



Géolinguistique

18 | 2018
Varia

Atlanti linguistici e archivi vocali di lingue locali e minoritarie in Italia

Atlas linguistiques et archives vocales des langues locales et minoritaires en Italie

Linguistic Atlases and Vocal Archives of Local and Minority Languages in Italy

Antonio Romano, Matteo Rivoira, Federica Cugno, Giovanni Ronco, Valentina De Iacovo e Valentina Colonna



Edizione digitale

URL: <http://journals.openedition.org/geolinguistique/290>

DOI: 10.4000/geolinguistique.290

ISSN: 2650-8176

Editore

UGA Éditions/Université Grenoble Alpes

Edizione cartacea

ISBN: 978-2-37747-073-0

ISSN: 0761-9081

Notizia bibliografica digitale

Antonio Romano, Matteo Rivoira, Federica Cugno, Giovanni Ronco, Valentina De Iacovo e Valentina Colonna, « Atlanti linguistici e archivi vocali di lingue locali e minoritarie in Italia », *Géolinguistique* [Online], 18 | 2018, online dal 01 décembre 2018, consultato il 12 settembre 2019. URL : <http://journals.openedition.org/geolinguistique/290> ; DOI : 10.4000/geolinguistique.290

Questo documento è stato generato automaticamente il 12 settembre 2019.

Géolinguistique

Atlanti linguistici e archivi vocali di lingue locali e minoritarie in Italia¹

Atlas linguistiques et archives vocales des langues locales et minoritaires en Italie

Linguistic Atlases and Vocal Archives of Local and Minority Languages in Italy

Antonio Romano, Matteo Rivoira, Federica Cugno, Giovanni Ronco, Valentina De Iacovo e Valentina Colonna

1. Introduzione

- ¹ Per documentare il progresso nel campo della conservazione dei dati linguistici occorrerebbe risalire all'invenzione-scoperta della scrittura, che per prima ha fornito alla nostra specie la possibilità di 'registrare' il parlato, e poi trattare di diverse tappe del tema della «cattura del parlato» nel corso dei secoli. Si può, invece, saltare più concretamente alla fine dell'Ottocento, quando si definiscono alcuni fortunati metodi di notazione grafica che permettono di rappresentare più o meno fedelmente le produzioni linguistiche orali (sulla scorta di una nuova concezione dell'antico principio *Scripta manent*). Verrebbe da pensare, inevitabilmente, ai primi sistemi di trascrizione delineati e promossi da studiosi come Alexander John Ellis, Karl Richard Lepsius, Alexander Melville Bell o Henry Sweet, ma occorre trattare soprattutto di quelli diffusisi nel campo della ricerca accademica attraverso l'opera di sodalizi internazionali, come quello tra Paul Passy e Daniel Jones, tra gli altri, che ha rafforzato la diffusione dell' «alfabeto» della International Phonetic Association, al centro di un protocollo convenzionale collaudato poi nel corso del Novecento e applicato alla descrizione di specimina di lingue diverse².
- ² Da un certo punto di vista, possiamo associare alle prime tradizioni di raccolta e conservazione dei dati linguistici anche i grandi progetti atlantistici, primo tra tutti l'*Atlas linguistique de la France* (ALF) concepito e realizzato da Jules Gilliéron all'inizio del Novecento, al quale si ispireranno l'*Atlante Italo-Svizzero* (AIS)³ di Karl Jaberg e Jakob Jud, pubblicato tra il 1928 e il 1940, e il successivo *Atlante Linguistico Italiano* di Matteo Bartoli

e, più tardi, Benvenuto Terracini, a tutti gli effetti la più grande opera nazionale di raccolta di dati dialettali, tuttora in corso di pubblicazione (v. § 2 e. ss.).

- 3 Superando le difficoltà di trattamento di testi che traspongono in un formato grafico solo una parte — sia pure quella linguisticamente saliente — dei contenuti di un messaggio orale, come risultato di un'operazione di trascrizione condotta con rigore variabile a seconda del raccoglitore, la diffusione di apparati di registrazione ha innovato profondamente la ricerca documentaria sul parlato (consentendo di avvicinarsi vertiginosamente al *Verba manent* annunciato da vari autori, tra i quali Pistacchi, Aquilanti & Baldi, 2006 e Ventura & Aquilanti, 2007).
- 4 Le prime iniziative di costituzione di raccolte di materiali orali in archivi sonori risalgono essenzialmente all'alba del Novecento e si riconducono alle intuizioni di alcuni studiosi sulle possibilità di sfruttamento dei sistemi di conservazione e riproduzione dei suoni di recente diffusione a quell'epoca⁴. Successivamente, nel corso del secolo, i progressi delle telecomunicazioni e le nuove tecnologie AV hanno consentito di codificare e archiviare efficacemente i dati su vari supporti e suggerito la creazione di fonoteche e videoteche presso i più importanti enti nazionali⁵. Il trattamento informatico dei dati affermatosi negli ultimi decenni ha indotto però a rinnovare nuovamente il settore, permettendo la gestione di corpora di dimensioni significative che, se da un lato beneficiano di possibilità di accesso e riproduzione pubblici in tempi e modalità un tempo impensabili, dall'altro pongono problemi sia tecnici (come quelli sui protocolli e sulle procedure di accesso) sia etici e giuridici (come quelli sulla riservatezza dei dati vocali e sull'autorialità delle registrazioni)⁶.
- 5 Questi temi sono stati oggetti di interessanti riflessioni nell'allestimento pubblico delle prime raccolte di dati linguistici promosse in ambito accademico sulla base di registrazioni d'archivio (v. ad es. i progetti «Le soffitte della voce» di A. De Dominicis & P. Mattana e «Grammo-foni» di S. Calamai & P. M. Bertinetto)⁷ che hanno aperto la strada a progetti più modesti come quello che propone il Laboratorio di Fonetica Sperimentale «Arturo Genre» (LFSAG, v. § 7).

2. Il dato linguistico trascritto

- 6 L'avvio degli studi sulle lingue senza, o con scarsissima, tradizione scritta si può collocare con una certa approssimazione alla seconda metà dell'Ottocento, ed è la logica conseguenza del nuovo atteggiamento verso le lingue e le culture dei popoli maturato nella temperie culturale romantica. Con lo sviluppo di studi e teorie volte principalmente a spiegare l'evoluzione delle lingue nel corso del tempo, si definiscono nuovi paradigmi scientifici che mutuano dalle scienze biologiche e fisiche i modelli esplicativi, da qui l'idea della lingua come organismo in evoluzione, o il tentativo di formalizzare i mutamenti linguistici, principalmente quelli fonetici, come «leggi ineccepibili» tentato dai cosiddetti «Neogrammatici». Da questi approcci, dai quali se ne svilupperanno altri come vedremo, nasce la necessità di documentare in modo quanto più preciso possibile l'aspetto fonico delle lingue, tanto più per quelle per le quali mancava una tradizione scritta.
- 7 Si definiscono successivamente i modelli interpretativi della geolinguistica o geografia linguistica menzionati in apertura, in parte proprio come superamento di alcune aporie rese evidenti dall'approccio eccessivamente rigido dei Neogrammatici messo alla prova con i dati raccolti sul campo. Gli strumenti tecnologici allora disponibili, che rendevano

sostanzialmente impraticabile la registrazione su un supporto durevole il suono, e il contesto culturale che vede maturare questo tipo di studi (la cultura occidentale e il necessario e costante riferimento a lingue tradizionalmente scritte con una scrittura alfabetica), fanno sì che l'orientamento verso sistemi di trascrizione di tipo alfabetico costituisca l'opzione privilegiata e la più diffusa⁸. Sino all'affermazione del sistema internazionale IPA (che non ha in ogni caso del tutto sostituito altri sistemi), le diverse tradizioni di studio hanno adottato sistemi parzialmente differenti, pur se tutti basati fondamentalmente sul presupposto che la produzione sonora umana sia interpretabile come successione di suoni discretizzabili, ai quali è possibile far corrispondere un determinato segno. Nello stesso ambito romano le diverse scuole nazionali hanno elaborato sistemi differenti, parzialmente legati alle ortografie delle lingue nazionali di riferimento. In Italia, le più importanti riviste di studi linguistici e dialettali adottarono e promossero sistemi parzialmente differenti, così l'*Archivio Glottologico Italiano* (1873-) di Graziadio Isaia Ascoli sceglie il sistema noto come *Böhmer-Ascoli*, che fu adottato anche dagli autori dello *Sprach- und Sachatlas Italiens und der Südschweiz*, in italiano *Atlante linguistico ed etnografico dell'Italia e della Svizzera meridionale* (AIS), e così *L'Italia Dialettale* (1924-) di Clemente Merlo⁹.

3. L'Atlante Linguistico Italiano

- ⁸ L'*Atlante Linguistico Italiano* è un progetto ideato nella prima metà del secolo scorso e tuttora in fase di elaborazione presso l'omonimo istituto dell'Università degli Studi di Torino, facente capo al Dipartimento di Studi Umanistici. La sua impostazione discende direttamente, come si è detto, da quella concepita da Gilliéron per il suo *Atlas linguistique de la France*, non diversamente da tutte le altre imprese geolinguistiche avviate nella prima metà del Novecento¹⁰. Le sue linee programmatiche, messe a punto negli anni Venti del secolo scorso, prevedono dunque il metodo dell'inchiesta diretta sul terreno condotta da un raccoglitore unico, affidata allo studioso friulano Ugo Pellis, e l'uso di una trascrizione fonetica, nonché la pubblicazione integrale dei materiali. Tuttavia, in seguito alla morte di Pellis, avvenuta nel 1943 dopo lo svolgimento di 727 inchieste, per giungere al compimento della raccolta fu necessario coinvolgere altri raccoglitori — Raffaele Giacomelli, Corrado Grassi, Giorgio Piccitto, Giovanni Tropea, Temistocle Franceschi, Michele Melillo — che operarono tra il 1953 e il 1965¹¹. Secondo una linea metodologica già inaugurata dall'*Atlante Italo-Svizzero*, all'indagine linguistica si è accompagnata fin da subito quella etnografica, volta ad illustrare il rapporto tra «parole» e «cose», nella duplice veste di illustrazioni e fotografie da proporre agli informatori per ottenere l'esatta denominazione degli oggetti rappresentati e di nuova documentazione acquisita sotto forma di disegni e riproduzioni fotografiche quando nel corso delle inchieste si incontravano altre tipologie di oggetti. Dal 1995, con la pubblicazione del I volume dell'Opera, a cui sono seguiti, allo stato attuale, altri otto tomi dei circa venti previsti, i materiali linguistici sono consultabili nella loro veste cartografica, secondo l'obiettivo principe dell'impresa, ossia cogliere la vita del linguaggio nel suo farsi e nel suo muoversi nello spazio e, dunque, nella storia¹². Su ogni carta linguistica, in corrispondenza dei singoli punti di inchiesta, sono riprodotte in grafia fonetica le risposte integrali fornite dagli informatori riguardo a un determinato concetto.
- ⁹ La necessità di fissare nel modo più preciso possibile le risposte ottenute, l'impossibilità di riascoltare a piacere il suono delle parole (magari confrontandolo con altre già udite) e

l'estensione della ricerca portarono Gilliéron a scegliere di affidare le inchieste a un unico raccoglitore (non linguista) proprio con l'intento di ottenere dati il più possibile confrontabili tra di essi, optando per un metodo di trascrizione detto «impressionistico», vale a dire basato sulla percezione del raccoglitore in quel determinato momento e contesto, senza rielaborazioni successive. Ne risulta una trascrizione che, come mostrano gli stessi taccuini di inchiesta, esclude a priori qualsiasi forma di normalizzazione o interpretazione in una prospettiva che oggi definiremmo fonologica. È questo l'approccio seguito inizialmente anche da K. Jaberg e J. Jud con l' AIS, le cui inchieste furono affidate al solo Paul Scheuermeier (il coinvolgimento di Gerhard Rohlfs e di Max Leopold Wagner è legato all'espansione del progetto di partenza). Ed è lo stesso scelto da Bartoli, poi necessariamente abbandonato alla morte di Ugo Pellis.

- 10 Questi, e i suoi successori dopo di lui, si mossero avendo a disposizione un sistema di trascrizione *sostanzialmente aperto, che integra soluzioni proprie al modello Ascoli-Böhmer e altre desunte da quello usato da Merlo per L'Italia Dialettale*. L'udito molto acuto di Pellis e il lungo tirocinio lo portarono ad allargare l'inventario di segni, modificando in parte anche le soluzioni adottate in corso d'opera. Il risultato fu un archivio di dati evidentemente ricchissimo, ma anche assai complesso da gestire. Solo grazie a un pluriennale lavoro condotto da Arturo Genre (cfr. Genre, 1978) è stato possibile razionalizzare e armonizzare le differenti soluzioni via via susseguites: attualmente uno speciale programma informatico, che si basa su alcune migliaia di istruzioni permette di traslitterare il dato annotato sulle schedine compilate a mano e digitalizzato, in vista della stampa sulle carte linguistiche (cfr. Rivoira, 2016).
- 11 Nel piano dell'*Atlante Linguistico Italiano* sono state incluse 72 inchieste piemontesi, svolte interamente da Ugo Pellis in quattro periodi nell'arco di pochi anni, dal 1929 al 1942, e corrispondenti a 69 Punti d'inchiesta, a cui si aggiunge quello oltreconfine di La Brigue (P. 94), comune passato in parte alla Francia nel 1947¹³. La documentazione presentata riflette dunque la situazione linguistica della Regione della prima metà del secolo scorso e va pertanto letta in relazione al variegato quadro economico, sociale e culturale del tempo, in parte ricostruibile grazie ai profili delle singole località compilati dal raccoglitore e contenuti nei *Verballi delle Inchieste*. Una parte di questi dati è confluita nelle carte linguistiche a colori di tipo analitico del recente *Piccolo Atlante Linguistico del Piemonte* (PALP) (Cugno, Rivoira & Ronco, 2018), elaborate a partire da una selezione di alcune voci significative dell'Opera maggiore. Accompagnate da un breve commento linguistico, esse illustrano in forma semplificata e con l'eliminazione di informazioni ripetitive le aree accomunate dallo stesso fenomeno linguistico e mostrano le peculiarità linguistiche delle numerose varietà dialettali parlate nella Regione nonché delle lingue di minoranza riconosciute e tutelate dalla legge italiana a livello nazionale e regionale.

4. L'*Atlante Linguistico Etnografico del Piemonte Occidentale* (ALEPO)

- 12 Alla documentazione linguistica del Piemonte offerta dall'impresa di portata nazionale dell'*Atlante Linguistico Italiano* si è aggiunta in anni più recenti quella resa disponibile dall'*Atlante Linguistico Etnografico del Piemonte Occidentale* (ALEPO), focalizzata sulle parlate galloromanze (francoprovenzali e occitane) delle vallate alpine della Regione. Il progetto, promosso negli anni Ottanta da Sabina Canobbio e Tullio Telmon e tuttora in corso presso il Dipartimento di Studi Umanistici dell'Università degli Studi di Torino¹⁴ ha finora

condotto alla pubblicazione, a partire dal 2004 del modulo dedicato a *Il mondo vegetale*, articolato nelle sezioni I-III. *Funghi e licheni* (2004), I-I *Alberi e arbusti* (2005), I-II *Erbacee* (2007)¹⁵, seguito nel 2013 da quello intitolato III *Il mondo animale*, I *La Fauna*, II *Caccia e pesca*. I materiali linguistici presentati sono stati raccolti mediante inchieste sul campo svolte tra il 1980 e il 1989 in 42 località di parlata galloromanza con l'inclusione di alcuni punti galloitalici situati nelle aree vallive e pianeggianti adiacenti all'area studiata allo scopo di analizzare le dinamiche tra centri linguisticamente non affini. La scelta metodologica di svolgere interviste con la tecnica della conversazione guidata integralmente registrata ha portato all'allestimento di un imponente *corpus* di parlato, impiegato però esclusivamente dalla redazione per il controllo dei dati. Questi ultimi infatti sono riprodotti sulle carte in trascrizione fonetica secondo il sistema dell'Alfabeto Fonetico Internazionale, in seguito a un'operazione di conversione della trascrizione originale, che era stata invece condotta secondo il sistema ALF-Rousselot, in continuità con gli atlanti regionali francesi delle limitrofe aree linguistiche transalpine.

- 13 Merita inoltre una segnalazione il saggio di Atlante parlato del Piemonte presentato nel 1995 all'interno della mostra intitolata *Il Piemonte linguistico* organizzata dal Museo nazionale della Montagna «Duca degli Abruzzi» di Torino e dalla Regione Piemonte (*Il Piemonte linguistico*, 1995). Infatti con questo progetto è stata testata per la prima volta su dati linguistici piemontesi la possibilità di rendere udibile la risposta originale dell'intervistato contestualmente alla visione sulla carta della sua trasposizione in grafia fonetica. La ricerca, benché all'epoca assai innovativa sul piano tecnologico, si limitò però all'elaborazione di cinque carte linguistiche «multimediali» presentate nel corso della mostra, allestite a partire da inchieste sul campo appositamente svolte in una settantina di località del Piemonte.

5. L'Atlante Toponomastico del Piemonte Montano (ATPM) e il progetto *Mapping Alpine Place-names for Upward Sociality* (MAPforUS)

- 14 Tra le basi di dati dialettali del Piemonte va infine segnalata quella dell'*Atlante Toponomastico del Piemonte Montano* (ATPM), un progetto di ricerca del Dipartimento di Studi Umanistici dell'Università degli Studi di Torino, avviato nel 1983 da Arturo Genre sulla base di una convenzione con la Regione Piemonte, il cui obiettivo principale è la documentazione e la salvaguardia dei *corpora* toponimici di tradizione orale gravemente minacciati dalle profonde trasformazioni economiche e sociali che hanno interessato le valli alpine nel corso del Novecento e relativi ai 553 comuni inclusi nel circuito amministrativo delle Comunità montane (ora denominate Unioni di Comuni). Si tratta di un'area che si distingue, sul piano linguistico e culturale, per complessità e varietà, annoverando al suo interno parlate pedemontane (piemontesi, liguri, lombarde), galloromanze (occitane e francoprovenzali) e alemanniche (walser). L'attendibilità e la capillarità della documentazione sono garantite da inchieste sul terreno condotte Comune per Comune secondo i principi metodologici della geolinguistica, ossia mediante interviste semidirettive svolte da raccoglitori locali attingendo esclusivamente a una pluralità di fonti orali appartenenti alla comunità indagata, scelte tra i depositari privilegiati delle conoscenze riguardanti il patrimonio toponimico locale. I toponimi, localizzati dagli informatori su carta topografica predisposta dalla Redazione, sono

trascritti secondo il sistema ortografico appositamente predisposto per ciascuna area linguistica e, per la loro pubblicazione, anche secondo il sistema dell'Alfabeto Fonetico Internazionale. L'indagine non si limita però alla mera compilazione di liste di toponimi, ma si estende alla raccolta di tutte le informazioni linguistiche e extralinguistiche utili alla ricostruzione del quadro referenziale delle singole denominazioni di luogo, della loro collocazione nello spazio sociale abitato e vissuto dalla comunità di parlanti che le ha prodotte, le usa o le usava e infine delle motivazioni sottese al processo della *nominatio*.

- 15 Dagli anni Ottanta a oggi l'attività di raccolta sul campo ha permesso all'*équipe* dell'ATPM di allestire una banca-dati di circa 73 000 denominazioni di luogo provenienti da 155 comuni piemontesi, corrispondenti a circa un terzo di quelli previsti nel piano d'opera, una parte delle quali confluite nelle 55 monografie pubblicate a partire dal 1990¹⁶. Ognuna di esse, composta da un volumetto con l'elenco lemmatizzato dei toponimi ordinati alfabeticamente (da 300 a 2 000 a seconda delle località) e dalle relative carte topografiche inserite in Appendice o in un cofanetto separato, rappresenta una fotografia sinottica della distribuzione areale dei toponimi di una data comunità, corredata da varie informazioni complementari necessarie per una corretta interpretazione dei dati linguistici presentati. L'ATPM costituisce quindi uno straordinario archivio lessicale su cui sperimentare vari paradigmi di ricerca linguistica perché in grado di documentare sia la varietà linguistica delle valli piemontesi sia le diverse forme di insediamento, di organizzazione e di sfruttamento del territorio.
- 16 Il suo impianto teorico, inoltre, ha rappresentato il *background* per la messa a punto del Progetto di Ateneo 2016 *Mapping Alpine Place-names for Upward Sociality* (toponomastica alpina georeferenziata per lo sviluppo di nuove forme di socializzazione) il cui obiettivo generale è la documentazione e lo studio della toponimia di tradizione orale di venti borgate delle valli alpine piemontesi mediante la sperimentazione di più rapide modalità di escussione della documentazione toponimica e la realizzazione di nuove forme di diffusione e circolazione dei dati che ne garantiscano una valorizzazione e una fruizione anche al di fuori dei tradizionali circuiti scientifici¹⁷. Per quanto concerne il primo aspetto, si tratta di testare nello specifico ambito toponomastico un modello di ricerca fondato su processi partecipativi di *crowdsourcing* che, se garantirà al lavoro sul campo una tempistica accettabile, potrà dare nuovo impulso alle indagini pianificate dall'ATPM. Il sistema *crowdsourced*, integrato alla piattaforma web, permette infatti agli utenti interessati di collaborare all'inchiesta inviando toponimi, brevi testi descrittivi, file sonori, fotografie e video georiferiti direttamente dal proprio smartphone o tablet, attraverso un modulo (vale a dire un'area di accesso riservata sul sito web) o un'applicazione dedicata da installare sul proprio dispositivo, funzionante anche in assenza di connessione internet, con l'aggiunta quindi di un sistema di registrazione in locale dei dati e una funzione di sincronizzazione in presenza di una connessione stabile. Riguardo alla modalità di presentazione dei dati raccolti, è previsto il passaggio dalle rappresentazioni cartografiche dell'ATPM a mappe interattive dedicate a ciascuna delle venti borgate, la cui funzionalità sarà così collaudata su repertori toponimici numericamente esigui, per valutarne la futura applicabilità all'intero corpus toponimico di una comunità e, progressivamente, all'intero inventario dell'ATPM. Le mappe sono costruite sui dati cartografici liberamente accessibili di *OpenStreetMap* (OSM) e basate su un sistema di georeferenziazione delle denominazioni di luogo. Queste ultime possono essere richiamate dall'utente che naviga sulla mappa e correlate a una serie di informazioni extralinguistiche, ad esempio materiale fotografico, documenti audio e video. In tal modo

sarà possibile rendere fruibile e valorizzare l'archivio sonoro dell'ATPM, costituito dai file audio delle singole denominazioni di luogo e attualmente impiegato ai soli fini redazionali, ovvero per la verifica della trascrizione ortografica e fonetica dei toponimi. A ciò si aggiunge la possibilità di correlare, attraverso lo strumento-mappa, dati generalmente soggetti a rappresentazioni (e quindi ad analisi) separate e quindi procedere a letture integrate, per esempio di dati toponimici e di dati geografici (desumibili da carte tematiche del territorio, da metacarte e soprattutto dai dataset informatici messi a disposizione dalle amministrazioni o dagli enti gestionali del territorio)¹⁸.

6. Il dato linguistico registrato

- 17 Le ricerche sul parlato richiedono una marcata interdisciplinarietà che ne arricchiscono gli ambiti applicativi: fonetica, ingegneria linguistica, dialettologia, etnolinguistica, pragmatica, analisi conversazionale, ma anche sociolinguistica e psicologia (cfr. De Dominicis, 2002). Il parlato interessa trasversalmente quasi tutte le scienze umane perché dipende dal comportamento individuale e offre una notevole dimensione di espressione culturale.
- 18 Come ricorda Baude (2006), la profonda trasformazione della società, con l'innovazione tecnologica e la globalizzazione dell'informazione, ha riportato in auge e rimesso in discussione la nozione di cultura scientifica (insieme di valori vs. saperi – procedimenti e pratiche – contesti decisionali, economici e/o politici). Si è andato modificando e incrementando il fabbisogno sociale, che ha cominciato a trovare risposte anche nell'ambito delle istituzioni pubbliche attraverso la creazione di centri culturali, scientifici, tecnici a partire dai primi anni Ottanta. In Italia, la legge n. 113 del 1991 ha, fra l'altro, dato impulso a forme di integrazione mediante reti telematiche delle attività delle istituzioni operanti nel campo della ricerca scientifica e tecnologica¹⁹.
- 19 Un esempio significativo, proprio riguardo all'accesso a dati linguistici, è costituito dall'ELRA (European Language Resources Association), un'organizzazione senza scopo di lucro nata nel 1995 al fine di fornire una gestione centralizzata di corpora e archivi e garantire la distribuzione di risorse e strumenti linguistici (affidata a un centro operativo denominato ELDA – Evaluations and Language resources Distribution Agency).
- 20 Se le prime iniziative sono state rivolte alla documentazione della variazione delle lingue nazionali (per l'italiano si vedano, tra gli altri API e la sezione italiana di C-ORAL-ROM)²⁰, un'attenzione crescente è stata dedicata più recentemente al patrimonio linguistico di realtà interessate da maggiore fragilità (come quelle recensite da Huma-Num – LACITO, 2013)²¹. Diverse sono le indicazioni recepite nelle esperienze applicative di progetti come *Pangloss* e *Ortolang*, fino a giungere alla definizione di un Virtual Language Observatory (VLO) e di un consorzio omnicomprensivo: CLARIN (European Common Language Resources and Infrastructure for Social Sciences and Humanities)²².

7. Archivio vocale del LFSAG

- 21 Il Laboratorio di Fonetica Sperimentale «Arturo Genre» (LFSAG), istituito sin dai primi anni '70 da Arturo Genre presso l'Istituto dell'Atlante Linguistico Italiano, ha ripreso le sue attività dal 2006, integrandosi progressivamente anche al Centro Linguistico

dell'Ateneo torinese (CLA). Dopo alcuni mesi di sospensione (e traslochi temporanei), per ragioni di agibilità dell'edificio che li ospitava, il LFSAG, così come l'Istituto dell'ALI, ha ritrovato una posizione a cavallo tra i due centri, in seguito alla riapertura di Palazzo Nuovo avvenuta il 9 febbraio 2017. L'evento ha reso possibile la riorganizzazione dei dati raccolti in più di dieci anni di attività e la pubblicazione *on-line* (sul sito < www.lfsag.unito.it/ark.html>) dei materiali relativi a lingue e dialetti diversi forniti da ricercatori permanenti o in visita, nonché da laureandi e dottorandi coinvolti nello studio di aspetti linguistici diversi. Oltre alle distinte sezioni dedicate a singole varietà dialettali d'Italia (tra le altre quella piemontese di Bagnolo Piemonte, CN, o quella salentina di Supersano, LE), sondate per mezzo di questionari diversi²³, la pagina web degli archivi ospita attualmente una sezione relativa a «Le voci di VINCA» (brani di parlato elicitato per immagini relative al progetto VINCA, *Varietà di Italiano di Nativi - Corpus Appaiato*) e una sezione riguardante provvisoriamente i dati di AMPER-ITA, la sezione italiana del progetto internazionale dell'*Atlante Multimediale Prosodico delle varietà romanze* (AMPER) il cui scopo è la descrizione della variabilità prosodica dialettale. Altre sezioni in fase di sviluppo riguardano il parlato ortofonico (*PhoneIt!*), letto, cantato e recitato (v. ora il progetto VIP)²⁴.

- 22 La sezione più ricca e interessante del sito è tuttavia quella delle registrazioni relative a versioni in lingue diverse del breve testo de «La tramontana e il sole», passato dai 130 campioni del 2013 ai quasi 400 del 2017²⁵. Nata inizialmente (sin dal 2007) con l'intenzione di documentare la variabilità dialettale in Italia e poi per soddisfare l'esigenza di testare alcune metriche ritmiche su un insieme differenziato di lingue, la collezione si è estesa a un considerevole numero di versioni di questo racconto, richiedendo attività di presentazione che sono ancora in fase di svolgimento. Al momento, infatti, per ogni varietà linguistica documentata (dai numerosi dialetti sardi e siciliani e ad alcune lingue minoritarie parlate in Italia a lingue minacciate di estinzione come il Gorane di Kassatcha, Ciad), oltre al file sonoro in formato compresso, mp3 a basso bitrate concesso con licenza *GPL Creative Commons*, il record contiene soltanto: un testo in trascrizione ortografica grezza (quando disponibile), una valutazione sommaria della qualità della registrazione, i crediti relativi ai raccoglitori del campione e l'indicazione di eventuali pubblicazioni associate (da menzionare nel caso se ne faccia un uso in opere pubblicate)²⁶. È allo studio la possibilità di corredare uniformemente i file sonori con un file di annotazione linguistica a più livelli (già disponibile per molti sottoinsiemi di dati, anche se non pubblicato sul sito)²⁷.

8. Conclusioni

- 23 In questo contributo riassuntivo abbiamo inteso sintetizzare le attività svolte da alcuni gruppi di ricerca che gravitano attorno all'Istituto dell'Atlante Linguistico Italiano e al Laboratorio di Fonetica Sperimentale «Arturo Genre» dell'Università di Torino e che mirano alla raccolta e al trattamento di dati linguistici registrati e archiviati in vari formati e per varie finalità. In particolare, alle raccolte più tradizionali si sono affiancate negli ultimi anni quelle di dati relativi a realtà linguistiche in rapida trasformazione che includono situazioni che occorre documentare con urgenza. In molti casi, il formato dell'unità di archiviazione non è più soltanto ristretto a una forma grafica o al semplice file sonoro registrato, ma comincia a essere rappresentato da un corredo d'informazioni organizzato schematicamente. L'allestimento di basi di dati sonori su parlate a rischio

d'estinzione o su lingue la cui rappresentazione richieda particolari accorgimenti (ad es. per via della complessità dello stesso diasistema linguistico o per la mancanza di grammatiche descrittive), induce a ricercare protocolli di annotazione robusti e collaudati, orientandosi verso quelli che godono di più ampio consenso nella comunità scientifica internazionale.

BIBLIOGRAFIA

ALI = BARTOLI Matteo, TERRACINI Benvenuto, VIDOSSÌ Giuseppe, GRASSI Corrado, GENRE Arturo & M. ASSOBRIO Lorenzo, *Atlante Linguistico Italiano*, Roma, Istituto Poligrafico e Zecca dello Stato (7 volumi pubblicati + materiali inediti c/o Istituto dell'Atlante Linguistico Italiano, Università di Torino, dal 1995).

AVOLIO Francesco (a cura di), 2016, *Archivi Etnolinguistici Multimediali*, «Quaderni del Museo delle Genti d'Abruzzo», n° 41 (numero dedicato agli atti del convegno *Archivi Etnolinguistici Multimediali*, Pescara, 5-6 ottobre 2012).

BAUDE Olivier, 2006, *Corpus oraux : guide des bonnes pratiques*, Paris, CNRS.

BAUER Roland, 1995, *Vivaldi-Sicilia. Documentazione sonora dei dialetti siciliani*, in G. Ruffino (a cura di), *Percorsi di geografia linguistica. Idee per un atlante siciliano della cultura dialettale e dell'italiano regionale*, Palermo, Centro di Studi Filologici e Linguistici Siciliani, pp. 543-550.

CALAMAI Silvia & BERTINETTO Pier Marco, 2012, *Per il recupero della Carta dei Dialetti Italiani*, in T. Telmon, G. Raimondi & L. Revelli (a cura di), *Coesistenze linguistiche nell'Italia pre- e postunitaria* (Atti del XLV Congresso internazionale di studi della Società di Linguistica Italiana, Aosta/Bard/Torino 26-28 settembre 2011), Roma, Bulzoni, pp. 335-356.

CLARIN = European Common Language Resources and Infrastructure for Social Sciences and Humanities (CLARIN consortium) + Virtual Language Observatory (VLO).

CUGNO Federica & CUSAN Federica (in stampa), *MAPforUS (Mapping alpine place-names for upward sociality): toponimia tradizionale e nuovi modelli di socialità*, in *Atti del Congresso internazionale di Dialettologia «Dialetto e Società»* (Sappada, 3-7 luglio 2017).

CUGNO Federica & MASSOBRIO Lorenzo, 2010, *Gli atlanti linguistici della Romània*, Alessandria, Edizioni dell'Orso.

CUGNO Federica, RIVOIRA Matteo & RONCO Giovanni, 2018, *Piccolo Atlante Linguistico del Piemonte (PALP)*, Torino, Istituto dell'Atlante Linguistico Italiano.

DE DOMINICIS Amedeo (a cura di), 2002, *La voce come bene culturale*, Roma, Carocci.

DE DOMINICIS Amedeo & MATTANA Pamela, 2007, *Il Progetto Bomarzo*, in L. Romito, V. Galatà & R. Lio (a cura di), *La fonetica sperimentale: metodi e applicazioni* (Atti del IV Convegno nazionale AISV, Cosenza, 3-5 dicembre 2007), Torriana (RN), EDK, pp. 405-411.

FRANCESCHI Temistocle, 1967, *Achèvement des recherches dialectales pour l'Atlante Linguistico Italiano (ALI). 120 enquêtes linguistiques à travers l'Italie*, in L. E. Schmitt (a cura di), *Verhandlungen des 2. Internationalen Dialektologenkongresses* (Marburg/Lahn, 5-10 settembre 1965), «Zeitschrift für Mundartforschung», vol. 32, n° 3-4, pp. 232-238.

FRONTINI Francesca & CALAMAI Silvia, 2017, *Speech Audio Archives and CLARIN Metadata*, comunicazione presentata all'International Workshop *Speech Audio Archives: Preservation, Restoration, Annotation, aimed at supporting the Linguistic Analysis* (Roma, Accademia Nazionale dei Lincei, 18-19 maggio 2017), in c. di p.

GENRE Arturo, 1978, *Tavola di unificazione dei segni di trascrizione fonetica di Ugo Pellis*, *Bollettino dell'Atlante Linguistico Italiano*, III Serie, n° 2, pp. 57-86.

GHIRARDINI Cristina, LECHLEITNER Gerda & LIEBL Christian (a cura di), 2014, *Friedrich Schürer's Recordings from Romagna (1914)*, Wien, VÖAW.

GRENET Sylvie, 2006, *Problématiques et enjeux du patrimoine culturel immatériel au ministère de la Culture*, communication donnée lors des rencontres « Patrimoine culturel immatériel et transmission : la polyphonie corse traditionnelle peut-elle disparaître ? » (Centre des musiques traditionnelles corses, Ajaccio 22 et 23 juin 2006). <www.iiaac.cnrs.fr/IMG/pdf/Grenet.pdf> (ultimo accesso 20 dicembre 2017).

HINRICHS Erhard & KRAUWER Steven, 2014, *The CLARIN Research Infrastructure: Resources and Tools for e-Humanities Scholars*, in *Proceedings of the 9th International Conference on Language Resources and Evaluation* (LREC, Reykjavík, 26-31 maggio 2014), pp. 1525-1531.

HUMA-NUM – LACITO, 2013, *La Tribune d'Huma-Num. Partage d'expériences*, «La lettre de l'INSHS», n° 25, pp. 26-29.

Il Piemonte linguistico, 1995, Torino, Museo nazionale della montagna Duca degli Abruzzi, Club alpino italiano.

KATTENBUSCH Dieter, 1995, *Atlas parlant de l'Italie par régions : VIVALDI*, in *Estudis de lingüística i filologia oferts a Antoni M. Badia i Margarit*, vol. 1, Barcelona, Publicacions de l'Abadia de Montserrat, pp. 443-455.

LECHLEITNER Gerda (a cura di), 1999, *Stimmporträts*, Wien, VÖAW (Sound Documents from the Phonogrammarchiv of the Austrian Academy of Sciences: The Complete Historical Collections 1899-1950).

JADE Mariannick, 2006, *Patrimoine immatériel : perspectives d'interprétation du concept de patrimoine*, Paris, L'Harmattan.

LORENZETTI Luca, MILIZIA Paolo & SCHIRRU Giancarlo, 2017, *Collecting Spoken Data from Upper-Southern Italian Dialects*, comunicazione presentata all'International Workshop *Speech Audio Archives: Preservation, Restoration, Annotation, aimed at supporting the Linguistic Analysis* (Roma, Accademia Nazionale dei Lincei, 18-19 maggio 2017), in c. di p.

MASSOBRIO Lorenzo, 2008, *Mi sono rampigato fora dai malanni, o squasi. Impegno e certezze di Pellis nel lavoro di preparazione dell'Atlante Linguistico Italiano*, in S. Perulli (a cura di), *Ugo Pellis. Un fotografo in movimento*, Udine, Società Filologica Friulana, pp. 63-72.

MINISTÈRE FRANÇAIS DE LA CULTURE ET DE LA COMMUNICATION, 2008, *Le patrimoine culturel immatériel*, «Culture & Recherche», n° 116-117. <www.culture.gouv.fr/var/culture/storage/lettre-recherche/cr116-117.pdf>.

- MONACHINI Monica & FRONTINI Francesca, 2016, *CLARIN, l'infrastruttura europea delle risorse linguistiche per le scienze umane e sociali e il suo network italiano CLARIN-IT*, «Italian Journal of Computational Linguistics», vol. 2, n° 2, pp. 11-30. <http://www.ai-lc.it/IJCoL/v2n2/1-monachini_and_frontini.pdf> [ultimo accesso 24 giugno 2017].
- MÜLLER Marcel, KÖHLER Carola & KATTENBUSCH Dieter, 2001, *VIVALDI – ein sprechender Sprachatlas im Internet als Beispiel für die automatisierte, computergestützte Sprachatlasgenerierung und -präsentation*, «Dialectologia et Geolinguistica», n° 9, pp. 55-68. <www2.hu-berlin.de/vivaldi/publikationen/vivaldi-sprechender-sprachatlas-2001.pdf> [ultimo accesso 20 dicembre 2017].
- Ortolang* = Outils et Ressources pour un Traitement Optimisé de la LANGue. <www.ortolang.fr> [ultimo accesso 24 giugno 2017].
- Pangloss* = Collection Pangloss, LACITO – CNRS. <lacito.vjf.cnrs.fr/pangloss/> [ultimo accesso 24 giugno 2017].
- PELLIS Ugo, 1936, *Relazione preliminare per l'edizione dell'Atlante Linguistico Italiano*, «Ce fastu?», n° 8-10, pp. 13-58.
- PISTACCHI Massimo, AQUILANTI Francesco & BALDI Francesco, 2006, *Verba Manent: Teoria e prassi della conservazione e promozione dei documenti sonori e video della Discoteca di Stato-Museo dell'audiovisivo (prima parte)*, «DigItalia. Rivista del digitale nei beni culturali», vol. 2, pp. 131-148.
- RIVOIRA Matteo, 2011, *L'Atlante Linguistico Italiano (ALI) et l'Atlante Toponomastico del Piemonte Montano (ATPM) : histoire, méthodes de récolte et de présentation des données, et perspectives de développement de deux entreprises géolinguistiques turinoises*, in F. Manzano (a cura di), *Mémoires du terrain : enquêtes, matériaux, traitement des données (Actes du colloque de Lyon, 12 et 13 mars 2009)*, Lyon, Publications du CEL, pp. 287-313.
- RIVOIRA Matteo, 2016, *L'archivio dell'ALI: informatizzazione dei dati e prospettive di sviluppo*, in F. Avolio (a cura di), *Atti del Convegno Gli archivi etnolinguistici*, pp. 49-60.
- ROMANO Antonio, 2007, *À propos de deux bases de données de parole publiées récemment : compte rendu de API – Archivio del Parlato Italiano et de C-ORAL-ROM*, «Rivista di Linguistica», vol. 19, n° 2, pp. 367-390.
- ROMANO Antonio, 2016, *La BD AMPER, La tramontana e il sole e altri dati su lingue, dialetti, socioletti, etnoletti e interletti del Laboratorio di Fonetica Sperimentale «Arturo Genre»*, in F. Avolio (a cura di), *Archivi Etnolinguistici Multimediali*, «Quaderni del Museo delle Genti d'Abruzzo», n° 41, pp. 225-240.
- ROMANO Antonio, DE IACOVO Valentina, CARPITELLI Elisabetta, BUCCI Jonathan & RONCO Giovanni, 2016, *Il dialetto di Corato (BA) in città e nelle comunità migranti (Isère, Francia)*, *Bollettino dell'Atlante Linguistico Italiano*, III Serie, n° 40, pp. 161-172.
- TERRACINI Benvenuto, 1955, *L'atlante linguistico italiano nell'opera di Matteo Bartoli e Ugo Pellis*, «Bollettino dell'Atlante Linguistico Italiano», Nuova Serie, n° 1, pp. 3-9.
- VENTURA Marina & AQUILANTI Francesco, 2007, *Verba Manent: Teoria e prassi della conservazione e promozione dei documenti sonori e video della Discoteca di Stato-Museo dell'audiovisivo (seconda parte)*, «DigItalia. Rivista del digitale nei beni culturali», vol. 1, pp. 81-94.

NOTE

1. L'articolo presenta una versione estesa di quello destinato al volume 5 di «Archivi e biblioteche in Piemonte», *I documenti sonori: Voce, suono, musica in archivi e raccolte* (a cura di D. Brunetti,

D. Robotti & E. Salvalaggio). Sebbene concepito e redatto congiuntamente da parte di tutti gli autori nelle parti generali (note e riferimenti), l'articolo è da attribuirsi principalmente a: MR (Introduzione, 1.1, 1.2), FC (1.2-1.4) e AR (Introduzione, 2.1, 2.2 e Conclusioni).

2. Riteniamo opportuno ridurre al minimo in questa sede il riferimento a quelle applicazioni nelle quali un sistema *IPA-like* (generalmente basato su etichette convenzionali) è usato ancora oggi per associare a un file sonoro una documentazione che ne codifichi i contenuti a vari livelli di annotazione. Per fare anche solo un esempio citiamo qui il sistema SAMPA – *Speech Assessment Methods Phonetic Alphabet* (v. SAMPA: *Computer Readable Phonetic Alphabet*, a cura di J. C. Wells, 1995-1998, <www.phon.ucl.ac.uk/home/sampa/index.html>) e la sua applicazione italiana nell'ambito del progetto CLIPS – *Corpora e Lessici di Italiano Parlato e Scritto* (<www.clips.unina.it>).

3. Il titolo originale dell'opera (consultabile al sito <www3.pd.istc.cnr.it/navigais-web>) è *Sprach- und Sachatlas Italiens und der Südschweiz*.

4. I primi progressi in quest'ambito sono ancora oggi al centro di ricerche e scoperte documentate in varie occasioni anche al grande pubblico (v. artt. apparsi su «La Repubblica», 27 marzo 2008: prima registrazione di una voce umana, ad opera di Édouard-Léon Scott de Martinville, 9 aprile 1860; 26 aprile 2013: *Ecco la mia voce, sono Alexander Bell*, documento sonoro del 1885 ad opera del figlio del già menzionato Graham Bell) e sono stati ricostruiti presso l'Istituto per la Conservazione dei Beni Sonori e Audiovisivi del Ministero dei Beni Culturali in occasione di una mostra (a cura di Francesco Aquilanti *et al.*, dal 2010) con l'intento di raccontare l'epopea della ricerca e della sperimentazione fono-discografica in Italia (<www.icbsa.it/index.php?it/569/fonografi-e-grammofoni>). Tra gli apparecchi presentati, spiccano senz'altro i celebri fonografi di Edison, introdotti tra il 1894 e il 1899 e subito adottati in molti centri europei che ambivano all'allestimento di raccolte di dati sonori per fini archivistici (tra questi ricordiamo l'Österreichische Akademie für Wissenschaften, le cui collezioni sono state incluse dall'UNESCO tra i documenti d'importanza universale nel Registro Mondiale «Memory of the World» (Lechleitner, 1999 e Ghirardini, Lechleitner & Liebl, 2014). Tra i pionieri in questo campo ricordiamo Joseph Nigra che aveva brevettato nel 1879 il suo sistema *Tinfoil*, una delle primissime macchine per l'incisione e la riproduzione della voce in Italia (cfr. Pistacchi, Aquilanti & Baldi, 2006 e Ventura & Aquilanti, 2007). I successivi sviluppi (anche della diffusione domestica) sono legati alla fortuna del grammofono (quello di Emile Berliner del 1887 o quelli britannici commercializzati a partire dal 1889-1890): grandi collezioni d'incisioni sono notoriamente quelle della Bibliothèque nationale de France (Département de l'Audiovisuel che ha assorbito gli *Archives de la Parole*, fondati nel 1911 da Ferdinand Brunot, e le fonoteche nazionali create negli anni '30). Gli sviluppi delle registrazioni su vinile, su nastro (Brüel & Kjær, 1942; Revox, 1948; Nagra, 1951; Philips, 1963, ecc.) e poi su supporto digitale (CD/DVD, 1982/1995) sono ben testimoniati nelle collezioni di documenti archiviati progressivamente presso la Discoteca di Stato e le collezioni AV della RAI e dei centri di documentazione etnomusicologica (v. tra gli altri, <<http://archiviosonoro.org>> e <www.mappasonora.it>).

5. All'importanza di raccogliere e rendere fruibile una documentazione audiovisiva storica sulla cultura immateriale (inoggettuale nella definizione di Alberto Cirese), sono dedicate altre importanti iniziative in Italia (molte delle quali documentate in Avolio, 2016) e all'estero (come nel caso del progetto VIVALDI, sin da Bauer, 1995 e Kattenbusch, 1995; cfr. Müller, Köhler & Kattenbusch, 2001). Le preoccupazioni generali sono invece ben espresse nel numero monografico di «Culture & Recherche» (2008) dedicato a *Le patrimoine culturel immatériel*.

6. Big Data e Open Data sono state al centro dei temi trattati nella Conferenza *Forum TAL 2014* «TAL & Open Data», Università degli Studi di Torino (Rettorato), 21-22 gennaio 2014, e nel corso del Seminario «Archiviazione, disseminazione e riuso dei dati: a che punto siamo? Uno sguardo all'indietro per andare avanti» svoltosi presso il Dip. di Culture Politiche Società, il 23 novembre 2017. Oltre agli snodi giuridici sulla proprietà del dato (a chi appartiene? Chi ne è responsabile? Il singolo individuo che produce il dato o chi lo raccoglie e lo sfrutta per scopi d'interesse collettivo

(o privato?) o sulla sua riservatezza (quando è opportuna o prescrittiva la vaghezza del dato?) e i risvolti relativi ai diritti di citazione o a un eventuale sfruttamento economico (quali sono le modalità di cessione dei diritti sul dato?), questi eventi hanno permesso di approfondire anche aspetti più tecnici (in molti casi già affrontati, in riferimento alla realtà francese, da Grenet, 2006 e, più in generale, da Baude, 2006). Alla necessità di provvedere in modo organico alla definizione di metadati, si associano quelle di fornire adeguate piattaforme di accesso e indicare adeguati approcci interpretativi. Questi incidono in generale sulle fasi della loro fruibilizzazione (*ingest*, indicizzazione, disseminazione). In particolare sembra rilevante associare ai dati un identificativo persistente (DOI) che consenta di risalire alla sua origine e garantisca la sua conservazione, la sua tutela e il recupero della sua collocazione storico-geografica originaria o lo storico dei suoi movimenti e della sua riproduzione (argomento su cui ha invitato a riflettere l'intervento di Diego Robotti della Soprintendenza Archivistica e Bibliografica del Piemonte e della Valle d'Aosta).

7. Si veda per questi De Dominicis & Mattana (2007) e Calamai & Bertinetto (2012). Più in generale, riguardo all'interesse degli archivi vocali per la linguistica e la ricerca sul parlato, si veda anche De Dominicis (2002).

8. Il *Visible Speech* di Alexander Melville Bell, un sistema di trascrizione non alfabetico, pur usato per scopi pratici (insegnamento della lingua ai sordi), non ha avuto importanti applicazioni nell'ambito della ricerca.

9. Anche la «Rivista Italiana di Dialettologia (RID)» nel 1980 promuove un nuovo sistema di trascrizione, in questo caso però si tratta di un sistema a base ortografica fortemente semplificato il cui scopo è quello di rendere con sufficiente precisione i suoni dei diversi dialetti senza rinunciare all'intelligibilità del sistema per i non specialisti.

10. La bibliografia riguardante la storia e l'organizzazione è assai estesa. Per una panoramica relativamente approfondita, si faccia riferimento a Pellis (1936), Terracini (1955), Franceschi (1967), Ronco (2004), Massobrio (2008), Cugno-Massobrio (2010, pp. 58-73), Rivoira (2011); cfr. inoltre il sito <www.atlantelinguistico.it>.

11. Costituiscono invece una novità metodologica l'inclusione delle parlate alloglotte, la struttura del questionario, articolato in varie sezioni per adattarsi meglio all'eterogeneità del territorio indagato e corredato di domande illustrate, e la selezione delle località da esplorare secondo parametri rappresentati dalle vie di comunicazione e dalla presenza o meno di determinati istituti amministrativi, scolastici ed ecclesiastici, allo scopo di rappresentare in giusta misura centri massimi, medi e minimi.

12. Dal 1995 al 2011 sono stati pubblicati otto volumi per i tipi dell'Istituto Poligrafico e Zecca dello Stato: I: *Il corpo umano*, di 96 carte con 282 voci; II: *Il corpo umano - Le sue funzioni*, di 109 carte con 290 voci; III: *Indumenti e abbigliamento*, di 95 carte con 178 voci; IV: *La casa e l'arredamento*, di 95 carte con 145 voci; V: *La casa e l'arredamento - La cucina*, di 132 carte con 219 voci; VI: *L'alimentazione*, di 90 carte con 179 voci; VII: *Le età dell'uomo - Infanzia e giovinezza, i giochi*, di 109 carte con 219 voci; VIII: *Le età dell'uomo - La scuola, l'età adulta, il matrimonio, la parentela*, di 107 carte con 245 voci. Il IX volume intitolato *La società - In strada, veicoli, in viaggio, alla fontana, all'osteria, racconti da osteria* e edito dall'Atlante Linguistico Italiano è previsto per il 2018.

13. Il numero delle inchieste non coincide con quello dei Punti perché il P. 29 Locana comprende anche l'inchiesta della sua frazione San Lorenzo e il P. 78 Pietraporzio include anche l'indagine svolta a Bersezio.

14. Informazioni dettagliate sul progetto e sui suoi attuali sviluppi sono reperibili sul sito <www.alepo.eu>.

15. Il piano completo dell'opera prevede la pubblicazione di altri otto volumi, ciascuno dei quali suddivisi in fascicoli: vol. II: *L'agricoltura e i suoi prodotti*; vol. III: *Il mondo animale*; vol. IV: *L'allevamento e le attività connesse*; vol. V: *Lo spazio e il tempo*; vol. VI: *L'uomo*; vol. VII: *La casa e le attività domestiche*; vol. VIII: *La società e la vita comunitari*; vol. IX: *Morfologia*. Oltre ai numerosissimi

studi pubblicati in corso di realizzazione dell'opera, si segnala P. Avena, E. Balbis & E. Miola, *Alepo I-I: lettura interpretativa di alcune carte*, Alessandria, Dell'Orso, 2008.

16. Per informazioni dettagliate sull'area di indagine, sulla metodologia di ricerca, sulle pubblicazioni e sugli sviluppi del Progetto si consulti il sito dell'ATPM (<www.atpmtoponimi.it>).

17. Per informazioni più dettagliate sul progetto e sui suoi sviluppi si rimanda alla sezione dedicata a MAPforUS sul sito dell'*Atlante Toponomastico del Piemonte Montano* (<www.atpmtoponimi.it>).

18. Un esempio di lettura integrata di questo tipo figura in Cugno & Cusan (in stampa).

19. Numerosi progetti si stanno definendo a livello internazionale per consentire un accesso multilingue a milioni di riproduzioni di opere, immagini di manufatti artistici, archivi fotografici, prodotti audiovisivi, musicali e letterari. Importanti tappe si stanno raggiungendo in vista della costruzione di una biblioteca digitale sostenuta dalla Commissione europea.

20. Tanto l'*Archivio di Parlato Italiano* quanto il *Corpus ORAL des langues ROManes* sono recensiti in Romano (2007). Un altro progetto italiano che ha contribuito a formare personale specializzato e definire linee operative nazionali è il già citato CLIPS – *Corpora e Lessici di Italiano Parlato e Scritto* (v. n° 2).

21. Come utilmente ricordato da Baude (2006), la digitalizzazione e l'accessibilità *on-line* dei dati linguistici sono essenziali alla valorizzazione del patrimonio scientifico e culturale, al dinamismo della creazione di nuovi contenuti e al miglioramento di nuovi servizi. Contribuiscono alla democratizzazione e all'accesso alla cultura, all'evoluzione della società dell'informazione e all'industria della conoscenza.

22. In merito a quest'ultimo si veda Hinrichs & Krauwer (2014) e, per il suo accoglimento in Italia, Monachini & Frontini (2016) e Frontini & Calamai (2017).

23. Questa sezione ospita anche i risultati del progetto Galileo 216-G14 dell'Università Italo-francese sul dialetto parlato dalle comunità italiane in Francia (cfr. Romano *et al.*, 2016).

24. VIP – *Voices of Italian Poets* sarà una piattaforma digitale di documenti sonori della poesia italiana contemporanea e di relative analisi fonetiche. L'archivio rientra tra gli obiettivi del Dottorato di Ricerca in *Digital Humanities* di Valentina Colonna (UniGE-UniTO) e offrirà diverse interpretazioni (a opera di poeti contemporanei, speaker, attori, critici) di testi della letteratura del Novecento per i quali si conta di recuperare anche le letture originali degli autori. L'archivio, in fase di costruzione, renderà accessibile a ricercatori e studiosi, nonché a un pubblico di appassionati, lettori e poeti, una rassegna di interpretazioni diverse e uniche del testo poetico. In modo analogo al rapporto tra composizione e interpretazione in musica, l'archivio si propone di individuare le potenzialità sonore della scrittura poetica partendo dall'osservazione sperimentale (mediante analisi ritmico-intonative) del rapporto tra il testo e le sue varie letture. Il progetto continua il lavoro dedicato nel 2016-2017 da Valentina Colonna alle «Prosodie del Congedo del viaggiatore cerimonioso di Giorgio Caproni». Otto di dodici registrazioni analizzate costituiscono il primo nucleo di archivio e di ricerca sull'argomento e inaugurano la piattaforma (le voci sono quelle di poeti contemporanei e speaker radiofonici che si sono resi disponibili nella lettura della prosopopea caproniana). Lo studio originario ha preso in analisi anche la lettura dell'autore, nucleo centrale e di partenza del lavoro, e di tre attori, tra cui Achille Millo, dedicatario del testo. Centrali in questo studio sono l'individuazione delle principali «categorie intonative» melodiche, l'indagine sul ritmo del verso libero italiano, partendo dall'osservazione del dato acustico in relazione alla struttura testuale, l'individuazione di un impianto retorico della prosodia e l'osservazione di tendenze dominanti davanti ai fenomeni retorici del testo (quali l'inarcatura, la rima, l'assonanza, la consonanza).

25. La raccolta di campioni in lingue diverse di versioni di questa favola esopica è perseguita in vari ambiti ricordati in Romano (2016). Tra le fonti di maggiore autorevolezza, ricordo la prima edizione dei *Principles of the International Phonetic Association* (1912) che, nelle edizioni successive, si arricchisce progressivamente d'illustrazioni di lingue e dialetti, fino a offrire negli ultimi anni

un insieme piuttosto coerente di campioni sonori resi disponibili ai soci sul sito dell'Associazione (<www.internationalphoneticassociation.org>). Ricordo ancora il campione di produzioni regionali offerto da C. Tagliavini (nei dischi in vinile allegati al volume *La corretta pronuncia italiana: corso discografico di fonetica e ortoepia*, 1966) e il *DOP - Dizionario di ortografia e pronunzia* di B. Migliorini, C. Tagliavini & P. Fiorelli (1969) che, dal 2008, vanta anche un'edizione *on-line* (<www.dizionario.rai.it>) nella quale è disponibile il testo di numerose versioni de «La tramontana e il sole» spesso corredate da file sonori. Più recentemente, sono tuttavia i lavori di L. Canepari che offrono la maggiore varietà di lingue descritte anche attraverso questo testo; alcune di queste descrizioni sono ora corredate da file sonori nel sito «Natural Phonetics» (<<http://canipa.net>>).

26. Il sito di pubblicazione di queste basi di dati sonori si orienta verso i modelli definiti dal CREM (<www.crem-cnrs.fr>) e dallo SLDR (DDL, Lyon + LPL, Aix-en-Provence) nell'ambito degli archivi aperti Huma-Num (v. § 6). Il formato originario dei dati è quello wav/PCM (16bit-16kHz); tuttavia, allo stato attuale, solo dati sono stati forniti pubblicamente, secondo modalità di presentazione provvisorie a scopo divulgativo. I ricercatori interessati a una migliore qualità possono farne richiesta ai gestori del sito.

27. Questi archivi sono destinati a un pubblico di utenti non specialisti, ma offrono una certa utilità anche agli specialisti, mirando a consentire agli utenti delle lingue così 'fotografate' di ritrovarsi linguisticamente e/o stimolare la riflessione sul dato linguistico e sui suoi eventuali contrasti con le loro intuizioni. Le registrazioni rese fruibili risultano in alcuni casi come prodotto di un riversamento di registrazioni originali di diversi periodi e di altri autori e, in molti casi sono già servite per saggi esplorativi su alcune particolarità delle lingue rappresentate. La vocazione del sito ad aprirsi a contributi di ricercatori diversi potrebbe suggerirne una trasformazione successiva in un vero e proprio *repository* oppure, se non altro, rifluire in una più ampia rete istituzionale o nel polo di aggregazione di *repository* italiani in fase di istituzione presso il CNR (ACCESS CNR-ISTI) sulla scia di altre iniziative tese a dare visibilità e accrescere l'accessibilità ai risultati della ricerca scientifica italiana, assicurando un accesso più ampio agli standard di interoperabilità e adottati standard tecnologici per facilitare l'archiviazione. Queste finalità sono ben incarnate al livello europeo dal consorzio CLARIN.

RIASSUNTI

In quest'articolo sono riassunte alcune considerazioni sull'utilità della trascrizione/ annotazione dei dati linguistici con particolare attenzione alla funzione di conservazione delle parlate dialettali o delle lingue minacciate di estinzione. Alla descrizione sommaria delle prime operazioni dialettologiche dei grandi progetti atlantistici, si associa un riferimento al crescente diffondersi di progetti di allestimento di archivi vocali e alla necessità di adottare protocolli condivisi in un'ottica di coordinamento internazionale (sul modello di quelli offerti da consorzi come ELRA o CLARIN).

Cet article prend en compte quelques considérations sur l'utilité de la transcription/ annotation de données linguistiques en accordant une attention particulière à la fonction de conservation des dialectes ou des langues dialectales menacées d'extinction. À la brève description des premières opérations dialectologiques des grands projets d'atlas, nous associons une référence à la diffusion croissante des projets de constitution d'archives vocales et à la nécessité d'adopter

des protocoles partagés en vue d'une coordination internationale (sur le modèle de consortiums tels que ELRA ou CLARIN).

This article summarizes some considerations on the usefulness of the transcription/ annotation of linguistic data with particular attention to the conservation function of dialectal dialects or languages threatened with extinction. To the brief description of the first dialectological operations of the major atlases projects, we associate a reference to the growing spread of projects to set up vocal archives and the need to adopt shared protocols with a view to international coordination (on the model of those offered by consortia such as ELRA or CLARIN).

INDICE

Mots-clés : atlas dialectaux, archives orales, transcriptions et annotations de matériaux linguistiques

Keywords : dialect atlases, oral records, transcripts and annotations of language materials

Parole chiave : atlanti dialettali, archivi vocali, trascrizione e annotazione di materiali linguistici

AUTORI

ANTONIO ROMANO

Istituto dell'Atlante Linguistico Italiano, Laboratorio di Fonetica Sperimentale «Arturo Genre»,
Università degli Studi di Torino
antonio.romano@unito.it

MATTEO RIVOIRA

Istituto dell'Atlante Linguistico Italiano, Laboratorio di Fonetica Sperimentale «Arturo Genre»,
Università degli Studi di Torino
matteo.rivoira@unito.it

FEDERICA CUGNO

Istituto dell'Atlante Linguistico Italiano, Laboratorio di Fonetica Sperimentale «Arturo Genre»,
Università degli Studi di Torino
federica.cugno@unito.it

GIOVANNI RONCO

Istituto dell'Atlante Linguistico Italiano, Laboratorio di Fonetica Sperimentale «Arturo Genre»,
Università degli Studi di Torino
giovanni.ronco@unito.it

VALENTINA DE IACOVO

Istituto dell'Atlante Linguistico Italiano, Laboratorio di Fonetica Sperimentale «Arturo Genre»,
Università degli Studi di Torino
valentina.deiacovo@unito.it

VALENTINA COLONNA

Istituto dell'Atlante Linguistico Italiano, Laboratorio di Fonetica Sperimentale «Arturo Genre»,
Università degli Studi di Torino
valentina.colonna@unito.it