticità scegliendo di impiegare il loro

¹³ Oltre ad essere strumenti al servizio delle comunità di parlanti, la mappa interattiva e il *Lexicon Alpinum* sono a disposizione di scienziati e accademici per la propria ricerca.

tempo libero in attività di cui, nella loro prospettiva, riscontrano immediatamente un vantaggio a breve o a lungo termine. Per questo, nella fase iniziale della promozione della piattaforma di *crowdsourcing*, *VerbaAlpina* ha profuso un grande impegno nel contattare diverse istituzioni dell'arco alpino con la preghiera di fare da tramite presso i parlanti. Si tratta perlopiù di istituzioni ed enti legati territorialmente alle Alpi oppure che si occupano di salvaguardia del patrimonio culturale, linguistico o storico locale. Si dava per scontato che la comunicazione avrebbe potuto dare maggiori frutti se mediata da enti e istituzioni locali, nelle quali (si pensava) i parlanti si identificano, rispetto alla comunicazione diretta tra *VerbaAlpina* (all'epoca sconosciuto ai più) e i parlanti poiché essi avrebbero capito che la documentazione della propria varietà avrebbe giovato sia a livello locale (per il proprio villaggio), sia a livello alpino. Tuttavia si è compreso ben presto che questa idea sarebbe stata limitata da diversi ostacoli. Molto spesso non è stata recepita alcuna risposta oppure è stata riscontrata poca predisposizione alla collaborazione. Le ragioni di tale atteggiamento possono essere diverse. Probabilmente il primo motivo è da ricollegare, ancora una volta, alla mancanza di tempo e di risorse da parte degli impiegati delle istituzioni che sono state contattate. Ciò che si riscontra però, considerato il contesto generale e ampliando la riflessione, è la mancanza di visione comune nella creazione di una community dalla quale gruppi di ricerca, istituti di cultura ed enti del territorio alpino potrebbero approfittare. Mentre il rapporto tra *VerbaAlpina* e il parlante è abbastanza diretto, ciò che blocca la richiesta di *VerbaAlpina* di fungere da intercessore presso i parlanti potrebbe essere una sorta di esitazione da parte degli enti esterni nell'affidarsi a una piattaforma online per la raccolta dei dati. *VerbaAlpina* ha dunque deciso di operare un cambio di direzione e di concentrare il proprio impegno sulla comunicazione diretta tra il progetto e i parlanti.

5.4 Prospettive

Il più importante (e più complesso) obiettivo che *VerbaAlpina* si propone è quello di creare una comunità internazionale di parlanti che partecipi al *crowdsourcing* in maniera costante nel corso del tempo. Si tratta di una prospettiva ambiziosa che vuole essere raggiunta progressivamente portando avanti e consolidando l'attività sul web e sui *social network*. Al fine di raggiungere questo obiettivo, si ritiene fondamentale ricorrere alla pratica piuttosto tradizionale del classico passaparola tra i parlanti con il conseguente sviluppo di una community organica e, contemporaneamente, conferire una struttura a tale insieme. Per fare ciò, è necessario continuare a sviluppare idee su come i *crowder* più coinvolti e interessati al progetto possano diventare uno strumento utile per diffondere il progetto, diventando a tutti gli effetti degli ambasciatori di *VerbaAlpina*. Ad esempio, si potrebbe conferire loro uno status speciale di *ambassador* o *top crowder*, dando la possibilità di gestire una piccola sub-community territoriale. Un ulteriore mezzo di comunicazione con i *crowder* potrebbe essere un canale *Telegram* di *VerbaAlpina* (da pubblicizzare adeguatamente sui *social network*), in cui tutte le persone interessate al progetto possono sentirsi libere di condividere i propri pensieri riguardo ai temi cari a *VerbaAlpina*. Un canale *Telegram* potrebbe essere anche un modo per interagire in modo diretto con più utenti in contemporanea, magari

proponendo una parola del giorno e chiedendo a tutti gli iscritti al canale di tradurla nel proprio dialetto.

6. Porsi all'ascolto

Nel corso di questo scritto si è presentata l'esperienza di *VerbaAlpina* ponendo l'accento su criticità, opportunità e sfide della piattaforma di *crowdsourcing*. Il fulcro della questione gira attorno all'idea di un'apertura all'inclusione dei parlanti da parte del progetto di ricerca. Una riflessione finale che desideriamo portare riguarda il rapporto tra *VerbaAlpina* e la comunità dei suoi *crowder* attuali e di quelli che lo diverranno in futuro. Lasciare spazio a persone non-scienziate non può e non deve limitarsi a chiedere loro qualcosa. È fondamentale assecondare richieste e necessità, ma soprattutto rendere queste persone partecipi attraverso attività di divulgazione e di pubblicazione dei risultati, attraverso momenti di dialogo e con opere di sensibilizzazione rispetto ai temi della dialettologia e della cultura locale. Inoltre, è importante prestare ascolto ai parlanti, farli sentire protagonisti di un'attività transnazionale e parte di una comunità panalpina. Le realtà accademiche spesso faticano a comunicare con persone esterne all'ambito poiché interessi, modalità e termini non sempre coincidono, ma lo sforzo va compiuto al fine di propagare l'idea di una scienza più inclusiva e meno esclusiva, abbracciando proprio *in toto* l'idea di *citizen science*.

Bibliografia

- AA. VV. 2001-. «Wikipedia, l'enciclopedia libera e collaborativa.», versione italiana.
<<https://it.wikipedia.org/>> 21.1.2022.
- COLCUC, Beatrice & Anna Rodella. 2021. *Ci racconti chi sei? VerbaAlpina, il Crowdsourcing e la vita nelle Alpi*. Online: VerbaAlpina.
<https://www.verba-alpina.gwi.uni-muenchen.de/?fragebogen=questionario-verbaalpina_it.> 15.2.2022.
- COLCUC, Beatrice & Christina Mutter. 2020. «Interoperabilité des données géolinguistiques à l'exemple du projet VerbaAlpina.» In *Bien Dire et Bien Apprendre – Revue de Médiévistique*, 35, 131–146.
<<https://doi.org/10.5282/verba-alpina?urlappend=%2Fwp-content%2Fuploads%2Fbdba-35-colcuc-mutter.pdf%3Fdb%3D202>>.
- CROWDSOURCING. 2022. «Piattaforma di crowdsourcing di VerbaAlpina» *VerbaAlpina* 21/2.
<<https://www.verba-alpina.gwi.uni-muenchen.de/crowdsourcing>> 15.2.2022.
- ECO MEDIATICA. 2022. «Eco mediatica di VerbaAlpina» *VerbaAlpina* 21/2.
<https://www.verba-alpina.gwi.uni-muenchen.de/it/?page_id=175&db=212> 15.2.2022.
- ECSA = European Citizen Science Association (ed.). 2013-.
<<https://ecsa.citizen-science.net/>> 31.01.2022.
- GILLIVER, Peter. 2020. *The OED's latest exercise in crowdsourcing*. Oxford: Oxford University Press.
<<https://public.oed.com/blog/the-oeds-latest-exercise-in-crowdsourcing/>>.
- JABERG, Karl & Jakob, Jud. 1928-1940. *Sprach- und Sachatlas Italiens und der Südschweiz (AIS)*, 8 vol., Zofingen: Ringier.
- KREFELD, Thomas & Stephan Lücke (ed.). 2014a. *VerbaAlpina. Der alpine Kulturraum im Spiegel seiner Mehrsprachigkeit*. München.
<<https://dx.doi.org/10.5282/verba-alpina>>.

- KREFELD, Thomas & Stephan Lücke (ed.). 2014b. «Lexicon Alpinum.» *VerbaAlpina* 21/2.
 <https://www.verba-alpina.gwi.uni-muenchen.de/?page_id=12180&db=212>
 15.2.2022.
- KREFELD, Thomas & Stephan Lücke (ed.). 2014c. «Timeline.» *VerbaAlpina* 21/2.
 <https://www.verba-alpina.gwi.uni-muenchen.de/it/?page_id=547&db=212>
 15.2.2022.
- KREFELD, Thomas & Stephan Lücke (ed.). 2014d. «API Dokumentation.» *VerbaAlpina* 21/2.
 <https://www.verba-alpina.gwi.uni-muenchen.de/?page_id=8844&db=212> 26.1.2022.
- KREFELD, Thomas. 2019a. «Umanistica digitale.» *VerbaAlpina* 21/2.
 <https://doi.org/10.5282/verba-alpina?urlappend=%2Fit%3Fpage_id%3D21%26db%3D212%26letter%3DU%2314>.
- KREFELD, Thomas. 2019b. «Trascrizione.» *VerbaAlpina* 21/2.
 <https://doi.org/10.5282/verba-alpina?urlappend=%2Fit%3Fpage_id%3D21%26db%3D212%26letter%3DT%2357>.
- KUNZMANN, Markus & Christina Mutter. 2021. «Zur Verwertbarkeit indirekt erhobener Sprachdaten: Erfahrungen aus dem Forschungsprojekt VerbaAlpina.» In *Linguistik grenzenlos: Berge, Meer, Käse und Salamander 2.0*, Lücke, Stephan et al. (ed.), online, versione 1.
 <<http://www.kit.gwi.uni-muenchen.de/?p=74940&v=1>>.
- LÜCKE, Stephan. 2020. «Lizensierung.» *VerbaAlpina* 21/2.
 <https://doi.org/10.5282/verba-alpina?urlappend=%3Fpage_id%3D493%26db%3D212%26letter%3DL%2341>.
- LÜCKE, Stephan. 2021. «FAIR-Prinzipien.» *VerbaAlpina* 21/2.
 <https://doi.org/10.5282/verba-alpina?urlappend=%3Fpage_id%3D493%26db%3D212%26letter%3DF%23128>.
- Mappa interattiva. 2022a. «Mappa interattiva del tipo di base lat. *butyru(m)*.» *VerbaAlpina* 21/2.
 <https://www.verba-alpina.gwi.uni-muenchen.de?page_id=133&db=212&tk=4041>
 14.2.2022.
- Mappa interattiva. 2022b. «Mappa interattiva di VerbaAlpina.» *VerbaAlpina* 21/2.
 <https://www.verba-alpina.gwi.uni-muenchen.de/?page_id=133&db=212> 15.2.2022.
- Mappa interattiva. 2022c. «Mappa interattiva di VerbaAlpina sulla provenienza dei crowder.» *VerbaAlpina* 21/2.
 <https://www.verba-alpina.gwi.uni-muenchen.de/?page_id=133&db=xxx&tk=4047&layer=5> 15.2.2022.
- MARTIN, Nicole, Stefan Lessmann & Stefan Voß. 2008. «Crowdsourcing: Systematisierung praktischer Ausprägungen und verwandter Konzepte.» In *Multikonferenz Wirtschaftsinformatik*, Bichler, Martin (ed.), 1254 – 1263, Berlin: GITO-Verlag.
- OGILVIE, Sarah (ed.). 2018. «Crowdsourcing the OED.» *Dictionary Lab*.
 <<https://dictionarylab.web.ox.ac.uk/crowdsourcing-oed>> 31.01.2022.
- Statistiche live 2017-. «Statistiche live dell'attività di crowdsourcing.» *VerbaAlpina* 21/2.
 <https://www.verba-alpina.gwi.uni-muenchen.de/it/?page_id=523&noredirect=it_IT&db=212> 15.2.2022.
- TOCHTERMANN, Klaus (ed.). 2018-. *Generic Research Data Infrastructure - GeRDI*. Kiel, Hamburg.
 <<https://www.gerdi-project.eu/>> 26.1.2022.
- Universitätsbibliothek Erlangen-Nürnberg (ed.). 2018-. *eHumanities – interdisziplinär*.
 <<https://www.fdm-bayern.org/ehumanities-interdisziplinaer/>> 26.1.2022.
- University Library LMU Munich (ed.). 2021-2022. *Discover*.

- <<https://discover.ub.uni-muenchen.de/>> 26.1.2022.
VerbaAlpina su Facebook 2016-. *Profilo di VerbaAlpina su Facebook*.
<<https://www.facebook.com/verbaalpina/>> 15.2.2022.
VerbaAlpina su Twitter 2016-. *Profilo di VerbaAlpina su Twitter*.
<<https://twitter.com/verbaalpina/>> 15.2.2022.
VerbaAlpina su Instagram 2016-. *Profilo di VerbaAlpina su Instagram*.
<<https://www.instagram.com/verba.alpina/>> 15.2.2022.
WILKINSON, Mark et al. 2016. «The FAIR Guiding Principles for scientific data management and stewardship.» *Sci Data* 3 (160018).
<<https://doi.org/10.1038/sdata.2016.18>>.
ZOONIVERSE. 2009-. *Progetto VerbaAlpina su Zooniverse*.
<<https://www.zooniverse.org/projects/filip-hr/verbaalpina/>> 15.2.2022.

Riassunto

Il progetto *VerbaAlpina* della Ludwig-Maximilians-Universität di Monaco di Baviera sta elaborando la sua terza fase, incentrata sul lessico alpino relativo all'ecologia e al turismo in area alpina. Dal momento che gli atlanti linguistici dell'arco alpino non contengono tale lessico, si è proceduto alla raccolta del materiale lessicale autonomamente utilizzando un'apposita piattaforma. Il presente contributo desidera dare comunicazione del contesto di utilizzo, dei successi e delle limitazioni di tale piattaforma e, nell'ottica della condivisione delle esperienze, si rivolge in prima linea a ricercatori che intendono intraprendere lo stesso percorso.

Abstract

In its third stage, the project *VerbaAlpina* of Munich University investigates the vocabulary of ecology and tourism in the Alpine area. Since the existing linguistic atlases do not include this vocabulary, *VerbaAlpina* collects the lexical material using a special platform. This paper aims primarily at informing about the context of use and discussing achievements and limitations of this platform. It is addressed to researchers who wish to adopt the same data collection methods.