

Raccogliere dati di varietà romanze in via di estinzione: il caso delle lingue ereditarie in America¹

Abstract

Questo breve articolo racconta l'esperienza di pianificazione e messa in atto di una raccolta sul campo di dati di varietà italiane ereditarie in Argentina, Brasile, Canada e Stati Uniti. La raccolta è parte della ricerca legata al progetto ERC *Microcontact. Language variation and change from the Italian heritage perspective* e ha avuto luogo nella primavera del 2019. L'articolo si sofferma su alcuni aspetti della preparazione dell'inchiesta sul campo: quelli burocratici e quelli di natura pratica. Dopo una breve introduzione, si presentano brevemente i profili delle comunità italo-romanze nelle Americhe, per poi passare alla descrizione dei protocolli necessari per effettuare un'inchiesta di questo tipo, e concludere con alcuni suggerimenti pratici da tener presenti quando si vogliono coinvolgere persone anziane e bassamente scolarizzate nella documentazione delle proprie lingue.

1. Introduzione

Il lavoro di raccolta dei dati per le inchieste dialettologiche è solitamente molto oneroso. Svolgerlo avendo come target una moltitudine di dialetti, per lo più nella loro versione ereditaria, è un'impresa al limite dell'impossibile. Questo contributo si propone di offrire agli interessati un riassunto della raccolta sul campo, avvenuta tra l'autunno del 2018 e la primavera del 2019, di dati di alcune varietà italiane in contatto con altre varietà romanze in Argentina, Brasile, Canada e Stati Uniti. La raccolta dei dati aveva come oggetto le lingue ereditarie, ovvero le varietà parlate dagli emigranti e dai loro figli e nipoti in condizione diglottica, in contatto con varietà ad esse molto simili strutturalmente: lo spagnolo rioplatense, il portoghese brasiliano, il québécois canadese, e con una lingua strutturalmente altra, cioè l'inglese degli Stati Uniti. La ricerca si è svolta nell'ambito del progetto ERC *Microcontact*, in via di conclusione presso l'Università di Utrecht, nei Paesi Bassi.

L'articolo riporterà le fasi iniziali del progetto, comprendenti la preparazione dei protocolli per la protezione dei dati sensibili, la progettazione e l'espletazione della ricerca sul campo. I risultati della ricerca sono stati riportati in diversi lavori, che non verranno riassunti in questa sede (si rimanda a D'ALESSANDRO 2021; D'ALESSANDRO / NATVIG / PUTNAM 2021; FRASSON / D'ALESSANDRO / VAN OSCH 2021; ANDRIANI ET AL. 2022A; ANDRIANI ET AL. 2022B; ANDRIANI / D'ALESSANDRO IN STAMPA; D'ALESSANDRO / FRASSON IN STAMPA).

Il resoconto è organizzato secondo lo schema seguente: nel §2 si fornirà una breve introduzione al progetto *Microcontact*. Il §3 offrirà una breve panoramica delle comunità italiane nelle Americhe, il §4 presenterà alcuni suggerimenti su cosa tenere presente nella progettazione di una raccolta di dati fuori dall'UE, mentre il §5 conterrà alcuni suggerimenti per la raccolta di dati con parlanti incolti, anziani, e con difficoltà visivo-uditive. Il §6 conterrà alcune considerazioni conclusive.

2. Il progetto

Microcontact. Language variation and change from the Italian heritage perspective è il titolo del progetto, finanziato dallo European Research Council, che ha come base empirica le varietà italiane parlate nelle Americhe dalle comunità di emigranti. Si tratta di un progetto svolto in prospettiva formale-teorica, che ha come scopo principale la comprensione dei meccanismi del mutamento linguistico in contatto e lo sviluppo di un modello formale per la descrizione/predizione dei risultati di tale contatto. Il modulo grammaticale

1. Si ringrazia lo European Research Council per il finanziamento di questa ricerca (ERC_CoG 681959_MicroContact). Molte grazie anche ai partecipanti al Convegno internazionale di studi dal titolo *Dialettologia ed etimologia. Metodologie e cantieri aperti* per l'interessante discussione.

investigato è la sintassi, solitamente più resiliente verso le sollecitazioni esterne e più stabile rispetto al lessico e alla fonologia (si faccia riferimento alla discussione sull'*inerzia* linguistica (LONGOBARDI 2001 *et seq.*; KEENAN 1994; 2002).

Il progetto si compone di diverse parti: un'inchiesta preliminare, la raccolta dei dati sul campo, l'analisi dei dati, la formulazione delle ipotesi concernenti alcuni fenomeni sintattici, l'enunciazione di alcune generalizzazioni che riguardano il cambiamento linguistico nella sua componente cognitiva e nella sua componente grammaticale. Questo contributo si soffermerà solo sulla metodologia d'inchiesta, e in particolare sulla raccolta dei dati.

Nel progettare uno studio sulla linguistica di contatto incentrato sulle varietà ereditarie, il primo problema è stabilire quale dei cambiamenti registrati rispetto alle varietà originarie sia dovuto al contatto e quale sia avvenuto spontaneamente. Spesso non è possibile determinare con certezza se un cambiamento sia spontaneo o influenzato dalla lingua di contatto.

In generale, il cambiamento avvenuto per contatto e quello diacronico, endogeno, hanno caratteristiche abbastanza diverse, e ciò è riflesso nella letteratura scientifica che li esamina.

Il cambiamento linguistico in contatto è considerato, solitamente, imprevedibile. Quando due lingue entrano in contatto, entrano in contatto due società, due culture, due forze politiche solitamente non equipollenti. Poiché ogni situazione di contatto è unica, spesso non è possibile estrapolare generalizzazioni che riguardino i cambiamenti in tutte le possibili combinazioni di lingue in contatto. L'esito del contatto sulla struttura di una lingua dipende da fattori esterni alle grammatiche coinvolte, come hanno reso ben chiaro gli studi di molti specialisti in linguistica del contatto (ANDERSEN 1988; THOMASON / KAUFMAN 1988; AIKHENVALD / DIXON 2001; 2007; ROSS 2001; HEINE / KUTEVA 2005). I fattori che influenzano in misura diversa il cambiamento in contatto possono riguardare anche la grammatica (quanto sono simili le lingue tipologicamente; quanto un termine, un morfema, un suono possa essere integrato nel sistema ospitante; la salienza pragmatica dell'elemento preso in prestito; l'esistenza di un gap percettibile, e così via; si vedano a questo riguardo AIKHENVALD / DIXON 2001; 2007). Oltre a questi fattori squisitamente grammaticali, vanno considerati i fattori cognitivi e quelli sociolinguistici, che svolgono un ruolo molto importante nella determinazione del cambiamento. Per ciò che riguarda l'aspetto cognitivo, cosa accadrà a una lingua in contatto è determinato dall'età in cui la lingua viene appresa, dalla qualità dell'input, dalle altre lingue conosciute dal parlante, dalla condizione in cui le lingue vengono parlate (ROTHMAN 2009; 2010; 2011 *et seq.*; SORACE 2011; 2012 *et seq.*). Tra i fattori sociolinguistici rientrano il prestigio delle lingue in contatto, il desiderio di preservare la propria identità o di adottarne un'altra, l'atteggiamento del parlante riguardo alla lingua (AIKHENVALD / DIXON 2001; THOMASON / KAUFMAN 1988; HEINE / KUTEVA 2005). Molti di questi fattori hanno un impatto imprevedibile: il cambiamento in contatto dipende dal momento storico, dalle culture in contatto, dall'atteggiamento dei parlanti molto più di quanto si possa pensare. Per queste ragioni, molti studiosi riassumono la formula del cambiamento linguistico nella massima *everything goes*, 'tutto è possibile'.

Il cambiamento linguistico diacronico è, diversamente, più sistematico. Gli studi diacronici individuano quasi sempre una traiettoria di sviluppo della grammatica, solitamente indicata come traiettoria di *grammaticalizzazione*. La grammaticalizzazione, in termini sintattici, corrisponde alla perdita progressiva di elementi funzionali o al graduale impoverimento semantico di una forma. Il processo di grammaticalizzazione di una forma linguistica è spesso percepito come direzionale e progressivo. In lingue non in contatto raramente si assiste alla comparsa dal nulla di una costruzione grammaticale o all'introduzione di un tratto grammaticale.

Sebbene le caratteristiche e le cause del cambiamento linguistico in contatto e in diacronia siano molto diverse, come si diceva, non è sempre possibile stabilire se un cambiamento sia dovuto o no al contatto. Questo problema diviene particolarmente spinoso se il contatto avviene tra lingue molto simili, come quelle appartenenti alla stessa famiglia linguistica. Secondo AIKHENVALD (2006:9), «if languages are genetically related, we expect them to develop similar structures, no matter whether they are in contact or not. And if genetically related languages are in contact, trying to prove that a shared feature is contact-

induced and not a chance result of Sapir's drift may be next to impossible»². Pur essendo questa affermazione perfettamente condivisibile, va considerato che Aikhenvald, come quasi tutti gli studiosi del contatto, considera la tipica situazione di studio, ovvero una situazione di 1:1 (una lingua in contatto con un'altra lingua). In questo caso è ovviamente impossibile stabilire se un cambiamento, avvenuto in una lingua in contatto con una lingua a lei simile, sia attribuibile al contatto tra le due varietà o sia avvenuto spontaneamente. Il quadro però cambia radicalmente se si considera una situazione 1:molte, cioè una lingua in contatto con molte lingue a lei tipologicamente identiche o molto simili: in quel caso il punto di debolezza dell'approccio comparativo, e cioè la sua somiglianza della lingua in variazione con la lingua di contatto, diventa un punto di forza. La lingua può essere osservata al microscopio, focalizzandosi soltanto su un tratto e seguendone lo sviluppo, tenendo tutto il resto invariato. L'approccio 1:molte è alla base del progetto *Microcontact*. Un tratto sintattico non viene controllato solo in contatto con una lingua, ma in contatto con molte varietà, identiche tra loro tranne per il valore di quel tratto. La metodologia del *microcontatto* (D'ALESSANDRO 2015; 2021) non si limita, dunque, alla sola osservazione di un tratto alla volta, ma prevede l'osservazione dello stesso tratto in contatto con molte varietà identiche, con l'eccezione di quel tratto, limitatamente a un determinato dominio sintattico.

Schematicamente, l'idea del microcontatto può essere riassunta come in Figura 1.

Figura 1. *Microcontact*

- Grammatiche: A, B, C, D, E
identiche (in un dato dominio) eccetto un tratto X
- Tratto X in grammatica A **in contatto con**
 - Gramm. B
 - Gramm. C
 - Gramm. D
 - Gramm. E
- Tratto X in grammatica A **in isolamento**
- Stesse condizioni al contorno

La Figura 1 illustra la metodologia di base della ricerca sul *microcontatto*: si tratta di isolare un tratto X, che abbia un determinato valore nella grammatica A, e poi controllarlo in contatto con le grammatiche B, C, D, E ..., ciascuna delle quali possieda un valore diverso per X (con idealmente due o più grammatiche che abbiano lo stesso valore). Le grammatiche A-E sono identiche, in un dato dominio sintattico, tranne per il tratto X. Altro elemento fondamentale è la possibilità di tracciare l'evoluzione diacronica di X: la grammatica A deve essere documentata, deve essere possibile descrivere l'evoluzione temporale spontanea (per quanto possibile) del tratto X.

Inoltre, è necessario escludere l'importo di tutti i fattori non grammaticali, e che dunque la condizione del contatto sia pressoché identica per tutte le coppie di lingue.

Una situazione del genere non può essere costruita in laboratorio. È necessario trovare delle lingue che siano abbastanza vicine tra loro da poter essere considerate parte di un "macrogruppo" sintattico, e che siano attestate per secoli. Le lingue della famiglia romanza sono tra le pochissime lingue al mondo (forse paragonabili soltanto alle lingue germaniche e alle lingue slave) che rispondano a tutti questi requisiti: la quantità di varietà romanze è vastissima. In particolare, le varietà romanze in territorio italiano, solitamente indicate come italo-romanze, hanno una lunga storia di attestazioni e documenti. Dai territori in cui si parlano queste varietà sono partiti, in diverse tornate milioni di emigranti, trasferitisi in ogni parte del mondo, entrando in contatto, in particolare, con altre varietà romanze.

La selezione è caduta, in origine, su 8 varietà italiane (veneto, piemontese, napoletano, abruzzese, palermitano, salentino, fiorentino, e senese), parlate da emigranti italiani nelle Americhe, entrate dunque in contatto con lo spagnolo (in Argentina), il portoghese (in Brasile), il francese (in Québec), e poi anche

² Se due lingue sono geneticamente correlate, ci aspettiamo che esse sviluppino strutture simili, indipendentemente dal loro essere in contatto oppure no. Se lingue geneticamente correlate entrano in contatto, è praticamente impossibile dimostrare che un loro tratto comune sia indotto dal contatto e non semplicemente il risultato della deriva linguistica descritta da Sapir?.

l'inglese (a New York), considerata come varietà di controllo. Le condizioni storico-sociali in cui tali lingue sono entrate in contatto sono pressoché identiche: si tratta di emigranti che hanno lasciato l'Italia per trasferirsi definitivamente nel Nuovo Mondo, quindi desiderosi e necessitanti di apprendere la lingua del Paese ospite. Si osservi, incidentalmente, che il modello *Microcontact* per questa ricerca è *1:molte* ripetuto per ogni varietà italiana; si potrebbe definirlo un modello del tipo *molte: molte*.

Ritourneremo brevemente sul profilo delle famiglie di emigranti italiani del secondo dopoguerra nel prossimo paragrafo. Per il momento soffermiamoci sulle varietà selezionate e sulla logica sottostante la loro selezione.

1.1. Il problema della "baseline"

Per comprendere e descrivere adeguatamente il cambiamento linguistico è fondamentale conoscere l'input ricevuto dai parlanti; ciò diventa indispensabile se non si può partire avendo come lingua di riferimento una varietà standardizzata. Questo problema è conosciuto come "problema della *baseline*", cioè problema nella scelta della grammatica di riferimento, rispetto alla quale quantificare il cambiamento. Spesso non è possibile accertare se ciò che appare come una deviazione rispetto alla varietà di riferimento sia effettivamente il risultato del cambiamento linguistico o piuttosto la continuazione di un tratto già presente nell'input per effetto della grande variazione dialettale presente sul territorio italiano. Il problema inverso è anch'esso non trascurabile: se un tratto presenta un valore diverso nella lingua ereditaria rispetto alla corrispondente varietà parlata in Italia, ciò può significare certamente che il tratto sia cambiato per effetto del contatto, ma potrebbe anche darsi che la varietà parlata oggi in Italia sia cambiata rispetto a quella parlata dagli emigranti al momento del trasferimento all'estero, parlata che costituisce l'input per i parlanti ereditari. In altre parole: ciò che appare come un cambiamento potrebbe essere già stato presente nell'input per effetto della microvariazione, o potrebbe essere il risultato di una deviazione dalla forma contemporanea della varietà ma in linea con la variante originaria oggi scomparsa.

Riuscire a trovare il bandolo della matassa nell'enorme variazione empirica davanti a cui ci si trova manipolando dati di questo tipo non è impresa semplice, e spesso è necessario procedere per supposizioni. La ricerca ha richiesto l'applicazione di metodi di ricerca diversi: la consultazione delle grammatiche, il metodo storico-comparativo, l'indagine sulla lingua dei parlanti di prima generazione per descrivere l'input ricevuto dai parlanti ereditari, e la consultazione di letteratura scientifica sulle lingue ereditarie.

Per la scelta della varietà di riferimento si sono cercate grammatiche delle varietà locali risalenti al periodo della partenza degli emigranti. Ovviamente non è stato possibile rinvenire grammatiche di tutte le varietà locali, per cui si è cercata la grammatica dell'area più vicina alla località dei parlanti. Nel caso in cui fossero note le caratteristiche sintattiche di due o più varietà contigue con quella documentata, si è proceduto con il metodo "storico-comparativo", incrociando gli esiti contemporanei con le attestazioni precedenti, e confrontando questi dati con le varietà in contatto. Se una forma si fosse trovata in due o più varietà contigue a quella considerata e anche nella varietà ereditaria, la conclusione era che molto probabilmente non ci fosse stato nessun cambiamento e la forma documentata fosse presente già nella varietà degli emigranti.

Il metodo "grammaticale" è consistito, come si è spiegato poc'anzi, nell'identificazione di aree grammaticali identiche tra loro nelle varie lingue, con l'eccezione di un tratto. Le aree sintattiche individuate sono il Sintagma Verbale (SV), il Sintagma Temporale (ST), e il Sintagma Determinante (SD). I fenomeni selezionati sono: l'accusativo preposizionale e la transitività (per il SV); l'accordo del soggetto, i clitici soggetto e il soggetto nullo (per il ST); i pronomi e gli aggettivi dimostrativi e la deissi (per il SD). Per quanto riguarda questi fenomeni, tutte le varietà romanze sono tipologicamente uniformi (si veda la classificazione tipologica delle lingue romanze secondo LEDGEWAY 2012). Per fare un esempio, uno dei fenomeni sintattici investigati è l'uso del soggetto nullo, permesso (in determinate condizioni contestuali) dall'italiano e dallo spagnolo e da tutte le varietà italo-romanze (comprese quelle con i clitici soggetto; si faccia riferimento a Rizzi 1986 per una discussione dettagliata sulla natura dei soggetti clitici romanzi); permesso parzialmente in portoghese brasiliano; vietato in francese, e così via. La domanda a cui abbiamo inteso rispondere è: cosa accade al soggetto nullo? Cosa accade ai clitici soggetto in contatto, specialmente

a quelli non pronominali delle varietà settentrionali? Cosa ci dicono questi dati sulla natura dei clitici e della distribuzione del soggetto nullo? Riguardo a questo tema, si vedano la tesi di dottorato di Alberto Frasson e altri lavori (FRASSON 2021; 2022; IN STAMPA; FRASSON / D'ALESSANDRO / VAN OSCH 2021)

I risultati della ricerca sono stati pubblicati in diversi articoli. In questa sede non ne parleremo, come già detto.

Prima di passare alla situazione delle comunità italiane nelle Americhe è bene precisare che le varietà elencate nel progetto non sono state tutte documentate, e che ad alcune se ne sono sostituite altre. In particolare, è stato impossibile rintracciare parlanti di toscano; i parlanti di salentino incontrati erano troppo pochi perché i dati sulla loro varietà potessero essere minimamente attendibili, per cui il salentino non è stato considerato. Alle varietà originarie si sono aggiunti invece il calabrese e il friulano, molto presenti nelle aree di investigazione.

Durante la prima ricerca sul campo si sono intervistati 74 parlanti in Argentina, 50 in Brasile, 36 in Canada, e 58 a New York. Ritourneremo su questi dati nel prossimo paragrafo. La seconda ricerca sul campo, pianificata per la primavera 2020, è stata sostituita da un'inchiesta online, di portata molto modesta, consistita in interviste ed elicitazioni con 61 parlanti in Argentina, 44 in Brasile, 4 a New York, e 3 in Belgio (Paese che ha sostituito il Canada come target, per ragioni che saranno spiegate nel paragrafo seguente). In Italia sono stati intervistati online 228 parlanti. I più, 389 italiani hanno risposto online a un questionario sull'accusativo preposizionale e 788 a vari questionari sui soggetti clitici.

3. *Le comunità italiane nelle Americhe*

Il progetto *Microcontact* prevedeva, in origine, due momenti di raccolta di dati sul campo: uno pianificato per la primavera del 2019, uno per la primavera del 2020. Durante la prima indagine si intendeva stabilire un contatto con le comunità linguistiche degli emigrati nelle Americhe e documentare il loro parlato, limitandosi a un'inchiesta sintattica superficiale. Per la verifica delle ipotesi e un'inchiesta più approfondita era previsto un soggiorno di diversi mesi nella primavera/estate del 2020. L'insorgenza della pandemia da COVID-19 ha costretto i ricercatori a cambiare i piani. Questo paragrafo tratterà della prima raccolta, dell'individuazione dei parlanti, e del profilo delle comunità incontrate sul campo.

Il profilo linguistico dei parlanti di nostro interesse era ben chiaro: si volevano individuare parlanti di prima generazione (persone nate in Italia e trasferitesi nelle Americhe negli anni '50-'60 del secolo scorso), i loro figli e i loro nipoti (parlanti ereditari). La selezione di almeno tre generazioni è dovuta al fatto che è necessario controllare l'input ricevuto dai parlanti ereditari di seconda o terza generazione, come spiegato in precedenza.

Individuare queste comunità e questi parlanti non è stata un'impresa semplice. Il primo passo, previsto dal progetto, prevedeva l'utilizzo di un atlante interattivo, dove i figli e i nipoti degli emigrati potessero caricare delle brevi registrazioni del parlato dei loro nonni o genitori, o di sé stessi. Insieme a una documentazione delle varie storie avremmo così potuto individuare dei parlanti con il profilo giusto, da intervistare successivamente in fieldwork. L'atlante interattivo è consultabile a questo indirizzo (<https://microcontact.hum.uu.nl/#home>)³; comprende un breve questionario che serve a delineare il profilo sociolinguistico del parlante e una breve registrazione audio, della durata di 5 minuti massimo, che i giovani della comunità, avvezzi alla tecnologia, dovevano caricare sul sito.

La struttura dell'atlante è piuttosto semplice, essendo stato disegnato tenendo presenti tutti i criteri che potessero renderlo facile da usare a chiunque fosse capace di navigare su internet. L'atlante è in cinque lingue: italiano, inglese, spagnolo, portoghese e francese (la lingua viene selezionata automaticamente a seconda della zona del mondo in cui il sito viene aperto. La lingua di default è l'inglese). Inizialmente si sono contattati tutti gli Istituti Italiani di Cultura, le ambasciate, le associazioni, i conoscenti: chiunque potesse aiutarci a diffondere la richiesta di queste registrazioni, in Argentina, Brasile, Canada, Stati Uniti, e anche Italia. L'atlante raccoglie al momento 340 registrazioni. Sfortunatamente le uniche registrazioni

³ Il software dell'atlante interattivo *Microcontact* è in open source, scaricabile dal sito <https://github.com/UUDigitalHumanitieslab/microcontact>.

sono arrivate dall'Italia, mentre il tentativo di coinvolgere i parlanti ereditari nella documentazione delle proprie parlate, a distanza e senza contatto diretto, è fallito su tutta la linea. Le registrazioni effettuate nelle Americhe sono frutto dell'inchiesta sul campo, che comprendeva una prima parte di parlato spontaneo nella varietà da documentare.

Il ruolo dell'Italia nella ricerca è stato piccolo ma importante: è ben noto che, nello stesso periodo in cui gli emigranti lasciavano l'Italia per l'America, gli italiani entravano in contatto con l'italiano, diventandone parlanti nativi. La condizione dei parlanti, fino agli anni '70 circa (secondo De Mauro fino all'introduzione della scuola media unificata), era quella di forte diglossia con una componente preponderante di lingua locale. L'italiano veniva usato pochissimo e solo in contesti formali.

Il target della nostra ricerca erano dunque gli italiani, parlanti nativi delle varietà elencate, con livello di istruzione bassissimo o nullo. Ciò che ci premeva evitare era l'influenza del terzo incomodo, l'italiano, che creasse una situazione di trilinguismo. I parlanti di prima generazione dovevano essere bilingui: varietà italo-romanza (a cui in corso d'opera si è aggiunto il friulano, retoromanzo) e lo spagnolo, il portoghese, il francese, e così via.

Le comunità di parlanti di origine italiana da noi contattate e visitate hanno caratteristiche molto diverse; ciascun Paese ha assistito a uno sviluppo diverso e ha testimoniato un uso ben distinto delle varietà italiane nelle diverse realtà. Qui di seguito si forniscono alcune informazioni sullo stato delle comunità incontrate nei Paesi ospiti. Un resoconto più dettagliato delle condizioni delle varie comunità linguistiche italo-romanze nelle Americhe è presentato in ANDRIANI *ET AL.* (2021A), a cui si rimanda per informazioni più precise.

3.1. *Le comunità italiane in Argentina*

L'emigrazione italiana in Argentina ha inizio tra il XVIII e il XIX secolo, e diventa fenomeno di massa durante la metà del secolo XIX. Fino alla I Guerra Mondiale si trovano ancora testimonianze di lavoratori temporanei, che si trattengono in Argentina per qualche anno per poi far ritorno in Italia, o anche stagionali, che si recano in Argentina a lavorare in inverno come braccianti per rientrare in Italia per la bella stagione. I protagonisti di queste prime ondate migratorie sono di provenienza per lo più settentrionale. L'emigrazione meridionale cresce invece fino a diventare imponente tra le due guerre, e nel secondo dopoguerra (FERRARI 2008; MAURIZIO 2008).

La mescolanza tra italiani e argentini dà origine a varietà ibride, come il famoso *cocoliche* (Bagna 2011). In alcuni casi si formano comunità omogenee, di emigranti della stessa area di provenienza; ciò avviene prevalentemente nelle città di Buenos Aires, Córdoba e Santa Fe. Sono inoltre documentate colonie linguistiche di genovese, nel distretto di Buenos Aires chiamato Boca, e di friulano, in Colonia Caroya, completamente scomparse negli anni '80 (MEO ZILIO 1990).

Nonostante queste premesse e la presenza massiccia di famiglie di origine italiana, raggiungere le comunità italiane in Argentina si è rivelato molto difficile. Il contatto iniziale, che nelle nostre intenzioni sarebbe dovuto avvenire attraverso l'atlante, non c'è stato. Il ricercatore del gruppo deputato alla raccolta dei dati in Argentina, dopo aver provato innumerevoli volte e senza esito a contattare associazioni di italiani e privati, si è recato a Buenos Aires per stabilire i primi contatti di persona, scoprendo che le varietà italiane erano state quasi interamente abbandonate già dalla prima generazione di parlanti, che ha trasmesso la propria lingua madre ai discendenti in misura molto modesta. Questa fase di pre-inchiesta sul campo ha fornito un contributo importante nello stabilire i rapporti con le comunità locali.

La situazione linguistica degli italiani in Argentina è abbastanza uniforme: le varietà sono quasi integralmente estinte, sostituite dallo spagnolo. I parlanti ereditari sono spesso monolingui di spagnolo, sebbene alcuni di essi abbiano seguito corsi di italiano (il che, nel nostro caso, ha comportato una difficoltà aggiuntiva nell'indagine linguistica). In generale, i rapporti con l'Italia sono stati mantenuti nella maggior parte dei casi, ma le lingue ereditarie sono sempre meno parlate.

3.2. *Le comunità italiane in Brasile*

La situazione del Brasile è molto simile a quella dell'Argentina per quanto riguarda le ondate migratorie, con la differenza che l'ondata migratoria che ha interessato l'Italia dopo la Seconda guerra mondiale ha

raramente investito il Brasile. Gli italo-brasiliani, presenti diffusamente sul territorio soprattutto negli stati meridionali di San Paulo e Rio Grande do Sul (CENNI 2003), sono quasi tutti di seconda o terza generazione, mentre gli emigranti di prima generazione sono molto pochi. A differenza che in Argentina, tuttavia, le comunità italiane sono ancora presenti sul territorio brasiliano, soprattutto quelle dei parlanti del cosiddetto *taliàn*, una varietà romanza a base veneto-friulana ancora usata in molte comunità.

Intorno agli anni '40 del secolo scorso una legge introdotta dal governo ha proibito l'uso di lingue diverse dal portoghese: ciò ha condotto al passaggio forzato al portoghese come lingua di comunicazione, infliggendo un grave colpo alle minoranze linguistiche e alle lingue ereditarie italiane. Ciò ha spesso comportato un "salto di generazione", per cui i nonni hanno trasferito la loro lingua ai nipoti, saltando la generazione intermedia, che non ha potuto impararla dai propri genitori.

La politica linguistica proibizionistica brasiliana ha avuto un effetto maggiore nei grandi centri, più sottoposti al controllo rispetto alle aree montuose. Il *paulistano*, la varietà a base italo-romanza meridionale parlata a Saõ Paulo, ancora documentata al volgare del XIX secolo, è ormai definitivamente estinto (CENNI 2003). Le varietà settentrionali, ancoratesi invece a territori più isolati e montuosi, sono riuscite a sopravvivere fino ai giorni nostri. Il veneto e il friulano sono ancora parlati ampiamente a Porto Alegre e nei suoi dintorni. Queste comunità sono state quelle con cui abbiamo avuto maggiori contatti; i parlanti di prima generazione che siamo riusciti a intervistare sono stati solo 7; la maggior parte degli intervistati (42) apparteneva alla terza o quarta generazione.

3.3. *Le comunità italiane in Canada*

Come per gli altri Paesi di emigrazione, il Canada, e in particolare il Québec, era stato selezionato perché abbastanza distante da non permettere, nel secolo scorso, contatti continui con l'Italia. Già prima di partire le aspettative sui parlanti ereditari italiani in Canada erano molto diverse rispetto a quelle che riguardavano l'America Latina, ma anche New York. L'emigrazione italiana in Canada è registrata già dal primo quarto del secolo scorso, ma è di misura molto minore rispetto a quella verso gli altri Paesi considerati dal progetto. In più, molta di questa emigrazione ha riguardato le grandi città, in particolare Toronto, ed è stata, per il Québec, soltanto marginale (REINKE 2014). La più intensa ondata migratoria verso il Canada è avvenuta tra il 1950 e il 1970 circa; il profilo linguistico dei parlanti italo-romanzi era, a quel punto, profondamente diverso rispetto a quello dei loro genitori: molti emigranti avevano completato le scuole in Italia, avevano avuto una esposizione più sistematica all'italiano.

Questo tipo di profilo non corrispondeva, ci si è subito resi conto, a ciò che cercavamo; nell'intenzione del progetto c'era l'uniformità sociolinguistica; pur essendo perfettamente consapevoli che le situazioni e i contesti linguistici sono tutti diversi, l'idea era più o meno quella di selezionare parlanti a bassissimo livello di istruzione, monolingui della varietà italo-romanza, che fossero stati esposti all'italiano molto marginalmente. Nessuno dei parlanti canadesi del Québec corrispondeva a questo profilo: essi erano per lo più scolarizzati; quasi tutti mandavano i propri figli a scuola di italiano. L'italiano L3 si è rivelato essere davvero troppo pervasivo perché i dati raccolti in Canada potessero avere alcuna rilevanza statistica. Inoltre, l'inglese era la seconda, se non la prima, lingua di tutti. Impossibile escluderlo dal quadro linguistico dei parlanti. Ciò significa che nel migliore dei casi i parlanti di varietà italo-romanza conoscevano almeno anche l'inglese, oltre al francese e spesso all'italiano, o all'*italianese* (VILLATA 2010), una *koinè* a base italiana molto simile a quella parlata a New York: un profilo completamente diverso e non comparabile con quello degli emigranti negli altri Paesi.

Il profilo linguistico più adatto allo studio del contatto con il francese si è rivelato essere, a una prima indagine, quello degli emigrati in Belgio, che era stato originariamente escluso per via del fatto che molti lavoratori erano stagionali, che i contatti con l'Italia erano molto più semplici da intrattenere anche nel secolo scorso per via della facilità di spostarsi dall'uno all'altro Paese. In realtà molti gruppi di emigranti stabilitesi in Belgio, e in particolare nella Vallonia, dopo la Seconda Guerra Mondiale, per lavorare nelle miniere, sono rimasti molto spesso isolati, e in contatto con l'unica altra lingua a cui erano esposti: il francese. Gli emigrati in Belgio sembravano dunque avere il profilo adatto per informare la nostra ricerca; la pandemia ha impedito che la ricerca sul campo prevista in Belgio avesse luogo; i dati sul microcontatto con il francese sono dunque i meno affidabili e i più limitati.

3.4. Le comunità italiane a New York

La comunità italiana a New York è tra le più studiate nel contesto degli *heritage studies*. Da un censimento effettuato negli Stati Uniti nel 2000 è emerso che circa un milione di persone ha dichiarato di parlare italiano. Non è chiaro se si tratti solo di parlanti ereditari, ma è certo che l'emigrazione italiana verso le grandi città degli USA è stata continua, a partire dalla seconda metà del secolo XIX: secondo l'ISTAT il numero di italiani emigrati negli USA nel 1880 era di 81249. Il trend di crescita dell'uso della lingua italiana da parte dei parlanti ereditari sembra però andare scemando negli ultimi anni: secondo l'American Community Survey l'uso dell'italiano è diminuito di un terzo negli ultimi 10 anni (per informazioni più dettagliate si vedano ANDRIANI *ET AL.* 2021A; CAVAIOLI 2008).

L'assetto linguistico dell'italiano negli USA è molto diverso rispetto a quello documentato negli altri Paesi; le differenze principali sono due: innanzitutto, le comunità "miste", di italiani provenienti da molte parti d'Italia, si sono formate praticamente subito, lasciando poco spazio alla sopravvivenza delle varietà locali. La convivenza, all'interno della stessa comunità, di varietà provenienti da tutto il territorio italiano ha portato alla formazione di una *koinè* sempre più tendente all'italiano dopo l'introduzione, nel 1965, della *Immigration Reform Law* che permise alle famiglie italiane di vivere e lavorare legalmente negli USA. La *koinè* italiana a New York prese il nome di *broccolino*, una varietà a base italo-romanza molto diversa da ciascuna delle varietà documentate in Argentina o Brasile, molto più uniformi rispetto a questa nuova varietà ibrida (HALLER 1987; 1991; 1997A; 1997B; 2002; DE FINA & FELLIN 2010). La situazione newyorchese è molto simile a quella riscontrata in altre metropoli, come Sydney o Toronto.

La seconda differenza riguarda la scolarizzazione in italiano: essendo il flusso migratorio negli USA continuo, le nuove comunità di *expat* sono andate a sovrapporsi alle comunità preesistenti. Sebbene il contatto tra di esse non sia regolare, si sono create moltissime scuole in lingua italiana, frequentate in egual misura dai figli di emigranti recenti e non. L'italiano ha informato questa *koinè* molto più di quanto l'abbia fatto con le varietà parlate negli altri Paesi. Sulla base di questi fatti è stato necessario, nella ricerca, considerare non solo il gruppo in contatto con la lingua inglese come gruppo di controllo, controparte del microcontatto ed esponente di macrocontatto, ma anche considerare che le varietà parlate sono completamente diverse, molto più uniformi, rispetto alla frammentazione trovata nel microcontatto.

4. La raccolta dei dati

Progettare una raccolta di dati extracontinentale richiede uno sforzo notevole. In questo paragrafo si parlerà delle norme burocratiche che è necessario rispettare durante un'inchiesta sul campo che abbia luogo fuori dall'Europa. Il prossimo paragrafo sarà invece dedicato ad alcune considerazioni pratiche sull'organizzazione di un'inchiesta linguistica che abbia come partecipanti persone anziane e poco scolarizzate che parlino una lingua diversa da quella dell'intervistatore.

Nel maggio del 2018 è entrato ufficialmente in vigore in tutta Europa il GDPR (*General Data Protection Protocol*, GDPR 2016; 2018), il protocollo per la protezione dei dati sensibili. L'implementazione del GDPR è stata lasciata ai singoli stati. Per la ricerca scientifica, ciascuna università si regola autonomamente. Al momento non esiste nei Paesi Bassi un modello nazionale condiviso per l'implementazione delle regole del GDPR; ciascuna università si è dotata di un *data protection officer* (DPO) che ha il compito di controllare che i protocolli per la raccolta e la conservazione dei dati siano conformi con il GDPR. Cosa si intenda per "conformità con il GDPR" è però lasciato all'interpretazione di ciascun DPO. Per i progetti finanziati dall'ERC, la responsabilità della corretta implementazione delle regole del GDPR è attribuita alle università ospitanti; la violazione delle regole comporta sanzioni; ciò ha portato a un inasprimento delle misure di controllo che va spesso contro i fini della ricerca stessa.

Uno degli articoli più importanti per la ricerca linguistica contenuti nel GDPR è l'articolo 4, che definisce «dato personale»:

qualsiasi informazione riguardante una persona fisica identificata o identificabile («interessato»); si considera identificabile la persona fisica che può essere identificata, direttamente o indirettamente, con particolare riferimento a un identificativo come

il nome, un numero di identificazione, dati relativi all'ubicazione, un identificativo online o a uno o più elementi caratteristici della sua identità fisica, fisiologica, genetica, psichica, economica, culturale o sociale; (art. 4 sub 1 GDPR 2018).

Ciò significa che il parlante, o chi fornisce i dati, non deve essere riconoscibile in alcun modo. Si potrebbe pensare che per proteggere i dati personali possa bastare rendere anonimi i questionari, ma non è purtroppo così semplice, poiché i parlanti sono i “proprietari” dei dati, e deve essere loro permesso ritirarne la concessione all'utilizzo, in qualsiasi momento. In altre parole, se un(a) parlante firma un consenso informato accordando al ricercatore il permesso di usare i dati da lui o da lei prodotti, lo fa sapendo di poter ritirare il consenso quando vuole. Ciò significa, nella pratica, che il ricercatore deve sapere a chi appartengono i dati; deve avere una lista, non pubblica, di corrispondenza tra i dati e i parlanti che li hanno prodotti (nel *data protection protocol* è necessario specificare come e dove sarà conservata questa lista, se in un armadio chiuso con il lucchetto, se in un server protetto dell'università, eccetera, e come questa lista verrà trasferita dal ricercatore sul campo al direttore della ricerca. Non è possibile, in nessun momento, utilizzare piattaforme tipo Dropbox per contenere questi dati. Idealmente, essi devono restare sul pc del fieldworker per poi essere consegnati a mano al direttore della ricerca. Se c'è il rischio che il pc venga rubato, come nel nostro caso, è necessario connettersi in remoto a un server protetto, a cui solitamente non è molto facile accedere soprattutto con una connessione non troppo stabile.

Inoltre, i dati devono essere controllabili, e in teoria deve essere possibile per un ricercatore A recarsi dal parlante B a controllare che abbia effettivamente detto quelle cose. Sebbene questo sia uno scenario avveniristico per la dialettologia italiana, il GDPR comporta di fatto che: 1. i parlanti diano il consenso informato per l'uso dei loro dati; 2. il *data protection officer* (DPO) controlli che tutti i dati siano pseudonimizzati (o anonimizzati), e 3. il ricercatore conservi una lista di corrispondenze tra i dati e i parlanti.

Nella raccolta dei dati in territorio extraeuropeo, l'impresa diviene quasi impossibile. L'ERC richiede che i Paesi da cui si esportano i dati diano il permesso per farlo; che ci sia un DPO locale che controlli che i dati siano manipolati nel modo corretto secondo le legislazioni del Paese ospitante e *secondo il GDPR*. Al momento, tra gli stati da cui è permesso esportare dati direttamente compaiono l'Argentina e il Canada (“Third Countries”, GDPR 2018); gli USA hanno stipulato altri accordi. Il Brasile non ha accordi. Ciò complica il quadro non poco, poiché non tutti i Paesi hanno università con DPO; è molto importante comprendere che il GDPR rende molto difficile la cooperazione con Paesi in via di sviluppo per via delle richieste a cui non tutti sono in grado di rispondere adeguatamente.

Un articolo del GDPR che molti trascurano ma che invece può essere di grande aiuto è l'articolo 6, comma f, che riguarda l'interesse legittimo, e recita:

f) il trattamento [dei dati, n.d.r.] è necessario per il perseguimento del legittimo interesse del titolare del trattamento [il DPO] o di terzi [DPO nel Paese della ricerca, ad esempio], **a condizione che non prevalgano gli interessi o i diritti e le libertà fondamentali dell'interessato** che richiedono la protezione dei dati personali, in particolare se l'interessato è un minore.

È il caso di aggiungere, a questo punto della narrazione, che molti parlanti ereditari non hanno alcun piacere a risultare anonimi. È spesso capitato ai ricercatori di ricevere delle rimostranze quando i parlanti apprendevano che i loro nomi non sarebbero comparsi nella ricerca, fatto da loro intepretato come una cancellazione della loro identità.

È molto importante anche riflettere sull'effettiva possibilità di anonimizzare dati linguistici. La voce è unica, proprio come le impronte digitali (testimonianza ne siano i vari tentativi di creazione di dispositivi per il riconoscimento vocale, *voive biometric technology*). Sebbene il dibattito scientifico sull'unicità della voce sia ancora in corso, certamente la voce di un parlante è altrettanto riconoscibile quanto il suo volto. Ci si chiede quanto si utile nascondere il nome dei parlanti quando si può diffondere la loro voce.

Un problema ulteriore creato dall'applicazione del GDPR riguarda il consenso informato, secondo il quale i partecipanti devono capire esattamente come saranno usati i loro dati, a quale fine, e firmare un modulo per il consenso (dal quale apprenderanno anche di avere diritto a ritirarlo). Il primo ostacolo all'applicazione di quella che sembra una procedura standard è costituito dalla richiesta della firma: il

nostro target, come più volte ripetuto, era costituito almeno in parte da persone analfabete; chiedere a persone con scarsa dimestichezza con la lettura e con la scrittura di firmare un modulo è culturalmente inappropriato e molto offensivo. Si è pensato di ovviare a questo problema con il consenso orale, registrato (sempre nell'assunto che la voce sia unica e che dunque il parlante possa essere individuato, se necessario, e il consenso considerato valido). Per la nostra inchiesta sono stati utilizzati moduli di consenso informato in italiano, inglese, francese, spagnolo, e portoghese. La partecipazione di membri della comunità che aiutassero le persone anziane a comprendere il più possibile gli scopi della ricerca è stata fondamentale. A questo proposito si faccia riferimento a LEIVADA ET AL. (2019), che riporta nel dettaglio i problemi burocratici che possono insorgere nella raccolta dei dati con informatori anziani e scarsamente istruiti.

5. *La raccolta dei dati con parlanti anziani e poco istruiti*

Uno degli obiettivi del progetto *Microcontact* è stabilire l'evoluzione della lingua dalla prima generazione, quella degli emigrati nati in Italia, alla seconda e a quelle successive. I parlanti ereditari hanno appreso la lingua in condizioni diverse da quelle dei monolingui dei Paesi di provenienza di genitori e nonni, ovviamente, e per questo è stato fondamentale documentare l'input, cioè le lingue parlate dalla prima generazione di emigranti, che poi le hanno trasmesse ai loro figli.

Il profilo ideale dei parlanti di prima generazione, come si è già detto, è quello di persone emigrate dopo la guerra, con meno esposizione possibile all'italiano. Il fenomeno dell'italianizzazione degli italiani, descritto così bene da DE MAURO (1963; 2014; 2018), ha riguardato la generazione immediatamente successiva a quella del nostro gruppo target, costituito da italiani dialettofoni, monolingui, con bassissima competenza attiva dell'italiano.

Sebbene i fenomeni linguistici individuati siano di alta frequenza e quindi riscontrabili nel parlato spontaneo, è stato necessario anche elicitarne alcuni dati. Il primo ostacolo ha riguardato la presentazione degli stimoli. Dopo attenta consultazione della letteratura riguardante la ricerca sulle lingue ereditarie si è deciso di organizzare le interviste nel modo seguente. Prima di iniziare l'elicitazione si chiedeva al parlante di raccontare, nella lingua italo-romanza target, per dieci minuti o più, la storia del suo arrivo nel Paese ospite. Solitamente questa parte aveva luogo in forma di conversazione tra due o più parlanti della stessa comunità, in modo che i parlanti fossero a proprio agio e che la lingua dell'interlocutore non interferisse con quella dell'intervistato. Dopo questo breve racconto si passava all'HALA test (O'GRADY ET AL. 2009) che controlla il livello di fluency nella lingua target tramite semplicissimi test lessicali e grammaticali. Ad esempio, il parlante vede un paio di occhiali sullo schermo e deve dire cosa sono. Oltre a misurare l'ampiezza del lessico, l'HALA test misura il tempo di reazione allo stimolo: più lungo è il tempo di reazione, più difficoltà ha il parlante a recuperare il lessema. Ovviamente, trattandosi spesso di persone anziane, si è tenuto conto del fatto che i tempi di reazione fossero molto più lunghi di quelli degli altri parlanti. Si passava poi all'elicitazione dei dati con dei task costruiti all'uopo: un task di scelta forzata, uno di ripetizione, uno di completamento della frase, e dei task interattivi.

Si è già detto che molti parlanti di prima generazione sono analfabeti o scarsamente alfabetizzati: ciò significa che un questionario scritto era impossibile da somministrare, anche perché molte di queste varietà non hanno uno standard di scrittura. Si è così pensato di presentare degli stimoli uditivi, nella lingua target. Le istruzioni per completare i task e i task stessi sono stati pre-registrati in tutte le varietà italo-romanze target, e presentati agli ascoltatori tramite il computer.

Un esempio di task di scelta forzata è il seguente. Il parlante sentiva le istruzioni nella sua lingua (per esempio, il siciliano); un esempio di presentazione di uno stimolo, tradotto dal siciliano in italiano, è riportato in 0. Le frasi erano ovviamente randomizzate.

(1) "ascolterete ora due frasi. Diteci quale vi sembra la migliore delle due"

- A. Ho visto Anna
- B. Ho visto a Anna

Il parlante rispondeva “la prima”, o “la seconda”. Per essere certi che il parlante non tirasse a indovinare e non scegliesse le risposte a caso, oltre a incrociare i risultati con quelli ottenuti negli altri test, si sono controllati anche i dati di parlato spontaneo. Le registrazioni spontanee sono servite, dunque, oltre a far entrare i parlanti nella modalità “lingua madre” a controllare che le loro risposte ai test fossero effettivamente un riflesso della loro lingua. Nella maggior parte dei casi alla scelta della variante è corrisposta quella usata dal parlante nel discorso libero. È anche capitato che il parlante facesse una scelta a caso; per capire quale fosse effettivamente il suo giudizio, al parlante veniva anche richiesto di ripetere la frase; la ripetizione è stata uno dei task più utili, perché il parlante ripete sempre la frase nella maniera a lui/lei più naturale (POLINSKY 2018, AALBERSE / BACKUS / MUYSKEN 2019).

Il task di completamento delle frasi è stato usato per controllare l’uso degli ausiliari (se presenti nella lingua). Al parlante veniva chiesto di completare la frase secondo lo schema: “Oggi non faccio questo perché l’ho fatto ieri”. I test sui deittici/dimostrativi si sono invece rivelati molto più difficili da spiegare e da usare del previsto. Erano state preparate delle vignette come quelle rappresentate nella Figura 2. Molto spesso i nostri informatori non sapevano interpretarle. In particolare l’idea del fumetto si è rivelata una scelta disastrosa: molti parlanti non sono stati in grado di capire che cosa gli si stesse chiedendo di fare⁴.

Figura 2. I dimostrativi



I test sui dimostrativi sono stati dunque effettuati tutti con l’ausilio di oggetti (un orsetto di peluche, posizionato in diversi modi rispetto al parlante e all’ascoltatore). I risultati della ricerca sono presentati in diverse pubblicazioni (si vedano ANDRIANI ET AL. 2021A, B). Come già riportato in altri lavori, avere come target parlanti anziani richiede una serie di attenzioni che non sono soltanto di natura scientifica.

5.1. Selezione di stimoli appropriati all’età e alla cultura del parlante

I parlanti anziani e bassamente scolarizzati hanno poca dimestichezza con l’uso del computer; oltre a ciò, va fatta una scelta sul tipo di stimolo da presentare. Come si è già ripetuto, gli stimoli scritti sono inappropriati, per via della bassa capacità di affrontare testi scritti. Per gli stimoli sonori si sono incontrati altri tipi di problemi: molto spesso gli intervistati anziani hanno difficoltà di udito e di vista. È capitato all’intervistatore di dover ripetere, a voce più alta, ciò che era detto nello stimolo. Questo ha provocato però un’interferenza linguistica, essendo gli intervistatori non sempre di madrelingua nella varietà investigata.

Un altro ostacolo è derivato dal proporre alle persone anziane delle frasi che loro consideravano inaccettabili dal punto di vista culturale. Volendo studiare la transitività e il massimo grado di agentività del soggetto, tra i verbi che si erano proposti nel questionario preliminare c’erano *uccidere*, *baciare*, e *assassinare*. Molti dei nostri parlanti si sono rifiutati di rispondere alle domande riguardanti frasi come *Maria ha ucciso / io ho ucciso Massimo* e *Maria ha baciato Giovanni/Giovanna*. Nel primo caso, alcuni parlanti ci hanno esplicitamente detto di non voler accusare nessuno, e anche di non voler dichiarare, di fronte a

⁴ Come al solito, le traduzioni sono per noi; gli intervistati vedevano solo le vignette.

estranei, di aver ucciso qualcuno, nel timore che queste registrazioni fossero usate contro di loro. Nel secondo caso, i parlanti ci hanno detto di non voler fare pettegolezzi, di non voler dire che “Maria” (nome molto comune e spesso appartenente a persone di loro conoscenza) avesse baciato nessuno che non fosse suo marito, o che una donna non bacia una donna, un uomo non bacia un uomo, e così via. La soluzione è stata sostituire al nome proprio, nel complemento oggetto, un nome di parentela, come “suo figlio/sua nipote”. “Maria bacia sua figlia” è una frase completamente accettabile. Il test, però, ne è risultato lievemente modificato, in particolare per le varietà meridionali che marcano i nomi di parentela con un possessivo enclitico. La sintassi dei nomi di parentela e del possesso inalienabile può interferire con la marcatura differenziale dell’oggetto (IEMMOLO 2009). Altre volte ai nomi con referenti umani si sono sostituiti nomi di animali, ad esempio “Il lupo ha ucciso l’agnello”. Anche in questo caso il problema è che il grado di animatezza degli animali è inferiore a quello degli umani, e dunque i sintagmi erano a livelli diversi della scala di animatezza di SILVERSTEIN (1976). Sono stati anche aggiunti alcuni verbi, come *sognare* e *vedere*, che però non sono prototipicamente agentivi. Verbi come *lodare* sono stati scartati per via del loro registro troppo elevato.

È molto importante, dunque, prima di sottoporre un questionario a parlanti appartenenti a culture e età diverse rispetto ai ricercatori, fare attenzione a tutto ciò che possa turbare la loro suscettibilità. Nel nostro caso, alcuni parlanti anziani residenti in Italia hanno risposto alle domande in un piccolo progetto pilota. L’aiuto delle comunità dialettofone in Argentina e Brasile, soprattutto da parte dei membri più giovani della comunità, nella scelta e nella formulazione delle domande, è stato fondamentale.

La partecipazione di alcuni membri della comunità alle inchieste dialettologiche non è una metodologia utilizzata solo nella ricerca sulle lingue ereditarie (si veda, a questo proposito, CORNIPS / POLETTI 2005). La presenza di uno o più membri della comunità durante l’intervista è fondamentale per la riuscita dell’impresa. Innanzitutto, come ben sa chi abbia effettuato inchieste dialettologiche, la presenza di un intervistatore non di madrelingua induce immediatamente il parlante ad adattare la propria lingua a quella dell’interlocutore. Nel nostro caso, la lingua verso cui si deviava era quasi sempre l’italiano (sebbene stentato), essendo i nostri intervistatori italiani. In secondo luogo, un intervistatore è uno sconosciuto, e difficilmente un anziano ammetterà a casa propria una persona sconosciuta e straniera senza la presenza di almeno una persona conosciuta. Molte nostre interviste hanno avuto luogo in bar e posti affollati, proprio per questa comprensibile diffidenza verso gli intervistatori. Non essendo il nostro un progetto di fonetica/fonologia, la qualità audio è di importanza limitata, tuttavia è a volte difficile capire esattamente cosa l’intervistato stia dicendo, soprattutto se interloquisce con diverse persone e con un grande rumore di sottofondo.

5.2. Consigli pratici e norme di sicurezza

L’ultimo aspetto da tener presente quando si faccia ricerca sul campo in comunità come quelle appena descritte, che cioè non sono in località poverissime ma possono non avere lo stesso livello di comfort delle località a cui si è abituati, è la propria sopravvivenza (e quella della propria attrezzatura): è consigliabile portare sempre con sé del cibo e dell’acqua (a volte gli intervistati ci hanno sommersi di cibo, altre volte è capitato che un fieldworker restasse ad aspettare un intervistato in mezzo al nulla per una giornata intera); è bene portare con sé una batteria di riserva per il computer e per il registratore, perché non sempre sono disponibili prese elettriche.

Se il fieldworker deve recarsi in zone pericolose o completamente isolate, è consigliabile istituire un protocollo di sicurezza per il fieldwork prima della partenza. Molte università hanno dei protocolli stabiliti (si veda, ad esempio, https://www.eur.nl/sites/corporate/files/manual_security_guidelines_def_2.dd_juli_2016.pdf). Un accorgimento molto semplice è per esempio spedire ogni sera un messaggio a tutti i membri del progetto, o al direttore del progetto. È consigliato, inoltre, che chi si rechi in un Paese straniero dichiari la propria presenza al consolato o alla polizia locale, in modo da essere rintracciabile in ogni momento. È essenziale avere un protocollo di evacuazione, cioè la possibilità di un ricercatore sul campo di rientrare immediatamente nel paese d’origine nel caso ci siano problemi.

È necessario, inoltre, stipulare un'assicurazione che copra tutti i costi di eventuali cure o rimpatri, in caso di necessità. In un caso, uno dei nostri ricercatori ha subito uno scippo – è stato essenziale, in quel momento, che ci fosse la possibilità immediata di rientro, necessaria soprattutto per la tranquillità psicologica della persona scippata.

Per finire, è bene essere preparati psicologicamente per il fieldwork. Si tratta di un'esperienza unica e molto interessante, ma va considerato che essa richiede molta resistenza, capacità di adattamento, capacità di sopportazione e di reazione alle avversità, e anche, molto banalmente, resistenza allo stress che può essere causato dal turbamento dei cicli del sonno (per via del fuso orario) o della diversa alimentazione. Si sconsiglia vivamente il fieldwork in paesi remoti per chi non abbia già esperienza di viaggi e permanenza prolungata in paesi stranieri.

6. *Riflessioni conclusive*

Uno studio linguistico che abbia come target varietà non parlate in territorio europeo va progettato e programmato meticolosamente. Le difficoltà linguistiche, logistiche e burocratiche sono molte, spesso il lavoro non può procedere per motivi completamente slegati dalla ricerca che si sta svolgendo. Nonostante il lavoro di raccolta e descrizione dei dati delle varietà italo-romanze nelle Americhe sia costato molto di più, in termini di tempo e fatica, rispetto a una raccolta equivalente in Italia, i risultati sono stati molto soddisfacenti. Va sottolineato come queste varietà non erano state mai documentate prima in un'inchiesta linguistica sistematica. Oltre al valore culturale di tale documentazione va considerato che una situazione di multicontatto come quella descritta è molto difficile da riscontrare, e che essa offre informazioni utili alla comprensione dei meccanismi del cambiamento in contatto, e del linguaggio umano, che nessuna lingua, considerata in isolamento, può fornire. Come notato da D'Alessandro *et al.* 2021, le varietà ereditarie sono importanti per ciò che ci dicono sul cambiamento linguistico ma anche in sé, in quanto sistemi linguistici perfettamente completi e solo lievemente diversi dagli originali.

Bibliografia

- AALBERSE, S. / BACKUS, A. / MUYSKEN, P. 2019: *Heritage Languages. A language contact approach*, Amsterdam, John Benjamins.
- AIKHENVALD, A.Y. / DIXON, R.M.W (eds.) 2001: *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*, Oxford, Oxford University Press.
- AIKHENVALD, A.Y. / DIXON, R.M.W (eds.) 2007: *Grammars in Contact: A Cross-Linguistic Typology*, Oxford, Oxford University Press.
- ANDERSEN, H. 1988: *Centre and periphery: adoption, diffusion and spread*, in FISIÁK, J. (ed.), *Historical Dialectology*, Berlino, Mouton de Gruyter, pp. 39–85.
- ANDRIANI, L. / D'ALESSANDRO, R. / FRASSON, A. / VAN OSCH, B. / SORGINI, L. / TERENGI, S. 2022: *Adding the microdimension to the study of language change in contact. Three case studies*, in «Glossa: a journal of general linguistics», VII.1.
- ANDRIANI, L. / CASALICCHIO, J. / CICONTE, F.M. / D'ALESSANDRO, R. / FRASSON, A. / VAN OSCH, B. / SORGINI, L. / TERENGI, S. 2022: *Documenting Italo-Romance minority languages in the Americas. Problems and tentative solutions*, in NEVINS, A. / COLE, M. (eds.), *Contemporary research in minority and diaspora languages of Europe*, Berlino, Language Science Press.
- BAGNA, C. 2011: *America Latina*, in VEDOVELLI, M. (ed.), *Storia linguistica dell'emigrazione italiana nel mondo*, Roma, Carocci, pp. 305–357.
- CAVAIOLI, F. J. 2008: *Patterns of Italian immigration to the United States*, in «The Catholic Social Science Review» XIII, pp. 213–229.
- CENNI, F. 2003: *Italianos no Brasil: “andiamo in 'Merica”*, San Paolo, EdUSP.
- CORNIPS, L. / POLETTI, C. 2005: *On standardising syntactic elicitation techniques*, in «Lingua» CXV, pp. 939–957.

- D'ALESSANDRO, R. 2021: *Syntactic change in contact: Romance*, in «Annual Review of Linguistics» VII.7, pp. 309-328.
- D'ALESSANDRO, R. 2015: *Microcontact. Language Variation and Change from the Italian Heritage perspective*, ERC-CoG project.
- D'ALESSANDRO, R. / NATVIG, D. / PUTNAM, M.T. 2021: *Addressing challenges in formal research on moribund heritage languages: A path forward*, in «Frontiers in Psychology» XII, 700126.
- DE MAURO, T. 1963: *Storia linguistica dell'Italia unita*, Biblioteca di cultura moderna, 585, Bari, Laterza.
- DE MAURO, T. 2014: *Storia linguistica dell'Italia repubblicana dal 1946 ai nostri giorni*, Bari, Laterza.
- DE MAURO, T. 2018. *L'educazione linguistica democratica*, Bari, Laterza.
- FRASSON, A. 2021: *Clitics are not enough: on agreement and null subjects in Brazilian Venetan*, in «Glossa: a journal of general linguistics », VI.1.
- FRASSON, A. 2022a: *The syntax of subject pronouns in heritage languages: innovation and complexification*, Tesi di dottorato, Università di Utrecht.
- FRASSON, A. 2022b. in stampa. *Structural Change in Relative Clauses and the Autonomy of Heritage Grammars*, in Guardiano, C. / Silvestri, G. (eds.), *Special Issue on Syntactic Variation and Change of Heritage Languages*, Languages.
- Frasson, A. / D'Alessandro, R / B. van Osch. 2021: *Subject clitics in microcontact. A case study from heritage Friulian in Argentina and Brazil*, in «Heritage Language Journal», XVIII.1, pp. 1-36.
- GDPR 2016: *EU General Data Protection Regulation (GDPR): Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)*, OJ 2016 L 119/1.
- GDPR 2018: *General Data Protection Regulation (GDPR)*. General Data Protection Regulation (GDPR) – Final text neatly arranged. Available at: <<https://gdpr-info.eu/>>
- HALLER, H. W. 1987: *Italian speech varieties in the United States and the Italian-American lingua franca*, in «Italia» LXIV.3, pp. 393–409.
- HALLER, H. W. 1991: *Atteggiamenti linguistici nelle comunità italo-americane*, in «Rivista di Linguistica» III.2, pp. 389–405.
- HALLER, H. W. 1993: *Una lingua perduta e ritrovata: l'italiano degli italoamericani*, Firenze / Roma, La Nuova Italia.
- HALLER, H. W. 1997a: *Dialects abroad*, in MAIDEN, M / PARRY, M. (eds.), *The Dialects of Italy*, Londra, Routledge, pp. 401–411.
- HALLER, H. W. 1997b: *Italian*, in FISHMAN, J.A. / GARCÍA, O. (eds.), *The Multilingual Apple. Languages in New York City*, Berlino - New York, Mouton De Gruyter, pp. 119–142.
- HALLER, H. W. 2002: *I dialetti italiani negli USA*, in CORTELAZZO, M. / CLIVIO, G. / DE BLASI, N. / MARCATO, C. (eds.), *Dialetti italiani: Storia, struttura, usi*, Torino, UTET, pp. 1083–1086.
- HEINE, B. / KUTEVA, T. 2005: *Language Contact and Grammatical Change*, Cambridge, Cambridge University Press.
- IEMMOLO, G. 2009: *La marcatura differenziale dell'oggetto in siciliano antico*, in «Archivio Glottologico Italiano» XCIV.2, pp. 185-225.
- KEENAN, E., 1994: *Creating Anaphors: An Historical Study of the English Reflexive Pronouns*, ms., University of California at Los Angeles.
- KEENAN, E., 2002: *Explaining the creation of reflexive pronouns in English*, in: MINKOVA, D. / STOCKWELL, R. (eds.), *Studies in the History of the English Language: A Millennial Perspective*. Berlino - New York, Mouton de Gruyter, pp. 325–354.
- LEDGEWAY A. 2012. *From Latin to Romance: Morphosyntactic Typology and Change*. Oxford, Oxford University Press.
- LEIVADA, E. / D'ALESSANDRO, R / GROHMANN, K. 2019: *Eliciting big data from small, young, or non-standard languages: 10 experimental challenges*, in «Frontiers in Psychology» X, 313.

- LONGOBARDI, G. 2001: *Formal syntax, diachronic Minimalism, and etymology: the history of French chez*, in «Linguistic Inquiry» XXXII, pp. 275-302.
- Maurizio, R. 2008: *Migración y desarrollo: el caso de Argentina*, in SOLIMANO, A. (ed.), *Migraciones internacionales en América Latina. Booms, crisis y desarrollo*, Mexico, Fondo de Cultura Económica, pp. 75–184.
- MEO ZILIO, G. 1990: *Estudios hispanoamericanos*, Vol. 1: Temas lingüísticos, Roma, Bulzoni.
- O'GRADY, W. / SCHAFER, A.J. / PERLA, J. / LEE, O.-S. / WIETING, J. 2009: *A psycholinguistic tool for the assessment of language loss: The HALA Project*, in «Language Documentation & Conservation» III, pp. 100–112.
- POLINSKY, M. 2018a: *Heritage Languages and their Speakers*, Cambridge, Cambridge University Press.
- REINKE, K. 2014: *Language contact in a multilingual setting. The attractive force of Italo-Romance dialects on Italian in Montreal*, in BESTERS-DILGER, J. / DERMARKAR, C. / PFÄNDER, S. / RABUS, A. (eds.), *Congruence in Contact-induced Language Change*, Berlino, De Gruyter, pp. 149–167.
- ROSS M. 2001: *Contact-induced change in Oceanic languages in north-west Melanesia*, in AIKHENVALD, A.Y. / DIXON, R.M.W (eds.), *Areal Diffusion and Genetic Inheritance*, Oxford, Oxford University Press, pp. 134–66.
- ROTHMAN, J. 2009: *Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages*, in «International Journal of Bilingualism» XIII.2, pp. 155–163.
- ROTHMAN, J. 2010. *On the typological economy of syntactic transfer: Word order and relative clause attachment preference in L3 Brazilian Portuguese*, in «International Review of Applied Linguistics» XLVIII.2–3, pp. 245–273.
- ROTHMAN, J. 2011. *L3 syntactic transfer selectivity and typological determinacy: The typological primacy model*, in «Second Language Research» XXVII.1, pp. 107–127.
- ROTHMAN, J. / Cabrelli Amaro, J. / De Bot, K. 2011. *Third language acquisition*. Cambridge, Cambridge University Press.
- ROTHMAN, J. / GONZÁLEZ ALONSO, J. / PUIG-MAYENCO, E. 2019: *Third Language Acquisition and Linguistic Transfer*, Cambridge, Cambridge University Press.
- SORACE A. 2011: *Pinning down the concept of “interface” in bilingualism*, in «Linguistic Approaches to Bilingualism» I, pp. 1–33.
- SORACE A. 2012: *Pinning down the concept of interface in bilingualism. A reply to peer commentaries*, in «Linguistic Approaches to Bilingualism» II, pp. 209–216.
- THOMASON, S. G. / KAUFMAN, T. 1988: *Language Contact, Creolization, and Genetic Linguistics*, Berkeley - Los Angeles, University of California Press.
- VILLATA, B. 2010: *L'italianese – L'italiano comune parlato a Montreal*, Montreal, L'òsna & Tron.